

# Applied Research in Child Development

NUMBER 4, FALL 2003

In this issue

## 4

### Assessment

In their efforts to improve education, legislators and policymakers have embraced the norm-referenced, group-administered standardized test, encouraging its use to measure everything from student learning to program and teacher quality. Can norm-referenced instruments improve curriculum and instruction? Research suggests that performance-based assessment can improve teaching and learning and boost performance, but many questions remain unanswered. Moving beyond measurement to improving educational outcomes will require a clear understanding of the objectives, extensive teacher training, and the political will to try a different approach.

FROM THE DEAN → 2

Q & A: A CONVERSATION WITH SAM MEISELS → 9

CONSIDERING OPTIONS IN CHICAGO → 12

BUILDING MULTIDIRECTIONAL BRIDGES THROUGH CLASSROOM ASSESSMENT → 14

NEWS → 17

## Assessment in the early childhood classroom

ASKING THE RIGHT QUESTIONS,  
ACTING ON THE ANSWERS

BY SUZANNE L. WAGNER

Assessment, “the process of obtaining information for the purpose of making evaluative decisions” (Meisels, 1994), has been integral to American education since before the one-room schoolhouse. With the standardization and expansion of the educational system—and the concomitant need to decide who was to be educated, when, and how—came the proliferation of tests and measures to accomplish assessment. The more sophisticated our pedagogical knowledge and the more diverse the student population, the broader the curriculum needed to be and the larger the array of tools that were needed to inform decision making.

Researchers have commonly understood that assessments are linked to particular purposes. Each tool in the assessment kit is intended to collect specific data for a specific purpose. For example, in 1998, National Education Goals Panel (NEGP) articulated four primary purposes of early childhood assessment: (1) to support learning and instruction; (2) to identify children for additional services; (3) to evaluate

programs and monitor trends; and (4) to provide information for accountability.

Each of these purposes contributes to the overarching goal of improving educational outcomes. Yet, clearly, the data needed for one purpose may be irrelevant to another. Results from an achievement test, for example, cannot diagnose a child’s learning disability. Recognizing that fact, the panel further provided two basic principles for analyzing the suitability of measurement tools. First, each should be reliable, valid, and fair for its purpose. Second, each should be age- and linguistically appropriate, both in content and in method of data collection (Shepard, Kagan, & Wurtz, 1998).

In spite of what we know about measurement tools and the data they provide, those outside the research community—including policymakers, administrators, and taxpayers—too often use information collected with one tool to shed light on something else entirely. The Iowa Tests of Basic Skills (ITBS), for example, were designed to measure how a student’s academic achievements at a given moment in time compare with those of others in a representative group. However, ITBS → [continued on page 3](#)

## From the dean



Frances Stott, Ph.D.

Assessment is one of the cornerstones of contemporary American educational reform, largely because of the belief that that by illuminating what children have—or have not—learned, we can improve student achievement. Yet any discussion of early childhood assessment these days is likely to provoke more heat than light. Too little attention has been paid to the purposes of assessment—particularly to improving and supporting teaching and learning. Too little research exists to support or refute competing claims. Assessment itself has become associated with too many agendas—from assigning blame to revamping curricula and defunding programs and schools. As often as not, it provokes fear, anger, defensiveness, and self-justification among those who most need to understand its lessons. As public education groans under the load of an increasingly multicultural, multilingual, and disadvantaged student population, politicians and administrators are implementing policies that challenge the knowledge and experience of teachers whose professional judgment was once acknowledged.

Alternative assessments have been criticized for their reliance on teacher judgment—for being rooted in relationship as well as standards. Yet relationship—engaging with another and acting on the continuous feedback loop of information that that engagement provides—is central to learning. Equally central is a well-educated, deeply committed teacher. Erikson was founded to educate such teachers and to defend the importance of their professional judgment.

The Institute has long considered alternative assessments to be critical tools for examining what young children know and can do. They are part of the curriculum in Erikson's graduate education programs along with more traditional assessment methods. They have also figured in Erikson's technical assistance and research efforts over the years. During the Schools Project, for example, a partnership between Erikson and nine public elementary schools in Chicago between 1987 and 1998, project staff encouraged partner schools to adopt alternative assessments to capture student growth and progress that could be directly linked to curriculum planning. A book on the project, *Effective Partnering for School Change: Improving Early Childhood Education in Urban Classrooms*, written by Erikson faculty members Jie-Qi Chen and Patty Horsch, University of New Mexico professor Karen DeMoss, and Project Match research associate Suzanne Wagner, is forthcoming from Teachers College Press. In 2001, Erikson professor Barbara Bowman and research associate Carol Horton conducted an inquiry into the current state of expert opinion and public practice with regard to the assessment of prekindergarten children, finding great support for tying assessment to curriculum. The resultant paper, published by Erikson as part of its Occasional Papers series, is cited in this issue of *Applied Research*.

Currently, the diagnostic performance-based assessment system developed by Chen and Erikson professor Gillian McNamee is being piloted by Erikson master's students pursuing teacher certification and by more than 100 early childhood teachers in the Chicago Public Schools. Efforts at Erikson around alternative assessment have increased since the appointment of Samuel J. Meisels as president in January 2002. Considered one of the nation's foremost experts on child assessment, Meisels is cochairing a Chicago Public Schools commission on performance-based assessment. This newsletter reports on these current alternative assessment activities that aspire to directly influence teaching and learning.

---

### Applied Research in Child Development

NUMBER 4, FALL 2003

Writers: Suzanne L. Wagner,  
David Schwartz, Pat Nedeau

Editor: Pat Nedeau

Design: Sorensen London, Inc.

Photography: Lloyd DeGrane, Oscar  
Moresi

*Applied Research in Child Development* is a semiannual publication of the Herr Research Center at Erikson Institute and is funded by a generous grant from the Spencer Foundation.

Send requests for information  
or changes of address to:

Erikson Institute  
420 North Wabash Avenue  
Chicago, Illinois 60611  
Phone 312-893-7160  
Fax 312-755-0928  
[www.erikson.edu](http://www.erikson.edu)

© 2003 Erikson Institute.  
All rights reserved.  
03-136/09-03/6M/PN

## Assessment continued from page 1

results are often cited by parents and teachers as proof of how good a student a child is, or how well a teacher teaches, or whether a particular curriculum should be used or abandoned.

### Early childhood assessment in an era of accountability

The nation's current preoccupation with accountability has muddied the assessment waters considerably. To be accountable, in the case of teachers and school districts, is to be answerable—to parents and children for good teaching, to school boards and advisory councils for the prudent and productive use of resources. “How well are you doing your job?” is a question every school district must answer, but it is only one question among many. There are equally important questions whose answers more directly influence improvement in teaching and learning, such as “What should you do next to help this student learn?” or “Which of your instructional techniques work best with this child?”

Accountability testing, which rose in popularity during the 1980s, initially involved a variety of tools, including tests, portfolios, exhibitions of students' work, student initiated projects, and teacher evaluations. Just one of these methods, however, the standardized test, quickly became the instrument of choice for bureaucrats struggling to answer the call for accountability in the schools. Standardized tests have the advantage of being familiar, inexpensive to administer, and easy to mandate and implement (Elmore, 2002).

Whether or not such tests can provide the information required to determine how well students are learning, the fact remains that knowing how well the public education system is doing its job is not the same as knowing what it takes to improve it. Furthermore, the current emphasis on accountability as the *raison d'être* of assessment has led to the inappropriate use of standardized tests, which are now being proposed as data collection tools for a purpose they were never designed to serve: improving instructional quality for young children.

Perhaps no clearer examples of the confusion of purpose and application, or means and ends, are the

recently reauthorized Elementary and Secondary Education Act, better known as the No Child Left Behind Act of 2001 (NCLB), and Good Start, Grow Smart, the early childhood initiative that debuted in April 2002. Both initiatives seek to improve education and both focus on accountability, to the exclusion of other methods, to accomplish this. Both also rely on the use of a one-shot, norm-referenced achievement test designed to gather a single type of information to make consequential decisions about another matter entirely, that is, how well a district, school, or teacher is doing.

