Empirical Studies of the Ounce and Work Sampling in Infancy and Early Childhood

riksor

Samuel J. Meisels Erikson Institute smeisels@erikson.edu

Co-Authors and Collaborators

WSHS • Yange Xue, Mathematica Policy Research • Melissa Shamblott, St. Paul Public Schools

Ounce • Kristy Beachy-Quick, Erikson Institute

- Xiaoli Wen, Erikson Institute
- Ilona Helin, Erikson Institute
- Lisa Michael, Erikson Institute

The best way to evaluate a child's performance is to study performance, not something else.

Curriculum-Embedded Assessments

Assessments based on information from children's typical classroom performance.

The Work Sampling System

A continuous progress, instructional assessment that helps teachers document and assess children's skills, knowledge, behaviors, and academic achievements from age 3 to Grade 6.

Preschool 4: Developmental Checklist Developmental Checklist

In this presentation Dr. Meisels will mention several assessments. He is an author of two of them—the Work Sampling System and the Ounce Scale—and a consultant to the company that publishes them, Pearson Education. These assessments will be cited solely to illustrate the points raised in the lecture, rather than to promote the assessments or imply an endorsement by any particular institution.

Preschool 4: Guidelines

The Work Sampling System Language and Literacy

A Listening

r (sor

r (sor

- 3 Demonstrates phonological awareness Demonstrates phonological awareness. Phonological awareness refers to the ability to hear and discriminate the sounds of language. Four year olds can attend to and distinguish the smaller units of sound within words with teacher support. They can begin to hear and discriminate syllables, the beginning sounds of words, and thyming sounds, prerequisite skills for being able to decode words when reading. Children show developing phonological awareness by: istening to the word the teacher says and then finding a word to thyme with it;
 begin the word of the finding hear is divide the support of the sound of the finding a word to the sound of the finding awareness the sound sound for the sound of the finding awareness the sound sound for the sound of the finding awareness the sound sound for the sound of the finding awareness the sound sound for the sound for the sound sound for the sound for the sound sound sound for the sound sound for the sound sound for the sound sound sound for the sound sound sound for the sound sound for the sound sound for the sound sound sound for the sound sound sound for the sound sound for the sound sound for the sound sound sound for the sound sound sound for the sound sound for

 - hearing the sound of the first letter in their own names and using this ability to sound out or "read" classmates' names that begin with the same letter;
 experimenting with words, giving them new beginning sounds;

 - capping out the number of syllables in their names; saying the sound of a letter when they see it in a new word because they recognize it from a familiar word (for example, the "s" sound in "stop");

 - bearing rhyming sounds, and creating rhyming words and nonsense words, such as "funny, bunny, runny, runny, sunny";
 chanting familiar rhymes with classmates during circle time or as they play.

The best way to improve a child's performance is to teach the child, not test the child.

Can Head Start and Pre-K **Teachers Use** Performance Assessments Accurately?



r (son

9rikson

Purpose of Study

To evaluate the reliability and validity of Work Sampling for Head Start (WSHS) as a measure of children's language, literacy, and mathematics.

Principal Study Questions

- 1. Is WSHS a reliable and valid means of evaluating children's achievement and progress?
- 2. What does the WSHS Checklist add to predictions of children's achievement over and above demographic variables?
- 3. Does WSHS identify children at risk for learning difficulties as accurately as normative measures?

Sample for the Study

- Pre-K: Head Start, school-, and community-based programs
- N = 112; Mean age = 4.6 years
- 16 classrooms (volunteer), 5 schools
- 62.5% African-American, 9% Hispanic; 95% low income; 54.5% male; 12% special needs

Outcome Measures

• Test of Early Reading Ability (3d edition)

ri (son

• Test of Early Mathematics Ability (3d edition)

Results of the WSHS Validity Study (Pre-K)

1. WSHS has high internal reliability (alphas = .90 – .94) and moderately high correlations between fall and spring ratings (.71 for literacy and .65 for math).

Results of the WSHS Validity Study (Pre-K), cont.

2. WSS correlates moderately well (.30 - .44) with standardized, individually administered psychoeducational tests of early reading and math (TERA and TEMA)

0	Mean Scores on Normative Tests (S.D.)							
	Fall	Spring						
TEMA-3	82.79 (13.42)	83.77 (14.41)						
TERA-3	90.27 (11.43)	90.67 (11.78)						

Results of the WSHS Validity Study (Pre-K), cont.

3. After adjusting for demographics, WSHS adds unique information (~20%) to predictions about children's early reading and math achievement.

Results of the WSS Validity Study (Pre-K), cont.

4. Data obtained from WSS accurately discriminate between children who are and are not at risk.

Conclusions - 1

)rí (son

r (sor

erikson

WSHS has excellent internal consistency and reliability within domains across time. **Conclusions - 2**

Brikson

r (Soi

WSHS is a valid and effective assessment of children's learning.



WSHS ratings accurately identify children at risk for learning difficulties in literacy and math.

