

2008 CCPRC Meeting  
Session 10 (Breakout B-3)  
July 31, 2008, 2:15-3:30 pm

## **Current Status and Challenges in Quality Measures Development: Linking Practices to Young Children's School Readiness across Domains of Development**

### **Description**

Building on the work of the April 2008 Quality Measures Roundtable, panelists discussed quality measures that assess caregiver supports for child development and school readiness, including the domains of language and literacy; math, science, and general cognition; social and emotional development; and health, safety, and nutrition. Panel members and session participants discussed the current status and limitations of current measures and areas for new measure development, for example, measures that address diverse populations. A State administrator discussed quality measures from the State perspective. Suggestions were sought about future measurement-related research and evaluation.

### **Moderator**

Wendy Robeson, Wellesley College

### **Panel Members**

Kimberly Boller, Mathematica Policy Research, Inc.  
Kathleen M. Dwyer, Office of Planning, Research and Evaluation, ACF/HHS  
Tamara Halle, Child Trends  
Susan Hegland, Iowa State University

### **Discussant**

Rod Southwick, Massachusetts Department of Early Education and Care

### **Scribe**

Laura Wandner, Child Trends

### **Documents in the Session Folder**

### **Discussion Notes**

- Recent analyses indicate that measures of specific practices may be slightly better predictors of outcomes for children than global quality measures, particularly when there is good correspondence between the practice and outcomes.
- More work is needed to identify these practices in different developmental domains.
- Measures are needed to understand diverse children in diverse settings.

### Literacy

- There are measures that exist, probably more in this domain than in others.
- Actual practices that are associated with these outcomes include:
  - Interactive book reading.
  - Conversations and real-world situations.

2008 CCPRC Meeting  
Session 10 (Breakout B-3)  
July 31, 2008, 2:15-3:30 pm

- Environmental print—amount of print.
- Availability of writing tools.
- No one measure covers all five key constructs.
- Measures for children younger than age 2 are limited.
- There are concerns about cultural and linguistic appropriateness, especially for non-native English speakers. This is one of the issues with which we had the most problems. What are best practices? There is real tension in the field on this issue. How should we address children with diverse backgrounds?

### *Questions about Literacy Measures*

Do we know from the language and literacy group if there will be analyses with a few measures that capture the key issues that need to be addressed? Are there efficient measures that we can ask States to use? Should States be using special measures or global measures?

- This is something that we need to address. The global measures are hard and expensive to use. We use them as researchers, but it would be expensive for State agency personnel to use them.
- Some tools that researchers are using are not helpful for practitioners, as some of them do not show how well a child knows his or her vocabulary.
- We never want to teach to a test.
- We are asked by States why we cannot use their assessments.

We want to get a sense of other current research regarding English-language learners (ELL). We want to know other ways to measure quality.

- Some people are working on the CSSOs and the LA 5. Mike Lopez has been creating a compendium on ELL measures.
- Anyone who is aware of quality measures that are not part of the compendium on *Research Connections* should tell Tamara Halle.

### Mathematics, Science, and Cognition

- People in the group—Sally Akins, Nikki Forry, and so forth.
- We were struck by how little research has been done in this area. We had to restrict the measures to prekindergarten measures. There is some research being done on K–3, but it is preliminary in nature.
- There are some promising approaches. Mathematica Policy Research, Inc. is doing a curriculum study in the classroom. We are looking to see what types of mathematics-related activities are occurring in the classroom. We broke this group into three groups and looked for a mix between the groups using Tamara's compendium and that of *Research Connections*'.
- How many measures are mathematics-specific? We found four measures: (1) Ellen Freedi, (2) Clements (COMET), (3) Starky and Klein, and (4) Ginsberg and colleagues. They shared a preliminary version.