NCLB requires that states implement accountability systems that apply to all public schools and students. These systems must include annual testing of all students in the third through eighth grades, and the students' test scores are tied to rewards and sanctions (so-called high-stakes) for school districts and schools, including funding, teachers' job retention, and children's retention in grade. Of particular concern to experts in early

childhood education, the pressure to perform well on these crucial “promotional gates” tests has already resulted in the testing of children before third grade in many districts and in early childhood programs.

Good Start, Grow Smart includes the development of a new accountability system for Head Start focused on early literacy, language, and numeracy skills. Beginning in fall 2003, every local Head Start program will be required to assess all four- and five-year-old children on indicators in these domains. Although the data will be used to guide program improvement and staff training, they will also be used to evaluate Head Start agency contracts.

Many experts argue that high-stakes testing itself is problematic. Using a single data source as the basis for consequential decision making is risky, at best, and educators as well as researchers at the National Academy of Sciences point out that the need to perform well or face dismissal, closure, or grade retention encourages a degree of “teaching to the test” that

-----  
**The fact remains that knowing how well the public education system is doing its job is not the same as knowing what it takes to improve it.**  
-----

impoverishes teaching and narrows the curriculum (National Research Council, 1999). In the case of young children, high-stakes testing is particularly dangerous.

### **The problem of assessing the young child**

Most experts believe that conventional, norm-referenced tests cannot fully or accurately measure a child's skills and knowledge and are particularly inappropriate at the preprimary and early primary level (Horton & Bowman, 2002; Kagan, Scott-Little, & Clifford, 2003). Early childhood is a period when change is more the norm than is stability (LoParo & Pianta, 2000). Young children have a limited ability to read and write and are best able to demonstrate their skills and knowledge through showing and talking, not through written language. Moreover, the nature of their learning is highly integrated, episodic, and nonlinear, so the breadth and depth of their skills and knowledge cannot be fully captured through a point-in-time, content-driven test.

Young children tend to be less able to adapt quickly to a new situation such as a formal test setting, making it even harder for them to perform well on achievement tests than it is for older children. Of particular concern in a multicultural society, the abstract and verbally mediated content of these tests is potentially biased against children whose culture and socioeconomic status differ from the norm group (Stallman & Pearson, 1990).

For these reasons and more, early childhood experts generally recommend alternative assessments for young children, assessments that can be designed to capture children's development in a much wider array of domains. Alternative assessments encompass a range of instruments and procedures, including teacher observations documented in anecdotal records or by using checklists, portfolios of children's work, and interviews. Many alternative assessments focus on children's performance of specific tasks, either in the regular classroom environment or in a more specialized setting, and so tend to be more readily usable by teachers to improve and individualize curricula and instruction.

In 2000, the National Association for the Education of Young Children (NAEYC) and the

National Association of Early Childhood Specialists in State Departments of Education recommended that assessments for young children share the following characteristics:

- Involve regular and periodic observation of the child in a wide variety of circumstances that are representative of the child's behavior in the program over time
- Rely primarily on procedures that reflect the ongoing life of the classroom and typical activities of the children; avoid approaches that place children in artificial situations, impede the usual learning and development experiences in the classroom, or divert children from their natural learning process
- Rely on demonstrated performance during real, not contrived, activities, for example, real reading and writing activities rather than only skill testing

### **Research on alternative assessments**

Alternative assessments' reliance on teachers' informed, professional judgments about what children know and can do has given rise to a host of criticisms. Some researchers have suggested that alternative assessment is too "soft," and doesn't motivate teachers or parents sufficiently to seek help for students whose performance is poor (Sattler, 2002). Others argue that teachers' judgments are too subjective to measure achievement reliably (Silverstein, Brownlee, Legutki, & MacMillan, 1983).

Some are concerned about the trustworthiness and consistency, or validity and reliability, of assessments based on teachers' knowledge (Hoge & Coladarci, 1989). Do teachers have enough specific knowledge to judge children's performance in various domains? Can they successfully factor in individual motivation or learning problems (Hoge & Butcher, 1984; Salvesen & Undheim, 1994)? Paradoxically, the critics suggest that the very individuals who observe and interact with students on a daily basis, and who are trained to do so, are the least likely to be able to evaluate their students' intellectual, socioemotional, and behavioral accomplishments.

The first to answer these and other criticisms in a systematic way was Samuel J. Meisels, one of the





one-dimensional assessments (Meisels, Liaw, Dorfman, & Nelson, 1995).

Performance assessment yields several key benefits. It calls for continuous recordings to monitor progress over time, so it can shape instruction (i.e., it is “formative,” as opposed to single-record assessments, which are “summative”). It provides a continuous picture of children’s accomplishments, not a snapshot such as point-in-time tests afford. Its objectives are found within the context

nation’s foremost experts on child assessment and current president of Erikson Institute. Meisels began his research at a time when group-administered achievement testing, normally conducted among older children, began trickling down to early elementary school and preschool children. The 1983 report *Nation at Risk* had pushed academic standards into the public spotlight. In response, even seriously flawed measurement instruments such as the Gesell School Readiness Test were gaining widespread popularity (Shepard & Smith, 1988).

In the early 1990s Meisels, then a professor at the University of Michigan School of Education and a research scientist at Michigan’s Center for Human Growth and Development, began to tackle the problem of creating an alternative assessment for young children that would allow them to better demonstrate their knowledge, skills, and personality. The vehicle he chose, known as performance-based or “authentic” assessment for its reliance on actual classroom activities, tracks children’s performance on tasks that are part of their daily experience. Meisels recognized that the assessment’s relative flexibility made it particularly appropriate for young children in the midst of developmental shifts too difficult to capture with

of the classroom, not outside it, so it is instructionally relevant. It integrates a broad range of curriculum activities. It can be modified, so it allows an individualized approach to achievement. While it is based on systematic standards of knowledge and curriculum development—that is, a set of standards that identify what a child at a given age should know and be able to do in a specific domain—the standards themselves may be defined and adjusted in relation to a specific classroom, teacher, or child. It evaluates the “higher order” thinking skills—analysis, synthesis, evaluation, and interpretation of facts and ideas—far better than group-administered achievement tests can.

The curriculum-embedded performance assessment Meisels and his colleagues formulated, ultimately called the Work Sampling System (WSS; Meisels, 1997; Meisels et al., 1994; 2001), was a watershed in early childhood assessment. Designed for preschool through sixth grade, WSS is comprised of three complementary elements: (1) developmental checklists, each accompanied by guidelines, to be used by teachers to observe and document children’s performance and behavior as they engage in regular classroom activities; (2) portfolios of each child’s work collected throughout the year; and (3) summary reports, based on the checklists and

portfolios and completed three times a year, which record and aggregate information about the growth and progress of each child for use by teachers, administrators, and parents. Together, the elements form a systematic, organized portrait of the child's classroom achievements.

The system covers a wide range of domains: personal/social development, language and literacy, mathematical thinking, scientific thinking, social studies, the arts, and physical development and health. WSS gives teachers a systematic method of observing and documenting children's performance as they engage

in a range of classroom activities. At the same time, it enhances teachers' instructional decision making and encourages them to focus on children's metacognition, problem-solving, and critical thinking. By encouraging closer and

more thoughtful observations of daily curricular performance, it seeks not only to paint an accurate portrait of individual achievement but also to improve teaching. Meisels has recently carried the methodology of performance assessment to infants and toddlers, having just released the Ounce Scale, an observational assessment for infants, toddlers, and families (Meisels et al., 2003).

