

MEASURING THE QUALITY OF EARLY CARE AND EDUCATION PROGRAMS AT THE INTERSECTION OF RESEARCH, POLICY, AND PRACTICE



Research-to-Policy, Research-to-Practice Brief OPRE 2011-10a
December 2010



Measuring the Quality of Early Care and Education Programs at the Intersection of Research, Policy, and Practice

Research-to-Policy, Research-to-Practice Brief OPRE 2011-10a December 2010

Submitted to:

Ivelisse Martinez-Beck, Ph.D., Project Officer
Office of Planning, Research and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services

Submitted by:

Martha Zaslow, Child Trends
Kathryn Tout, Child Trends

Contract Number: HHSP233200500198U

CONTRACTOR

Project Director: Martha Zaslow, Child Trends
4301 Connecticut Ave NW
Washington DC, 20008

Suggested Citation:

Zaslow, M., Tout, K., & Martinez-Beck, I. (2010). *Measuring the Quality of Early Care and Education Programs at the Intersection of Research, Policy, and Practice*, OPRE Research-to-Policy, Research-to-Practice Brief OPRE 2011-10a. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

This Research-to-Policy, Research to-Practice brief series focuses on issues related to the development and refinement of measures to assess the quality of early childhood settings. The views represented in this brief do not necessarily reflect the opinions of the Office of Planning, Research and Evaluation of the Administration for Children and Families.



OVERVIEW. Quality measurement in early care and education programs now plays a central role in early childhood policy and program initiatives such as state Quality Rating and Improvement Systems (QRIS). This Brief discusses the changing context of quality measurement and delineates a set of challenges that must be addressed as measures are used for purposes such as making ratings of quality levels widely available, and using quality ratings to make funding and programmatic decisions. These challenges include: alignment of measures to program goals, development of guidelines to ensure reliable measurement, weighting multiple components of quality, validating measurement across the full range of quality and across types of early childhood settings, cost-effectiveness, and developing an appropriate data collection infrastructure.



In developmental research, there is a long tradition of measuring the quality of children’s early care education environments to identify the factors that contribute to quality and, in turn, to understand how quality contributes to children’s development (Burchinal, 2010, this series). Some early care and education practitioners have also used measures of quality as self-assessment tools as they engage in quality improvement activities (Buysse & Wesley, 2005; Mashburn, Pianta, Downer & Hamre, 2007; South Carolina ECERS Quality Study Team, 2006; Smith, Sarkar, Perry-Manning, & Schmalzreid, 2006). It is only recently, however, that the measurement of quality has taken on a new role as central in early childhood policy. The purpose of this Brief is to describe how quality measures are now being used in policy and programmatic decisions, and to delineate a set of challenges that must be addressed about using quality measures in these new contexts.

It is important first, however, to briefly note how quality has been defined and measured. Reviews of research on the measurement of quality and associations between quality and child outcomes (Friedman & Amadeo, 1999; Halle, Vick & Anderson, 2010; Vandell and Wolfe, 2000; Gootman & Smolensky, 2003), identify two major approaches to defining and measuring quality: (1) *process quality*, involving interactions with children and structuring of the environment so that care is emotionally responsive, stimulating and safe; and (2) *structural quality*, involving “regulable” features of the environment, such as group size, ratio, and staff qualifications, that can increase the likelihood of positive process quality. These two approaches to quality apply to both center-based care (including child care centers, pre-kindergarten, preschools and Head Start), and home-based care (including licensed family child care and family, friend and neighbor care), though the particular ways in which they are manifested and measured differ by setting as well as by age of child (infants and toddlers; preschool-age children).