- How many science-specific measures exist? We found two. (1) Lauren Van Eden reported on one—Head Start on Science. It is a little general, but it has some specific science items. (2) Kim Renemin—she has a lead on a measure.
- How many measures are general cognitions-specific? Peter Ornstein presented results that first-grade supports with mnemonic devices have outcomes at grade five. The researchers are looking at mother effects and at the effects of class on the child. He is trying to work with younger children. Diamond also has had some research and CLASS has some constructs.
  - There are some tools emerging from the laboratory on this subject.
- Combining all of them—we have about 10–15 measures with the (1) Early Childhood Environmental Rating Scale-Extension (ECERS-E) and (2) the Early Childhood Classroom Observation (ECCO).
- What are the challenges and limitations?
  - Researchers say money.
  - Researchers do not agree as to what constructs should be on the scales. The field needs to decide on a work focus.
  - Worries that providers do not have ability in the area. We need to help providers get over their fears and to help train providers to learn to teach children the mathematics skills.
- Adaptations of measures.
  - It seems like it should be fairly easy to extend the measures (e.g., CLASS). They are going to try to use the assessment with 2-year-olds.
  - What counts as science—e.g., cooking, body parts. What does this look like? These are not scary domains. We just need to think about how to measure this area.
- Going back to Peter’s work—we need to look at mother-child interactions. This issue could be a good place to look for the next steps.

### *Questions*

How do you see us measuring the quality of home-based settings?

- The general cognition piece, special development, numeracy, and exploration could be fairly well adapted. We are looking for them in the context of developmentally appropriate practices.

Are they there?

- We are worried about floor effects. I just read Pianta’s paper in *Science*. Only 25 percent of time is devoted to science in the fifth grade. We have to think about how often we will see these skills in a classroom setting. I think we are going to have problems finding associations. We will have to think about what kind of interventions will affect the settings.
- We sometimes found mathematics settings in classrooms, but we had problems finding science settings in classrooms.

2008 CCPRC Meeting  
Session 10 (Breakout B-3)  
July 31, 2008, 2:15-3:30 pm

- It seems like we used to focus more on literacy and that mathematics and science are being stressed more now which means that we may see more mathematics and science material.

How does this translate to practitioners?

- They will not call it this. Researchers may be considering math and science differently.

How are we going to encourage practitioners to stress certain things, and how should they stress them? It may just be rephrasing.

### Socioemotional

- There are few items that focus on this issue.
- There are prospects of developing future measures based on current new measures. We will want to link them to the child's socioemotional outcomes; approaches to learning, social competence, emotional competence, and disruptive behaviors; teacher interactions, routines, the physical features of the environment, and institution supports.
- They are building the matrix.
- Challenges and limitations—the constructs are not extremely well defined. The measures that are available need some work. When we talk about socioemotional development, how much of it is related to the child's relationship with the teacher or parent? We need to be more specific between the outcomes and support. There is also a challenge with the setting. Is there a difference between what we expect in home-based or center-based care or the time of day?
- Should we look at measurement between teacher report and observations?
- We need to take into consideration the child's age and cultural context. Does he or she need different things?

### *Questions*

I am wondering about the identifying analysis.

- Sometimes the children have different types of behaviors. We could set up an activity and see how the child reacts. Is this a motivation or is it desirable for assessing children? This is an open question.
- Concern for socioemotional support, what is the unit of analysis? Are observations in which we set up a structured activity with the provider and different children the appropriate way to assess children (as opposed to being there 2 days)?
- We are going to be using the emotional availability scale in the centers. The test will be a half-hour in length.
- The frequency can be assessed that way. We also want to measure the quality. Are you doing the right type of support?

My concern is how does the child's temperament affect the relationship? Can that be measured?

- If we are really interested in child's outcomes, we do not want to just measure teacher/caregiver behaviors at the group level. Each child has different interactions with the teacher. Not all the children will get the same curriculum, especially children with disabilities and ELL children.
- There are different tradeoffs when using the CLASS (the average child's experience) versus a time-sampling measure (e.g., measuring a child at a specific time—can look at the average child's experience). It is hard to get enough sweeps for the individual child. Need to think about how easy to train, cost, etc.