#### Researching performance-based assessment

To address the crucial issues of validity and reliability, Meisels and his colleagues at the University of Michigan launched carefully designed research studies using a classical psychometric model. They found that, for a field trial of the kindergarten version of WSS, both the internal and the inter-rater reliability data demonstrated the system to be "an extremely dependable assessment, given the sample size and design" (Meisels, Liaw, Dorfman, & Nelson, 1995). Teachers in this field trial participated in training sessions on three occasions over the course of the school year and also received classroom visits from project staff

approximately once a month, which included consultations on WSS. Subsequent research—with a larger sample and with children in kindergarten through third grade—confirmed and extended the earlier findings on the trustworthiness and consistency of teacher observations made using WSS (Meisels, Bickel, Nicholson, Xue, & Atkins-Burnett, 2001; Meisels, Xue, Bickel, Nicholson, & Atkins-Burnett, 2001). While the same level of accuracy may not hold for other alternative assessments, Meisels's research on the Work Sampling System demonstrates that performance-based assessment using teacher observations can yield valid, consistent information on children's learning.

Meisels and his colleagues also made research into WSS's effects on teachers' practice and children's achievement part of their larger investigation into the system's validity. Their study examined the trajectory of change in scores on the ITBS of low-income, largely minority, urban third- and fourth-grade children who had been enrolled in classrooms where WSS had been in use for at least three years. The results, recently published in *Education Policy Analysis Archives* (<http://epaa.asu.edu/epaa/v11n9>), are telling.

Two groups of demographically matched students who were administered the ITBS were compared with one another and with all other students in their grade levels in the Pittsburgh Public Schools who were also administered the ITBS. One of the target groups was composed of students who had been exposed to WSS for three years prior to being administered the ITBS; all other students had no experience with WSS. The ITBS was administered to all children by their teachers in the spring of 1997 (end of third grade) and the spring of 1998 (end of fourth grade). A total of 2,708 students received the ITBS reading assessments in both 1997 and 1998, and 2,564 students took the ITBS math assessments in both years. A longitudinal design was selected to focus on the trajectory of change from third to fourth grade as a way of capturing growth over time within students.

Results showed that students using the Work Sampling System had substantially greater gains in reading from one year to the next than students in the demographically matched comparison group (25:1) or in the school district as a whole (8:1). The students

-----  
"The solution is not to reject testing and assessment. . . . Indeed, high-quality teaching calls for high-quality assessment—the two are inextricably linked."  
-----

also had greater gains in math, though the gains were not as significant when compared with gains by the matched contrast group (Meisels, Atkins-Burnett, Xue, Nicholson, Bickel, & Son, 2003).

### Linking assessment with improvement: A complementary approach

While WSS, like other alternative assessments, was not intended to be used for accountability purposes, Meisels's latest research suggests that the Work Sampling System can make a significant difference in the achievement typically measured by conventional, group-administered accountability tests, particularly in reading.

The evidence suggests an intriguing possibility: improving scores on accountability examinations by using curriculum-embedded performance assessments to enrich teaching and learning. Standardized tests may furnish a child with a percentile ranking, but cannot be used effectively to improve learning. In a study of 18 states with high-stakes testing programs, the learning levels of the students in all but one state were at the same level as before the testing policies were implemented (Amrein & Berliner, 2002). Such tests do not provide teachers with the constant, direct feedback that is the basis of effective teaching. Further, they “rob teachers of their sense of judgment about how to help children develop to their optimal potential” (Meisels, 1993). Yet they provide important data about year-to-year achievement differences in students, data that are crucial to reporting accountability in public educational programs.

Alternative assessments, specifically curriculum-embedded, performance-based assessments, evaluate students' actual classroom performance on an ongoing basis in terms of standards-infused criteria. These criteria in turn suggest next steps in curriculum development that are consistent with advancing progress toward attainment of the defined standard.

In a commentary published in *Education Week*, Meisels (2003) explained it this way: “The solution is not to reject testing and assessment. . . . Indeed, high-quality teaching calls for high-quality assessment—the two are inextricably linked. To improve teaching, we need comprehensive, classroom-based evidence

about what children are learning that can be translated easily into meaningful instructional strategies to enhance teaching and improve learning. Systematic, well-researched, observational assessments, whose results can be aggregated across programs, can accomplish this.”

The research literature currently contains few studies of the impact of curriculum-embedded performance assessments on group-administered achievement test scores (Borko, Flory, & Cumbo, 1993; Falk & Darling-Hammond, 1993). Yet the research that Meisels and his colleagues report suggests that adding a system of performance assessment to conventional, norm-referenced testing may be the key to improving learning. If it is, the most important challenge in the coming years may be not just to continue to demonstrate performance assessments' validity, reliability, and effectiveness, but also to integrate them into the prevailing accountability system in ways that preserve their usefulness and integrity.

### References

- Amrein, A. L., & Berliner, D. C. (2002). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives*, 10(18).
- Borko, H., Flory, M., & Cumbo, K. (1993, October). *Teachers' ideas and practices about assessment and instruction: A case study of the effects of alternative assessment in instruction, student learning, and accountability practice* (CSE Technical Report 366). Los Angeles: Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Elmore, R. F. (2002, April). Unwarranted intrusion. *Education Next*. Retrieved July 16, 2003, from <http://www.educationnext.org/20021/30.html>
- Falk, B., & Darling-Hammond, L. (1993, March). *The primary language record at P.S. 261: How assessment transforms teaching and learning*. New York: National Center for Restructuring Education, Schools, and Teaching.
- Hoge, R. D., & Butcher, R. (1984). Analysis of teacher judgments of pupil achievement levels. *Journal of Educational Psychology*, 76, 777–781.
- Hoge, R. D., & Coladarci, T. (1989). Teacher-based judgments of academic achievement: A review of the literature. *Review of Educational Research*, 59, 297–313.

