

2011 STAM – CCPRC Meeting
Workshop B5
November 16, 2011, 11:30 am-12:45 pm

Measuring Quality: New Findings on Thresholds and Implications

Description

Policymakers are asking for more and better evidence about the nature of the association between measures of quality in ECE settings and developmental outcomes of children in these settings. Are the associations linear, such that an increase in setting quality predicts an improvement in children's outcomes, or might there be thresholds of quality above which associations are stronger? Two projects are investigating these questions through an analysis of large-scale data sets. Margaret Burchinal presented the findings from a meta-analysis that examines threshold quality-outcome linkages in multiple studies, including the National Institute of Child Health and Human Development Study of Early Child Care, National Center for Early Development and Learning Multistate Study, and the Head Start Family and Child Experiences Survey (FACES 2006), on both cognitive and social-emotional outcomes. Julia Torquati presented threshold analyses using EHS and QUINCE data on multiple developmental measures. The two studies use different methodological approaches to uncover new patterns of association. The discussant discussed the implications of this work for decision-making at the State and national levels.

Facilitator

Louisa Tarullo, Mathematica Policy Research, Inc.

Presenters

Margaret (Peg) Burchinal, University of North Carolina at Chapel Hill
Julia Torquati, University of Nebraska – Lincoln

Discussant

Deborah Cassidy, North Carolina Department of Health and Human Services

Scribe

Rebecca Starr, Child Trends

1. Documents in Session Folder

- “Thresholds of Quality: What Does It Mean in the Trenches?” Deborah Cassidy
- “Testing Thresholds of Quality on Child Outcomes Globally and in Subgroups;” Julia Torquati, Helen Raikes, Greg Welch, Ji Hoon Ryoo, and Xiaoqing Tu
- “Factor Analyses of Quality Measures;” Xiaoqing Tu, Helen Raikes, Julia Torquati, Greg Welch, and Ji Hoon Ryoo (Handout)

2. Summary of Presentations

- **Summary of Presentation #1:** Peg Burchinal

- Peg described work in connection with OPRE-funded Q-DOT project which involved testing for thresholds in associations between child care quality and outcomes. While many experts believe that good programs can make a big difference in child outcomes, several meta-analyses suggest that the associations between quality and outcomes are very modest. One of the possible reasons may be that quality has to reach a certain point before impacts on child outcomes are observed. This is what Q-DOT is designed to study.
 - Different thresholds may exist: perhaps outcomes get better as you move lower quality programs into moderate- to high-quality range; what about high-quality programs (relation not as linear?); or maybe it's not until you get the program into the moderate- to high-quality ranges that impacts on child outcomes can be observed.
 - For Q-DOT testing, the hypothesis was that there would be stronger associations between quality measures and outcome when quality was in the higher range.
 - Testing involved secondary data analyses of large data sets (criteria for inclusion were that the data must have school readiness measures, baseline and endpoint measures for preschool-aged children, and direct assessment of classroom quality). Data sets included: FACES, Early Head Start follow-up, More-at-Four, NCEDL, PCER, My Teaching Partner, NICHD, and Miami-Dade County Literacy Intervention studies.
 - Analyses included quadratic and linear HLM models and meta-analyses across projects; using a piecewise quality model, they estimated separate linear relationships for low- and high-quality programs. Were slopes different? They also looked at effect sizes: how much change in outcomes (in SD units) can we expect with one SD increase in classroom quality?
 - The researchers used three models starting with quadratic regression analysis; if that was non-significant, they used a linear model; and finally, they used the piecewise quality approach.
 - Summary of findings: looking across studies, there is a trend, but the trend does not reach significance. There is some evidence for thresholds, especially with instructional quality (CLASS) and some evidence of domain specific effects (not replicated). Overall, the results are not a confirmation of “good enough” quality, but instead suggest that we may need to focus on getting programs up to a certain “active range” where there is a relationship to child outcomes; and that within that active range, we need to continue encouraging improvement.
 - Caution: the thresholds were established conceptually and may or may not be the best cut-points.
- **Summary of Presentation #2: Julia Torquati**
 - Building on previous research that indicate modest effect sizes between quality and child outcomes, Julia and her colleagues aimed to identify thresholds of ECE quality necessary to positively and optimally influence children's development. Their examination used two datasets: Early Head Start (EHSREP) and QUINCE. Quality measures included ITERS (EHSREP) only; ECERS-R; FDCRS; and CIS (in both family and center-based programs).

- Using GAM analysis (linear/non-linear) and spline analyses, they examined whether thresholds differ for vulnerable subgroups, anticipating that different types of thresholds might exist.
- Conclusions:
 - There were more non-linear associations than linear associations across settings and measures. Many associations between quality measures and children’s developmental outcomes are non-linear.
 - GAM analysis is a useful tool for identifying associations that are best represented by non-linear models.
 - Need to keep in mind that identified thresholds represent the beginning of a sensitive range, so it’s important to improve quality within the sensitive range.
- **Summary of Presentation #3: Discussant: Deborah Cassidy**
 - Implications of this work for QRIS and work in the trenches:
 - Cut-points used for levels in QRS (and on scales) are often arbitrary and assume linear outcomes; we need further evidence to establish better cut points in QRIS.
 - We often start too low; perhaps our starting point should be raising quality to the “active” level.
 - Cut-points can be different for different measures (types of thresholds); e.g., differences across the ERS measures.
 - Encourage continuing improvement even for programs at higher levels of quality. Often quality at higher levels has most impact on child outcomes. How do we keep them at higher levels? How do we keep making improvements?
 - Further evidence:
 - North Carolina (NC) data suggests a relationship between star level and child outcomes (knowledge perception and social skills) but only between 1, 2, and 3 stars and 4 and 5 stars (grouped together). Each star level is not meaningfully discriminating child outcomes.
 - High stakes – in NC, reimbursements are based on star level, etc. In North Carolina, only 3- to 5-star programs are eligible for subsidy dollars. It is difficult for programs to improve when lower reimbursement rates continue the cycle of low quality.
 - Remaining questions: how can we modify our QRIS to reflect the possibility of needed flexibility in cut-points based on program types, income and/or ethnicity of children? What does this information communicate about the calibration of instruments used to measure quality?

3. Summary of Discussion with Presenters and Participants

- Meta-analyses reveal a modest association between quality and child outcomes. In the past, it has been assumed that the relation is linear. Research is happening to examine the association between quality and child outcomes in a non-linear way. There is some evidence for “thresholds” of quality such that the relation between quality and child outcomes is different for higher quality programs than for lower quality programs.
- This has implications for QRIS and how to target quality improvements for programs.

4. Discussion

- QRS levels focus on structural aspects/standards. When is it appropriate to expect changes in child outcomes? When can we look at the associations between quality and child outcomes? What is it about social-emotional outcomes? Are they more unstable as outcomes?
- Across studies, we need to think about how to move programs to the higher quality range. Process aspects are harder to measure. What is it that gets programs there? It is not education. We don't really have a clear understanding of what it takes to improve quality at high range.