### Health

- Phones—physical activity, health, oral, nutrition, emotional mental health, safety. We looked to see if we possess the measures related to school readiness.
- Most of our tools are checklists; a lot of them come from Caring for Our Children.
- Primarily used in consultation as opposed to research. However, they are starting to be used for quality. People in the group—Sue Aronson, Jonathan Kotch, and Marilyn Krajicek.
- We realized that we had different definitions of how we thought about health.
- Individual items have strong research findings, such as sudden infant death syndrome research.
- Lack of interrater reliability. They use words such as “some.” Getting interrater reliability will be difficult. They do not have test-retest reliability except for some research tools.
- There is consensus that the tools need structural and process components, such as handwashing and fences around the classrooms.
- Kotch has a study that notes the need for structural and training components, for example, different diapering centers and motion-censor faucets.
- Some centers object to tools that assess things that are hard to change. However, it is important to make changes.
- Challenges and limitations:
  - Different standards for different ages.
  - Center- versus home-based care.
  - Lack of quality regarding the studies. We do not have much research connecting to health problems.
  - We lack tools with test-retest reliability.
  - We have heard that checklists are not being implemented consistently.
  - We lack internal consistency. Why would we expect handwashing to be correlated to something outdoors. We should not expect it to correlate. However that is how we learned about reliability. I have found some articles (e.g., Kim Boller) about scales. Scales have been constructed that look at the effects (e.g., depression—loss of appetite, not sleeping). The scales should correlate with each other. The health measures reflect differently; we would look at the causes of health. We should not apply the need for internal consistency with them. Kim has been

- arguing that different models should be put in places for causal questions. The scales at which I have been looking have about 50 items.
- The test is the extent to which a high score on this measure relates to child outcomes. Child outcomes that have been used include viral infections, injury incidents, and child attendance.
  - The problem is that these are low incidences. We need a large number of centers and children to look at this issue, which is also affected by home characteristics.
  - Where do we need to:
    - Build collaborations.
    - Build logic models.
    - Clarify the language on the checklists.
    - Demonstrate that we can get the tests to become reliable.
    - Build a measures compendium.
    - Do something to address this problem. What we can do depends on what we can measure.
    - Be able to show that health affects a child's ability to learn.

### *Questions*

We do not have a measure on child outcomes that measures health outcomes and statistical obesity. What is the place of the measure of the environment? Do you think that they should be part of a global measure of quality? Should we be thinking about not eliminating the health measures but developing them on their own?

- If you look at the ECERS, we are not using them now.
- I am worried about the cost, but the more reliable and consistent the measure, the better we will be able to measure and assess the effects.
- The health items can be assessed at the same time as the other domains. I am concerned that health is not being addressed. People are buying the blocks but are not focusing on the health. We need to create a logic model.
- We need to think about the index versus scale items. It is difficult to think about the environment as a possible effect. The index may be more important. That way you could look at physical health. The field has been looking more at scales. Are environments really a scale? They may be a number on an index. The ECERS is an index, not a scale. We use the ECERS because it has a high reliability.
- We have a major problem with our health measures, but we have been hesitant to move toward an index.

At the State and Territory Administrators Meeting we discussed how toxins can affect children, especially toddlers. However, there is little research on this subject. It is something that we need to look at when one assesses the environments. There is a new initiative in California to teach providers to pay attention to environments and to know what to do and what not to do.

2008 CCPRC Meeting  
Session 10 (Breakout B-3)  
July 31, 2008, 2:15-3:30 pm

Are our measures, which were created for elementary schools, accurately capturing the same issues in child care? People are avoiding this issue because there are a lot of regulations, and they are focusing on preparing children for kindergarten.

- How do you fund the implementation of a lot of these tests?
- If we are setting up a quality rating system (QRS), how do we decide on the scales, how expensive is it, and how do we improve our QRS?
- How are we able to report on these five domains and to show their effects on the children?