- Horton, C., & Bowman, B. T. (2002). *Child assessment at the preprimary level: Expert opinion and state trends* (Occasional Paper No. 3). Chicago: Erikson Institute.
- Kagan, S. L., Scott-Little, C., & Clifford, R. M. (2003). Assessing young children: A policy brief that addresses what policy makers need to know and do. In Scott-Little, C., Kagan, S. L., & Clifford, R. M. (Eds.), *Assessing the state of state assessments: Perspectives on assessing young children*. Tallahassee: SERVE.
- LoParo, K. M., & Pianta, R. C. (2000). Predicting children's competence in the early school years: A meta-analytic review. *Review of Educational Research*, 70(4), 443–484.
- Meisels, S. J. (1993, July 5). Remaking classroom assessment with the Work Sampling System. *Young Children* 48, 34–40.
- Meisels, S. J. (1994). Designing meaningful measurements for early childhood. In B. L. Mallory & R. S. New (Eds.), *Diversity and developmentally appropriate practices* (pp. 202–222). New York: Teachers College Press.
- Meisels, S. J. (1997). Using Work Sampling in authentic assessments. *Educational Leadership*, 54(4), 60–65.
- Meisels, S. J. (2003, March 19). Can Head Start Pass the Test? *Education Week*, p. 44.
- Meisels, S. J., Atkins-Burnett, S., Xue, Y., Nicholson, J., Bickel, D. D., & Son, S. H. (2003, February 28). Creating a system of accountability: The impact of instructional assessment on elementary children's achievement test scores. *Educational Policy Analysis Archives*. Retrieved July 16, 2003, from <http://epaa.asu.edu/epaa/v11n9/>
- Meisels, S. J., Bickel, D. D., Nicholson, J., Xue, Y., & Atkins-Burnett, S. (2001). Trusting teachers' judgments: A validity study of a curriculum-embedded performance assessment in Kindergarten–Grade 3. *American Educational Research Journal*, 38(1), 73–95.
- Meisels, S. J., Dombro, A. L., Marsden, D. B., Weston, D. R., & Jewkes, A. R. (2003). *The Ounce Scale*. New York: Pearson Early Learning.
- Meisels, S. J., Jablon, J., Marsden, D. B., Dichtelmiller, M. L., & Dorfman, A. (1994). *The Work Sampling System* (3rd ed.). Ann Arbor, MI: Rebus Inc.
- Meisels, S. J., Jablon, J., Marsden, D. B., Dichtelmiller, M. L., & Dorfman, A. (2001). *The Work Sampling System* (4th ed.). Ann Arbor, MI: Rebus Inc.
- Meisels, S. J., Liaw, F. R., Dorfman, A., Nelson, R. F. (1995). The Work Sampling System: Reliability and validity of a performance assessment for young children. *Early Childhood Research Quarterly*, 10, 277–296.
- Meisels, S. J., Xue, Y., Bickel, D. D., Nicholson, J., & Atkins-Burnett, S. (2001). Parental reactions to authentic performance assessment. *Educational Assessment*, 7(1), 61–85.
- National Research Council (1999). *High stakes: Testing for tracking, promotion, and graduation*. Washington, DC: Author.
- National Association for the Education of Young Children (NAEYC) and National Association of Early Childhood Specialists in State Departments of Education (NAECS/SDE). (2002, November). Early learning standards: Creating the conditions for success. Retrieved on July 16, 2003, from [http://www.naeyc.org/resources/position\\_statements/position\\_statement.pdf](http://www.naeyc.org/resources/position_statements/position_statement.pdf)
- Salvesen, K. A., & Undheim, J. O. (1994). Screening for learning disabilities. *Journal of Learning Disabilities*, 27, 60–66.
- Shepard, L., Kagan, S. L., & Wurtz, E. (1998). *Principles and recommendations for early childhood assessments*. Washington, DC: National Educational Goals Panel.
- Shepard, L. A., & Smith, M. L. (1986). Synthesis of research on school readiness and kindergarten retention. *Educational Leadership*, 44, 78–86.
- Silverstein, A. B., Brownlee, L., Legutki, G., & MacMillan, D. L. (1983). Convergent and discriminant validation of two methods of assessing three academic traits. *Journal of Special Education*, 17, 63–68.
- Stallman, A. C., & Pearson, P. D. (1990). Formal measurement of early literacy. In L. M. Morrow and J. K. Smith (Eds.), *Assessment for Instruction in Early Literacy* (pp.7–44). Englewood Cliffs, NJ: Prentice Hall.
- Sattler, J. M. (2001). *Assessment of children: cognitive applications* (4th ed.). San Diego: Jerome M. Sattler.

### About the author

Suzanne L. Wagner is a research associate with Project Match.

### Note to lead article

Erikson Institute does not endorse the WSS or Ounce Scale or any other commercially-produced assessment.



## A Conversation with Sam Meisels

Sam Meisels's introduction to child development—indeed, his introduction to children—came from a part-time placement in a Montessori preschool classroom in Harvard Square. “I had not seen a young child, as far as I know, since I was one myself,” he once wrote of the experience, “but from almost the first moment I walked into that preschool my life changed.” The two-afternoon-per-week commitment became a consuming passion, launching a career that has spanned nearly three decades. The one-time student of moral philosophy left Harvard with a doctorate in education, a year teaching science to preschoolers, and two years' experience teaching kindergarten and first grade in a public school.

A deep appreciation for the practical as well as the epistemological aspects of child development has marked his subsequent career. In addition to university teaching and research in the Department of Child Study at Tufts, Meisels directed the Eliot-Pearson Children's School there, later taking a leave of absence to become a senior adviser in early childhood development in the Developmental Evaluation Clinic at Boston Children's Hospital. At the University of Michigan, where he held positions at the Institute for the Study of Mental Retardation and Related Disabilities and the Center for Human Growth and Development, Meisels built on the foundation of his experiences as teacher, researcher, practitioner, and policy analyst. Currently an emeritus professor of education, Meisels has distinguished himself as a leading scholar on methods of developmental assessment for young children; the effect of standardized tests on children; developmental consequences of high-risk birth; and the impact of state and federal policies on the families of children with disabilities. He assumed the presidency of Erikson Institute in January 2002.

**I have heard that your interest in child assessment stems from your own experience as a child. Is that true?**

It's true that I always knew as a child that I understood more than I could ever show on one of those tests. Sometimes I thought I knew how to do what they were asking of me, but I couldn't do it the way they asked me to do it. Sometimes I just couldn't figure out what the test was calling for. This frustration—this sense of being judged unfairly—has certainly been a very significant motivation for me to create assessments like the Work Sampling System that have, in a sense, many doors, many windows, many ways in. There shouldn't just be one right way to do things, because with most of the things we want children to do, with the exception of multiplication tables and addition and subtraction, there are lots of right answers. Even with mathematics, there are many different ways to get to the right answer.

**How have attitudes about alternative assessments, including performance-based assessments, changed over the years?**

**Does the current focus on group-administered, standardized testing make it a better or worse climate for these kinds of assessments?**

I've seen some significant changes over the past 10 to 15 years. Beginning at the end of the 1980s and in the early 1990s, there was a tremendous explosion in interest in alternative assessments, performance-based assessments being the principal alternative that was used. Offices of alternative assessments were established in state departments of education all across the country. Then people started asking, “Well, where's the data that show these assessments are valid?”

Unfortunately, as is often the case in American education, the methodology and

criteria used to evaluate the existing model were then placed on top of the new model, and it didn't work, it didn't fit. Hence, the new model was rejected.

Researchers often assemble a group or battery of assessments in their studies as if they are all faceless and nameless, though they all have a function. To me, each of

those assessments is like a person sitting around your table. You know something about them. Some you like, some you don't like, some you like for this reason, some you don't like for that reason; some of them are brilliant, and some of them are fakes. We need to

apply at least as much discrimination to the tests we adopt as to the guests we invite to a dinner party. It matters to me how we assess kids, and for what purpose, and what our choices are.

Despite the fact that we now have federal legislation calling for high-stakes tests and assessments, states are still coming to me and asking for help because they know that using the old models of group-administered achievement tests will not accomplish their goal of improving education. And

they do respect the fact that we have more research on Work Sampling than exists on any other performance-based assessment at any level. So state departments of education are, in many cases, exercising the discrimination I was just talking about.

**Are you concerned about the specter of alternative assessments like your Work Sampling System being used in a high-stakes context?**

I am, and that certainly will happen. And when it does happen, I will tell them, "You will ruin it." This isn't just a parent saying, "Don't spoil my child." I'm trying to say that when high stakes are applied to performance assessments, the assessments won't mean the same thing anymore. Performance-based assessments like the Work Sampling System are all based on teacher judgment, and teacher judgment is going to be influenced by the context, the framework in which one is working.

**Your mentioning teacher judgment brings me to my question about psychometric reliability, a question that dogs performance-based assessments. Over the years, you yourself have spent a great deal of time on research to test the reliability and validity of the Work Sampling System. What have you found about the behavior and attitudes of individual teachers in this regard?**

There's a lot of local adaptation with Work Sampling. One of the first field trial sites we had was located in Michigan, and I went there quite often to observe the initial implementation. I also sent my best trainer there every four to six weeks throughout the entire school year. Yet, at the end of the school year, when I came in and watched the teachers, I found each one using WSS differently. We call this "street-level adaptation."