Functional Assessment

Assessments that focus on everyday, naturally occurring, practical behaviors and accomplishments

The Ounce Scale Basic Characteristics

- A relationship-building tool
- Enhances parent and provider knowledge, control, and empowerment
- Helps differentiate and expand parents' and providers' perceptions of babies

Observation Record Toddlers I: 18–24 months

Toddlers I #34 months

I. Personal Connections: It's About Trust How toddless show that

How does the toddler show that the presence of famility adults helps her do things on her own? Toddlers need to see the adults who are important to them as they play. Although their isoreneed anothily allows them to be more active and solf-relian to the more active and solf-relian.

e holder might do one or more of the following: ain courage—or the slike at the pain, but only after sitting peake here caregiver for a while. The start of the slike at the pain, but only after sitting start (show—brief peaks) and the slite start of the slite slite slite slite slite slite slite slite slites, start here briefly, and then go back to the same plate. How when needs you—go of to explore something while a sufficient slite pain to explore something while a slite slite slite pools. The slite slite slites slites for slike a conforting book—giner at a caregiver for resources take a comforting book—giner at a norther noon.

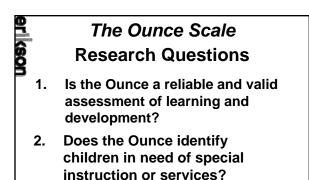
ddler show that the presence of familiar adults helps her do things? (Inclu

Developme Toddlers I: 18			
QA	Child		
The Toddlers I	Date of birth Age		
Scale Developmental Profile	Teacher/Home Visitor		
	Assessment completed by		
	Today's date		
24 Months	11.1		
SOCIAL and EMOTIONAL I. Personal Connections: It's About Trust • Seeks support of familiar adults to try things (p. 42) • Acts cautiously around unfamiliar adults (p. 43)	Comments:		
 II. Feelings About Self: Learning About Me Expresses own ideas, interests, and feelings (p. 43) Tries to manage own behavior (p. 44) 	88		
 III. Relationships With Other Children: Child to Child Watches and plays briefly with other children (p. 45) Shows awareness of other children's feelings (p. 46) 			

How accurately do Early Head Start teachers use observational assessments?

r (son

r (son



3. Does the Ounce have an impact on teaching and caregiving?

Sample for the **Validation Project**

- Birth 42 months, cross-sectional
- N = 251 children in 8 age groups
- 124 classrooms (volunteer) from 7 **Early Head Start programs**
- 63% African-American, 26% Hispanic; 92% low income; 55% male; 11% special needs (IFSP)

Outcome Assessments

- Bayley Scales of Infant Development-II Mental and Motor Scale;
- Preschool Language Scale-4;

er (sor

• ASQ:Social Emotional (completed by teachers)

Results of the Study

- erikson 1. Reliability: Moderate (overall Cronbach alpha = .65 [range = .19-.89])
 - 2. Validity: Correlations between Ounce **Developmental Profiles and criterion** measures of mental, motor, and language are moderate (.28 - .32) but higher for social-emotional (.47).

Correlations between Ounce Scales and Criterion Measures

Ounce Scale	ASQ:SE	PLS-4	Bayley-II Mental	Bayley-II Motor
Social- Emotional	.47**			
Language		.32**		
Cognition			.31**	
Motor				.28**
*p < .05. **p <.01				

Results of the Study - 2

- 3. The more seriously delayed the child, as rated by the Ounce, the less competently the child performed on the outcomes.
- 4. The Ounce significantly predicted performance over and above demographics, especially on social-emotional.

Does the Ounce Distinguish Between Children At-Risk and Not At-Risk?

More than 70% of the infants and toddlers who are at risk, as defined by the standardized assessments, were correctly identified.

"Needs Development" Rating

Teachers appeared to systematically avoid the rating, *Needs Development*, particularly for younger children in the sample.

Teacher Interview Study: Description of Sample

- 21 teachers from five EHS centers (73% response rate) + 7 program managers
- All completed at least two Observation Records and Developmental Profiles that were used in the Ounce Scale Validation Study
- Range of teachers' education: highest = Masters degree, lowest = high school
- 16 African American, 3 Hispanic, 9 White (all female)

Teacher Interview Study: Content of Questions

- 1. How teachers use the different components and ratings of the Ounce Scale
- 2. How teachers use the Ounce Scale for planning, setting goals, and working with parents
- 3. Training and support teachers were given in use of the Ounce Scale
- 4. Global questions, e.g., has the Ounce Scale changed them as teachers, do they like it?

Reasons For Not Using "Needs Development"

1. Teachers viewed this rating negatively

r (son

erikson

- 2. Teachers had difficulty applying the standards to atypical behavior
- 3. The age range may have made caregivers reluctant to say that a child needs development

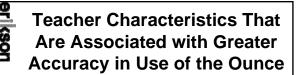
eriksor

Reasons For Not Using "Needs Development", cont.

- 4. No intermediate rating levels were available
- 5. Teachers wanted to protect parents from "bad news"

How Did the Ounce Affect Teachers and Caregivers?

- 1. Assessing children
- 2. Knowledge and understanding of child development
- 3. Setting goals
- 4. Communication with parents
- 5. Relationships with co-workers



- Higher levels of education
- More training in early care and education
- Being relatively new to working with infants and toddlers
- Longer use of the Ounce Scale

Understanding the child within context is key to understanding the child