

# **Head Start Assessment: Past, Present, and Future**

---

**Samuel J. Meisels**

**Erikson Institute**

**[smeisels@erikson.edu](mailto:smeisels@erikson.edu)**

© Samuel J. Meisels, 2005

# Milestones in Head Start Testing

---

- The Westinghouse Study (1969)
- Head Start Measures (1980s)
- FACES (1997 – present)
- Impact Study (1999 – present)
- Outcomes Framework (1999–present)
- National Reporting System (2003 – present)

# *What Policymakers Want to Know*

---

**"Are Children  
Learning?"**

---

**"Are public funds being  
used wisely?"**

# Two Myths of Head Start Testing

---

1. *School readiness can be assessed with a single test for everyone*
2. *High stakes tests are useful for evaluating early childhood programs*

# **The Myth of Measuring Readiness**

---

**Readiness can be assessed by a common set of indicators that all children will be expected to achieve by the start of school.**

# Readiness: A “Mischievous Half-Truth”

---

*The idea of readiness is a mischievous half truth . . . largely because it turns out that one teaches readiness or provides opportunities for its nurture, one does not simply wait for it.*

**Jerome Bruner, 1966**

# Readiness Within Context

---

*The appropriate policy question  
is not what children need to  
know or be able to do when  
they get to school, but what  
schools need to do to meet the  
social and educational needs of  
the children who walk through  
their doors.*

*Deborah Stipek, 2001*

# **Problems with School Readiness Tests**

---

- Early development is episodic and uneven
- Social knowledge components are typically culturally-biased
- The concept of “readiness” is relative
- Test items often imply teaching
- Content is inconsistent with teachers’ views of school success
- Validity of the tests is poor

# The Myth of High-Stakes Testing

---

The quality of a Head Start program can be evaluated by administering a high-stakes test to all children.

# **Purposes of the National Reporting System**

---

- 1. Program self-improvement**
- 2. Targeting training and technical assistance**
- 3. Program monitoring and accountability**

# **Letter Naming:**

## ***Fall 2004 NRS Data***

- ***33% of the children tested know no English letters***
- ***56% of the children tested know no Spanish letters***

# ***Psychometric Problems With The NRS***

- No external validity
- Construct underrepresentation
- Items are designed poorly ('Construct-Irrelevant Variance')
- Lack of adaptation for non-English/non-Spanish, bi-lingual, and special needs children

*“I can’t for the life of me understand why anyone would think it’s a bad idea to assess whether a program is progressing in crucial academic areas.”*

--Wade Horn, *National Journal*, 2/19/05

# **GAO Report (May 2005)**

## ***What GAO Found***

---

**The Head Start Bureau has  
not shown that the NRS ...  
results are valid measures  
of the learning that takes  
place [in Head Start.]**

# **GAO Report (May 2005)**

## ***Measurement-Driven Instruction***

---

**There is a concern that local Head Start programs will alter their teaching practices and curricula based on their participation in the NRS....at least 18% of grantees changed instruction during the first year to emphasize areas covered in the NRS.**

# **GAO Report (May 2005)**

## ***Reliability and Validity***

... the Head Start Bureau has not [validated] component tests and determin[ed] the reliability and validity of the NRS results across time....The potential exists that the NRS will produce results that are not useful for program evaluation.

# **GAO Report (May 2005)**

## ***Problems with the Spanish test***

---

**HSB cannot use the results from the Spanish version of the NRS for accountability purposes because it has not been demonstrated that this version produces reliable and valid results or that its results are comparable to those from children tested in English.**

# **GAO Report (May 2005)**

## ***Training & Technical Assistance***

---

**NRS results are of limited use  
to target training and  
technical assistance to the  
classrooms where  
assistance is needed most.**

# **GAO Report (May 2005)**

## ***Accountability***

---

**The NRS by itself does not provide sufficient information to draw conclusions about the effects of Head Start grantees on children's outcomes—information that would support use of the NRS for Head Start grantee accountability.**

# *A Better Way to Answer Policymakers' Questions*

---

*Conducting program evaluations that are faithful to the contexts of teaching and learning, instead of administering high-stakes tests to individuals*

# **Accountability Testing**

---

**Tests used to determine how well individual children are performing in relation to a normative sample.**

# ***What We Can Learn From Accountability Tests***

---

- How a particular child's performance compares to that of a normative sample.
- How much progress the child made since the last time (s)he was tested.

# **Program Evaluation**

---

**Procedures designed to demonstrate how effectively a program has achieved its goals.**

# *What We Can Learn From Program Evaluations*

---

- Which program variations may be related to child outcomes
- How the program interacts with family, community, and child contexts
- Whether the services produce anticipated outcomes

---

**High stakes  
testing does not  
promote learning  
in early childhood**

---

**Understanding  
the child within  
context is key to  
understanding the  
child**

---

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

---

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

# *The Mismeasure of Man*

---

“There are...few injustices deeper than the denial of an opportunity to strive or ever hope by a limit imposed from without, but falsely identified as lying within.”

Stephen Jay Gould (1981), p. 28