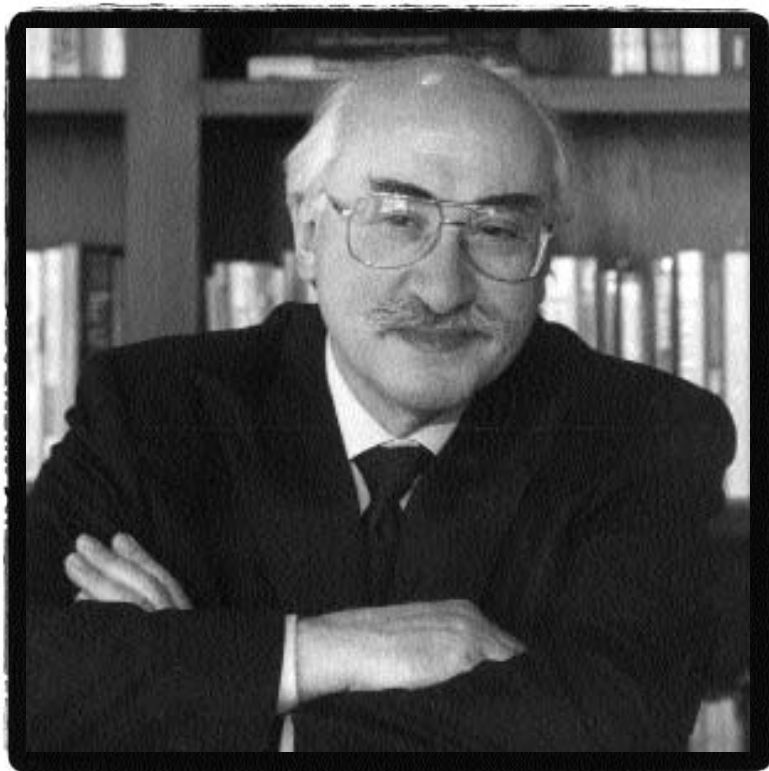


Photo: Oscar Morse

We actually studied this at another site in Pittsburgh. For her dissertation, one of my doctoral students from the University of Michigan, Julie Nicholson, interviewed the kindergarten through third-grade teachers in our study in order to investigate the consequential validity of WSS. What we learned was that there was a great deal of variation among these teachers. Also, there was variation in how much they liked and didn't like WSS, how much they supported it and didn't support it, and so on. But interestingly, we not only had these interpretive data from the teacher interviews, we also had empirical data on how the children were doing. Regardless of the variability, there was still a common level of implementation that resulted in certain things for children that were generalizable. So there can be a lot of variability and still an approach like WSS can work, but this depends on how you construct the assessment so that it has room for personal adaptation.

#### **Do teachers get training when schools or districts implement the Work Sampling System?**

Ideally, yes, but in reality, not always. One superintendent told me, "When I buy a car, I don't let the salesman tell me where I can drive it." That was hard to hear, but it's the truth. To some extent, you create something like this and then you've got to let go. You can't prevent people from doing what they want with the instrument you've created.

The recommended training for Work Sampling is three days of intensive training, on site, with everyone who is going to be part of it, and then every six to eight weeks

after that, a one-day follow-up training session, usually for a year. Sometimes this goes on for several years. What we try to do is transfer the capacity from us to them. Nobody wants all of that knowledge to be bottled up in New York or Ann Arbor or Chicago. They want to have it where they are, where they need it, especially because of teacher turnover. Of course, if only one or two elements of WSS are used, the training is much less intensive.

#### **How can schools justify the cost of teacher training, particularly in tight budget times?**

I don't think schools can afford not to train teachers in assessment. Preparation for use of performance assessments takes time, and that takes money. Is that wasteful? Not if it results in better teachers, because better teachers will yield better-educated students. The most wasteful, inefficient thing we can do now is ignore how important it is for teachers to be able to individualize their instruction based on children's performance.

#### **Can data from performance assessment be used for purposes other than curricular enhancement?**

Absolutely. WSS, which is one of the most well-researched and widely adopted performance assessments for young preschool and elementary children, has been used successfully to generate data for a variety of purposes. In both Kansas and Maryland, for example, state administrators have used the Work Sampling System to evaluate state prekindergarten programs. The Michigan Department of Education has used alternative assessments developed by High/Scope to evaluate its state preschool program.

#### **I understand that you are currently cochairing a Chicago Public Schools commission on performance-based assessment. Chicago is well known for its high-stakes accountability system based on students' scores on standardized achievement tests, so what are the expectations for this commission?**

I think both Arne Duncan [Chicago Public Schools (CPS) chief executive officer] and Barbara Eason-Watkins [CPS chief education officer] are concerned that there isn't enough focus on instruction. Barbara, from her own experience as a very gifted principal here in Chicago, and as a teacher, knows that the better we teach children, the better they will perform on tests.

Teaching children better, however, doesn't mean teaching them the tests. One of the better-known test corporations advertises that more copies of its test-prep booklets were sold last year than all copies of *Black Beauty* and *Charlotte's Web*. Can you imagine? It's an enormous distortion, and CPS is trying to work against it.

The commission isn't going to literally create an assessment, but we will recommend certain criteria, and we may recommend certain approaches and instruments. I think it's an extremely positive sign. The pendulum may be moving slightly.

# Considering options in Chicago:

THE COMMISSION ON IMPROVING CURRICULUM-BASED ASSESSMENT

BY PAT NEDEAU

In October 2002, Chicago Public Schools CEO Arne Duncan convened the Commission on Improving Curriculum-based Assessment, or CICA, a 21-member blue-ribbon panel charged with creating an alternative assessment system to work alongside standardized tests in the 437,000-student district. Chaired by Erikson president Samuel J. Meisels and

Donald M. Stewart, president and CEO of the Chicago Community Trust, the commission includes practicing teachers, principals, representatives of the Chicago Teachers Union, community leaders, academics, and CPS administrators. Their goal is to improve

teaching and learning, but to do so, they must successfully navigate the competing political agendas and legislative mandates that make change so difficult.

Like other districts around the country, the Chicago Public Schools has traditionally relied on only a few instruments to collect the data used to make critical decisions about its children and schools. Aggregated data from these tests become the canary in the mine: When gross achievement disparities emerge among schools or student groups, administrators take corrective action, such as reallocating resources.

First among the instruments used to measure school improvement or its lack

have been such norm-referenced standardized tests as the Iowa Tests of Basic Skills. Test publishers warn against putting too much weight on normative test scores, but the choice in Chicago, as in other big-city districts, has been dictated by economics. Unfortunately, data gathered from norm-referenced testing alone are insufficient to gauge whether a system is moving in the right direction or to improve instruction.

One reason for this is a problem of scale. “In a norm-referenced test, the larger the local sample, the more it behaves like the national sample,” Kate Nolan, staff director of CICA and a research associate at Erikson, explains. “The whole country is working on improving their test scores on an instrument that relies on a norm-referenced curve to chart progress.

“Imagine the whole country decides to go on a diet, and you’re going to measure your weight-loss success or failure by comparing your position relative to everyone else’s in a national sample. Your position at the start is in the 50th percentile—half the people in the sample weigh more than you, half less. If everyone loses weight, including you, you could still be in the 50th percentile at the end of the diet. If you don’t have a measure of your actual before and after weights, but only a before and after percentile position, how do you show that you lost weight? If you’re at the lowest end of the bell curve, you might be able to detect a change in position, but if you look at the results on the standardized tests for Chicago in the past year, they’re not at the low end. They’re approaching that 50th percentile.”

Duncan and his chief education officer, Barbara Eason-Watkins, hope that a system

“Imagine the whole country decides to go on a diet. . . . If everyone loses weight, including you, you could still be in the 50th percentile at the end of the diet. . . . how do you show that you lost weight?”

of well-developed assessment tools will not only reveal progress—or its lack—but provide useful feedback for classroom improvements. They seek an assessment system that can give immediate feedback to teachers, one that teachers can use to make “just-in-time” adjustments in instruction. Implementing such a system, however, is only half the battle. Realizing improvements in curriculum and instruction through assessment will also require a public information campaign and professional development. First, teachers, administrators, parents, and community advocates will have to be taught how to use assessment data responsibly and how to extract information that’s useful for improving instruction out of these data. Second, they will need to understand how to translate this information into richer and more demanding curricula and instruction, providing more entry points for more students.

The commission, whose recommendations are scheduled to be presented in December, is currently debating the proper use of curriculum-based assessment data relative to standardized test data. While some commissioners would prefer to eliminate standardized tests, federal and state laws make it virtually impossible to do so. A three- to five-year phase-in of curriculum-based assessment would allow educators to analyze how the two systems might work together, based on empirical data.

Yet the two-system approach has pitfalls. Inevitably, the system that educators perceive as carrying the most weight, from the perspective of decision making, will be given the most attention in the classroom. It’s a reality that could compromise the results of the experiment and doom alternative assessments to failure.

Experts also point out that the two approaches may produce conflicting data. The public—including the politicians and administrators responsible for school improvement—will want to know which data set is accurate, and the answer will not necessarily be self-evident. Returning to the diet metaphor, if the dieter does not lose weight but loses inches, which data set—pounds or inches—is “true” and should be used to decide whether the diet is working?

“It’s inevitable: For some students, the standardized tests are going to say one thing and the classroom-based assessments will say another,” says Nolan. “Part of teaching principals, administrators, parents, and community members what to do with assessment data is helping them to make sense of the conflicting data. Parents and teachers alike are going to have to be patient and study the data before they know for certain how to interpret it.”

Whether the recommendations of the commission will be successfully implemented will depend on many factors, not the least of which are funding and political will. The risks attendant on changing the system are real, and the costs are not limited to materials development or purchase. Teacher education may be the single highest cost to any commission proposal. Few teachers in the classroom today were taught to understand the connection between assessment and curriculum and instruction. Many teacher certification programs do not cover the subject. Developing teachers’ abilities to *use* the tools being placed at their disposal will require new approaches to teacher preparation as well as solid and extensive professional development for teachers already in the system.

All members of the commission agree that involving teachers in the assessment design and selection process is not only wise but necessary. The Chicago Teachers

Union plans to use some of their larger meetings as fora for teachers to contribute their ideas. Teachers from the Chicago Foundation for Education, whose members are CPS teachers interested in developing professional aspects of teaching, will act as an expert advisory group. Another idea is to use the operational structure established by CPS to implement the standards-based curriculum initiative and allow teachers to be part of those groups as they do their work.

“Schools are not places, by and large, where it’s easy for teachers to share the outside work that they’re doing, but we have to try,” Nolan says. “Teachers deserve to be part of the conversation, and they must feel ownership if the system is going to work.”

Ultimately, Nolan and the commission believe that the only model for real change in the testing environment is one that’s based on collaboration. “It may be possible to take a stand, to say, for example, that under third grade you shouldn’t be using standardized tests for anything other than research purposes—certainly not to make decisions about children.

“But until we build a system that is going to be almost as easy to understand, almost as inexpensive to use, and almost as reliable as high-stakes, norm-referenced tests, those tests are going to continue to carry the day. It’s our responsibility, then, to try to develop and build sets of alternative policies and practices over time, in dialogue with all constituents.”



# Building multidirectional bridges through classroom assessment

-----  
Bridging is not intended to assess school readiness. . . or school achievement. . . It attempts to assess only what children are learning (content) and how they are learning it (process) at any moment in time. . .  
-----

Learning—as measured by any method of assessment—is inseparable from teaching. To examine one without considering the other is to miss an opportunity to reinforce them both. That thinking, coupled with a healthy respect for the realities of the classroom, launched Erikson professors Jie-Qi Chen and Gillian McNamee on a project to devise a new assessment tool for the early childhood classroom. The result, *Bridging: A Diagnostic Assessment for Teaching and Learning in Early Childhood Classrooms*, paints a detailed portrait of what and how children between the ages of three and eight learn. As important, it invites teachers to think about what and how they teach.

The name “Bridging” is particularly apt. “We want this assessment tool to help teachers build multidirectional bridges—between children’s curiosity and cognitive strengths and the intellectual demands of school; between *what* and *how* they learn; between curricular areas; between the particulars of a few children in the assessment sample and the larger teaching and learning issues that face all the children in the classroom,” explains Chen. “Ultimately, we’re

looking for ways to bridge between assessment and curriculum.”

## The basics of Bridging

At the heart of Bridging is a set of 19 curricular activities using materials familiar to most early childhood classrooms: building symmetrical designs with pattern blocks, dictating and then dramatizing stories, moving to music, and drawing self-portraits, for example. Familiarity with both the activity and materials reduces anxiety for students and teachers alike and minimizes the time and effort usually required to implement new or different activities. Each activity has been carefully designed to reveal young children’s cognitive strengths and approaches to learning in one of five broad curricular areas: language and literacy; physical, natural, and mechanical sciences; performing arts; numbers and geometry; and visual arts. By focusing on a variety of subject areas rather than a few, Bridging supports education that strives for multiple ways of learning, doing, and achieving.

Each Bridging activity also includes key concepts and skills in the content area; materials, procedures, and strategies for

### Bridging activities

|                      |                         |                       |
|----------------------|-------------------------|-----------------------|
| Child Picks a Book   | Pattern Block Pictures  | Shadows and Light     |
| Teacher Picks a Book | Pattern Block Pinwheels | Nature Display        |
| Story Dictation      | Pattern Block Puzzles   | Building a Model Car  |
| Story Dramatization  | Counting                | Moving to Music       |
| Crayon Technique     | Subtraction             | Playing an Instrument |
| Self-Portrait        | Estimation              | Singing a Song        |
|                      | Fair Share              |                       |



Photo: Lloyd DeGrane

*Chicago Public School teachers met regularly to discuss their implementation of Bridging activities and their findings with respect to individual students and to reexamine the links between a particular activity and key concepts and skills.*

implementation; and a “bridge to curriculum” analysis that helps teachers clarify and review the skills and strategies the child possesses and is working on and suggests ways to extend activities in order to further children’s learning in upcoming days and weeks.

Teachers observe and document a child’s performance on each activity with the help of a detailed rubric. The rubrics reflect ten levels of development of children’s knowledge, skills, and abilities in relation to key concepts in the different subject areas represented by the activities. They afford teachers the opportunity to better understand a child’s place on a developmental continuum in varied subject areas.

Rather than providing age-group norms, Bridging asks teachers themselves to identify expected performance levels for a given grade or age at a particular point in time, usually October and May. Teachers then assess children in their classroom in relation to this norm. The method puts



Photo: Lloyd DeGrane

teachers’ expertise and experience to work and gives them the opportunity to think about what they know. In this way, Chen and McNamee hope Bridging will improve teachers’ ability to see children for who they are and to think about what they do when they come to school.

Bridging is not intended to assess school readiness by focusing on what children are ready to learn prior to any formal instruction, nor is it designed to examine children’s school achievement by focusing on what they have learned as a result of a particular course of instruction. It attempts

to assess only what children are learning (content) and how they are learning it (process) at any moment in time, with the clear understanding that children’s working approaches are not stable but may change over time and vary depending on the nature of the activity and its social structure. The tool identifies a total of 14 working-approach variables across the 19 activities. Half of the variables are evaluative qualities (e.g., a child’s tolerance for frustration or proclivity for planning); the other half portrays descriptive qualities (e.g., playfulness, pace of work).

Evaluative qualities, which describe aspects of a working approach that either promote or hinder a child’s performance, are scored numerically on a scale of one to five. The higher the score, the more adaptive, goal-oriented, organized, and social—and therefore more conducive to classroom learning—the working approach. Descriptive qualities, which capture interesting individual differences, are not scored but are noted. McNamee explains the importance of documenting the working approach: “It is a significant complement to the performance level because it can indicate the circumstances under which a child shows strengths and special interest, and it can help teachers plan the conditions under which each child can thrive as a learner.”

Bridging activities provide multiple ways for children to demonstrate mastery of concepts and skills. A child’s understanding of a narrative structure, for example, may be demonstrated through a conventional book reading activity in Bridging, such as Teacher Picks a Book, or through one of several alternatives Bridging provides, such as Story Dictation and Story Dramatization. By the same token, both Crayon Technique and Pattern Block Pinwheels are activities that assess the child’s ability to create visual images, but they use very different media, each with its

own representational qualities. By consistently linking assessment and key concepts, Bridging strengthens teachers' understanding of those concepts and their development in the early childhood years.

The activities lend themselves to a variety of social structures to suit the needs of different classrooms and individual children, as well as to elicit a child's best performance. Some activities, such as Moving to Music and Building a Model Car, can be implemented in a large group, providing an effective way to engage an entire class while assessing only a few children. Most activities can also be carried out in small groups and dyadic structures to allow for different kinds of interaction patterns and a closer view of children's thinking. While not a comprehensive curriculum package, Bridging is built around activities that can be readily carried out as part of the curriculum in every type of preschool, kindergarten, and primary classroom.

### **Preliminary data analysis reflects complexity of learning and teaching**

In 2002–03, Chen and McNamee introduced Bridging to 60 early childhood teachers in Chicago Public Schools. All teachers attended a weeklong intensive training on the assessment tool and process during the summer and then monthly seminars throughout the school year, totaling 45 hours. Each teacher conducted the assessment with four children in her classroom at the beginning of the year (October) and at the end (May). Chen and McNamee's research team interviewed 20 of these teachers in depth at the beginning and end of the school year and is comparing them with a control group. Similar data has been collected among preservice teachers. Finally, the researchers conducted pre- and post-assessment of 40 preschool and kindergarten children on all 19 Bridging activities.

-----  
**No one factor—subject area, activity structure, type of material, key concept emphasis, social structure, or working approach—overrides all others in determining performance.**  
-----

Data analysis is incomplete, but several trends are emerging reflecting the complexity of learning and development among young children, as illustrated by the following preliminary findings.

Every child observed in Bridging activities presents a “jagged profile,” supporting the hypothesis that, with respect to children's strengths and patterns of development in different areas, diversity is the rule rather than the exception.

Children's performance scores among Bridging activities are relatively independent of one another. This independence is as much due to the nature of the assessment as to the relative autonomy of the intelligences involved. No one factor—subject area (e.g., language versus math), activity structure (e.g., open-ended versus highly structured), type of material (e.g., pattern blocks versus paper and pencil in the area of visual arts), key concept emphasis (e.g., spatial or part/whole relationship versus number concepts in math), social structure (e.g., one-on-one versus small or large group), or working approach (e.g., impulsive versus goal-oriented)—overrides all others in determining performance.

In general, the older the child, the higher the performance score. Preschool children, however, as a group, show a more

rapid rate of development within a year than do kindergartners.

Performance level scores on different Bridging activities do not develop at the same rate. For some activities, a child may proceed with rapid development initially and then slow down, whereas for others the development pattern is opposite.

There is a relationship between young children's task performance and working approaches. More specifically, some approaches seem to facilitate performance better than others. Children tend to use some working approaches in their stronger areas and others in their weaker areas. Age, gender, and the activity's social structure all effect working approach.

A second trend concerns the teachers who are using Bridging, particularly those at the kindergarten level. They are reporting (1) more knowledge about children's cognitive strengths and weaknesses, (2) greater ability to identify key concepts for varied curricular activities, (3) more ideas about using a child's identified strengths and vulnerabilities to design appropriate curriculum activities to help the child, and (4) deeper understanding of the impact of various factors, such as subject areas, the group learning process, and materials, on children's learning and classroom teaching. These understandings and skills are the very basis of effective teaching. If the final data analysis supports these preliminary findings, Bridging may contribute to the improvement of performance for both teachers and students.

*Bridging: A Diagnostic Assessment for Teaching and Learning in Early Childhood Classrooms was developed with the generous support of the Polk Bros. Foundation, Field Foundation of Illinois, and McDougal Family Foundation.*

# News from the Herr Research Center

## Research update

Associate professor Jon Korfmacher was guest editor, with Sydney Hans, of the November issue of *Zero to Three*, which focused on research that seeks to understand the paraprofessional relationship. The pair coauthored “The Professional Development of Paraprofessionals.” Korfmacher also contributed “The Helping Relationship in a Teen Parenting Program,” written with Isabela Marchi.



Jie-Qi Chen

The NAEYC conference in New York included a number of presentations by Erikson faculty. **President Samuel J. Meisels** presented “Using Work Sampling for Head Start with the National Head Start Outcomes Framework” and “A New Assessment for Birth to 3-year-olds: The Ounce Scale” with Dorothea Marsden. **Dean Frances Stott** presented “Caregiver Goals and Societal Expectations: Perspectives from the Four College Early Childhood Consortium” with consortium faculty, while **Professor Jie-Qi Chen**, with doctoral student **Ann Masur**, presented “Putting Assessment, Teaching, and Learning Together: Erikson Institute’s Performance-based Assessment Instrument.” She and senior research associate **Charles Chang** also presented “Computer Technology and Early Childhood Teachers: A Shocking Reality and Possible Changes.” Finally, **Professor Joan McLane** presented “Play in Early Childhood Settings: What Matters Most?”

*Love to Read*, a collection of essays on developing and enhancing the early literacy skills of African American children edited by **Professor Barbara Bowman**, was published by the National Black Child Development Institute in winter 2002.



Jon Korfmacher



Linda Gilkerson

## “‘Does not.’ ‘Does too.’ Thinking About Play in the Early Childhood Classroom,” **Joan Brooks**

**McLane’s examination of teachers’ attitudes about play, their understanding of its purpose and potential, and how they incorporate it into their classrooms, has recently been released as part of the Herr Center’s Occasional Paper Series. To request copies, contact the Office of Communications, 312.893.7160.**

Senior research associate **Daniel Scheinfeld** reported strong gains in reading comprehension scores on the Iowa Tests of Basic Skills (ITBS) at two schools, Seward and Brentano, where he has been working to implement literature circles. Seward recorded a 13.2 percent gain, moving from 36.9 percent of students at or above grade level in May 2001 to 50.1 percent in May 2002. Brentano recorded a 8.1 percent gain, from 33.9 percent to 42 percent, in the same period. The average CPS gain on the ITBS in that period was 2 percent. Illinois Standards Achievement Tests showed similar gains: 12.7 percent and 9.3 percent respectively between April 2001 and April 2002. Both schools were awarded the title “School of Distinction” by Chicago Public Schools as a result.

In February, the *Education Policy Analysis Archives* published “Creating a System of Accountability: The Impact of Instructional Assessment on Elementary Children’s Achievement Test Scores,” coauthored by **Samuel J. Meisels**. The study examined the trajectory of change in scores on the Iowa Tests of Basic Skills of low-income, urban third and fourth graders who had been enrolled in classrooms where Meisels’s Work Sampling System was used for at least three years.

On March 1, **Jana Fleming**, a senior research associate at Erikson, became executive director of child development studies for the City Colleges of Chicago. Fleming’s task is to strengthen the City Colleges’ child development degree programs, including unifying standards between the teacher education programs and child development laboratory schools operated by the City Colleges and further developing practicum experiences available to students.

**Associate professor Aisha Ray** participated in the Illinois Association for Infant Mental Health Spring Seminar Series, Contexts of Parenting: What is Universal and What is Unique, in March. She and Northwestern University Medical School copresenter Craig Garfield, M.D., presented “Fathers in Society: Implications for Infants and Families.”

Also in March, **Professor Linda Gilkerson** formally launched the Fussy Baby Network, a resource, support, and consultation network for families with infants who cannot be comforted or who have sleep or feeding issues during the first year of life. Network services include a home visiting program, clinic, telephone support or “warm” line, Internet support, and parent groups. Funded by the Doris Duke Charitable Foundation and Irving B. Harris Foundation, the network is Erikson’s first formal clinical outreach.



In March, **Jon Korfmacher** participated in “Capacity of Children to Report for Themselves,” a panel discussion hosted by the American Academy of Pediatrics. At the biennial meeting of the Society for Research in Child Development in Tampa in April, he was a panelist for “Research Agenda Versus Community Agenda: How Do Prevention and Intervention Research Programs Successfully Navigate This Issue?”



Robert Halpern

In spring, **Professor Robert Halpern** gave the keynote address at “After-School 2010: Considering the Future of After-School Programming in the United States” at the J. Paul Getty Museum in Los Angeles.



Joan Brooks McLane

**Frances Stott** presented “Supporting Grieving Families When Child Abuse is Suspected” at Children’s Memorial Hospital in February. She also presented on a plenary panel on “play with violent themes” at the Playing for Keeps conference at Yale University in March. In May, she was guest editor, with representatives of the three other members of the Four College Consortium (Bank Street, Pacific Oaks, and Wheelock), of the May issue of *Zero to Three*, which focused on “Caregiver Goals and Societal Expectations.” Stott and **Robert Halpern** also coauthored “Listening to the Voices of Families: Thoughts, Hopes, and Fears in a Latino Community,” which appeared in the issue.

**Samuel J. Meisels** addressed a national audience in Washington, D.C., in April at “Hearts and Minds: The State of Babies and Toddlers,” a dialogue with the nation’s leading experts on early childhood organized by Zero to Three. Meisels was also active nationally in the debate over proposed changes in Head Start, providing counsel to Senators Daschle,

Kennedy, Clinton, Jeffords, and others as well as giving briefings to House and Senate staffers on Capitol Hill. In March, his commentary “Can Head Start Pass the Test?” was published in *Education Week*. In June, Meisels released the Ounce Scale, an assessment for infants, toddlers, and families. The project is the culmination of six years of work and research.

**Linda Gilkerson** gave a plenary talk, “Reflective Process: Seeing, Feeling, Remembering,” at the Michigan Association for Infant Mental Health in May.

**Joan Brooks McLane** presented “Thinking About Play in the Classroom” at the annual meeting of the Jean Piaget Society on June 6 in Chicago.

Research associate **Kathleen Kostelny** gave a presentation at the 8th International Symposium on the Contributions of Psychology to Peace, in Sunne, Sweden, in June. She also gave the keynote address, “Children in Trauma: Living in a World of Danger,” at the 6th annual Crisis Intervention Conference in Waukesha, Wisconsin, in September.

In July, **Robert Halpern’s** book *Making Play Work: The Promise of After-School Programs for Low-Income Children* was published by Teachers College Press. The book outlines the evolution of after-school programs and their role in the lives of children, providing a framework for reflecting on broader contemporary issues, such as the effects of poverty on children in the United States and current directions and expectations for the future of after-school programs. “Physical (In)Activity,” Halpern’s study of active play and sport among low-income children and youth, was published in August by the Robert Wood Johnson Foundation’s After School Project.

*Effective Partnering for School Change: Improving Early Childhood Education in Urban Classrooms* by **Jie-Qi Chen, Patricia Horsch, Suzanne Wagner, and Karen DeMoss** will be published by Teachers College Press in November 2003. The authors explore the challenges and promise of university-school partnerships as a vehicle for improving the public schools. The book draws on lessons learned from Erikson’s Schools Project—an 11-year initiative involving nine public elementary schools in low-income Chicago neighborhoods.

This November’s NAEYC conference in Chicago will highlight the work of a number of Erikson faculty, including **Aisha Ray**, who will copresent a symposium on fathers and early childhood practitioners, and **Linda Gilkerson**, who will present “Developmental Screen in Infant/Toddler Child Care Programs: Challenges and Effective Practices” with Cynthia Lashley and Ann Culter. Gilkerson’s instructional modules on brain development for early childhood faculty, produced in collaboration with members of the Erikson Faculty Development on the Brain Project and published by Zero to Three, will be previewed.

Research associate **Carol Horton** and **Robert Halpern** coauthored “The Design and Implementation of Baltimore’s After-School Strategy: An In-Process Look,” a report and program evaluation sponsored by Safe and Sound. Horton also is coauthoring a report for the DeWitt Wallace–Reader’s Digest Fund, tentatively titled “National Evaluation of the Public Libraries as Partners in Youth Development Initiative,” as well as writing a series of research briefs for the Center for the Study of Social Policy in Washington on the



subject of how early childhood education programs can help to prevent child abuse and neglect.

**Linda Gilkerson** is cochairing the steering committee for the Illinois Early Intervention (EI) Social/Emotion Pilot, a state-funded effort to enhance social/emotional development, parent/child relationships, and family support in early intervention. Social/emotional specialists will be added at three locations of Child and Family Connections, the entry point for EI services in Illinois.

---

### **Herr Research Center at Erikson Institute**

The Herr Research Center, established in 1997 with a gift from the Herr family, is the hub of research activities at Erikson Institute. Its mission is the development of knowledge from applied research that contributes to a significant improvement in the quality, effectiveness, and equity of education and services for children and families. The center provides technical assistance and funding for the development and implementation of a wide variety of research projects, promotes the dissemination of research findings, and sponsors conferences and seminars.

Dedicated to addressing the interests and needs of an increasingly diverse society, center-supported research initiatives work with populations that vary in age, race, and ethnicity, with a primary focus on programs and populations in disadvantaged communities. The center is committed to providing a sound and useful base of information to guide the understanding of complex social issues such as changing family and societal needs and families in stress as well as the nature and efficacy of services for children and families.

### **Current research projects**

Caregiving Consensus Groups with Latina Mothers  
Children and Violence Project  
Computer Training for Early Childhood Teachers Project  
Doula Support for Young Mothers Project  
Project (in collaboration with the Department of Psychology at the University of Chicago)  
Erikson Arts Project  
Faculty Development on the Brain Project  
Fathers and Families  
Fussy Baby Network  
The Helping Relationship in Early Childhood Interventions Project  
Bridging: A Diagnostic Assessment for Teaching and Learning in Early Childhood Classrooms  
Project Match  
Reggio Emilia Project  
Schools Project  
Teacher Attitudes About Play  
The Unmet Needs Project

### **Publications available from the Herr Research Center**

Applied Research in Child Development  
Number 1, After School Programs  
Applied Research in Child Development  
Number 2, Father Care  
Applied Research in Child Development  
Number 3, Welfare Reform  
“Lessons from Beyond the Service World,”  
Judith S. Musick, Ph.D.  
“Harder Than You Think: Determining What Works, for Whom, and Why in Early Childhood Interventions,”  
Jon Korfmacher, Ph.D.  
“Child Assessment at the Preprimary Level: Expert Opinion and State Trends,”  
Carol Horton, Ph.D., and Barbara T. Bowman, M.A.  
“‘Does not.’ ‘Does too.’ Thinking About Play in the Early Childhood Classroom,”  
Joan Brooks McLane, Ph.D.

### **Faculty**

Samuel J. Meisels, Ed.D., President, Erikson Institute  
Frances Stott, Ph.D., Vice President/Dean of Academic Programs, Erikson Institute; Acting Director, Herr Research Center  
Barbara T. Bowman, M.A.  
Jie-Qi Chen, Ph.D.  
Linda Gilkerson, Ph.D.  
Robert Halpern, Ph.D.  
Jon Korfmacher, Ph.D.  
Joan Brooks McLane, Ph.D.  
Gillian Dowley McNamee, Ph.D.  
Aisha Ray, Ph.D.  
Sharon Syc, Ph.D.

### **Senior research associates**

Toby Herr, M.Ed.  
Kathleen Kostelny, Ph.D.  
Daniel Scheinfeld, Ph.D.

### **Research associates**

Leatha D. Asbury, M.Ed.  
Diana Brooks  
Rosanne DeGregorio, M.Ed.  
Jana Fleming, Ph.D.  
Carol Horton, Ph.D.  
Kate Nolan, Ph.D.  
Sandra Scheinfeld, Ph.D.  
Suzanne L. Wagner, M.A.

### **Senior research adviser**

Charles Chang, M.A.

# Erikson Institute

A Graduate School in Child Development

420 North Wabash Avenue

Chicago, Illinois 60611-5627

[www.erikson.edu](http://www.erikson.edu)