Policy and practice statements of such national organizations as the National

Association for the Education of Young Children (NAEYC) and the American Academy of Pediatrics provide descriptions of high quality early childhood settings. For example, NAEYC’s quality standards for the accreditation of center-based settings focus on assuring that young children have experiences that promote positive developmental outcomes. Paralleling the dual focus on process and structural quality noted above, the standards address either children’s immediate experiences in early childhood settings (through standards for relationships, curriculum, teaching, assessment of child progress and health) or facets of quality that provide “an effective and durable support structure for a quality program” (through administration, including fiscal environment, leadership and management; and partnerships, including relationships with families and community relations) (NAEYC, 2005, p. 7).

The Changing Context of Quality Measurement

In the past decade, public and private funding for early care and education has reached unprecedented levels (Adams, Tout & Zaslow, 2007). Over this same period, concerns have grown about the quality of early care and education settings and their capacity to support school readiness among children most at-risk for problems as they enter school. As a result of these concerns, states and local communities are investing a modest but significant portion of early care and education funding in initiatives aimed at quality improvement, as well as at educating parents about how to find high quality care (Pittard, Zaslow, Lavelle & Porter, 2006).

One increasingly prevalent approach to informing consumers involves the development of state or local quality rating and improvement systems (QRIS), in which different facets of quality are summarized into overall ratings indicating differing levels of quality (Tout, Starr, Soli, Moodie, Kirby & Boller, 2010; Tout, Zaslow, Halle & Forry, 2009). Currently, over 20 states have a statewide QRIS in place and numerous others are conducting pilots or actively exploring implementation of a QRIS (National Child Care Information and Technical Assistance Center, 2010).



Measuring quality in a consistent and accurate way is central to the success of quality improvement initiatives. Measurement of quality can help policymakers diagnose problems, set goals for improvement initiatives, allocate resources toward the most pressing needs and track change over time. A number of states, as well as federal programs such as Head Start, have used quality measurement for these diagnostic and descriptive purposes (U.S. Department of Health and Human Services, 2003; Barnard, Smith, Fiene & Swanson, 2006; South Carolina ECERS Quality Study Team, 2006; Tout & Sherman, 2005). In addition, quality measurement can be used in evaluating the effectiveness of quality improvement initiatives or of early childhood programs such as state pre-kindergarten (Barnard et al., 2006; Frede, Jung, Barnett, Lamy & Figueras, 2007).

In some states and local communities, QRIS ratings are used not only for providing information to parents but also for determining the level of reimbursement by child care subsidy agencies or other public funding streams (Stoney, 2004; Tout & Zaslow, 2004). These conditions, with financial incentives and consequences linked to quality ratings, clearly require precise measurement at the level of the individual program. Yet, it is unclear whether the level of precision in measurement needed in this context differs from the precision required for averaging quality across settings in a research study, and if so, what the guidelines should be. This context of policy and accountability in which quality measurement now occurs requires an examination of measures in a new light.

New Measurement Challenges

Aligning Measures with Goals

A first challenge in the measurement of quality in a policy context is the selection of measures that map onto goals for quality measurement (Zaslow, Tout, Halle & Forry, 2009). States that have developed QRIS have started the process of development by drawing together a group of stakeholders to discuss the goals of their program, the aspects of quality they will emphasize, and the core components they will include in distinguishing different levels of quality (Mitchell, 2005).

As more states have developed QRIS, a difference is emerging regarding the relative emphasis placed on supporting children's progress towards school readiness. States developing their systems more recently have focused increasingly on instruction as a facet of quality and a goal of supporting children's school readiness in addition to a focus on global quality (Child Trends, 2008; Tout, Zaslow, Halle & Forry, 2009).

This distinction in goals has important implications for the selection of quality measures. Measures of quality developed more recently, such as the Classroom Assessment Scoring System (CLASS; Pianta, La Paro & Hamre, 2007) and the Early Literacy and Language Classroom Observation (ELLCO; Smith, Brady, & Anastasopoulos, 2008) focus much more on instructional quality and stimulation for early language and literacy development, whereas measures developed earlier, such as the environment ratings scales (ERS) (Harms, Clifford & Cryer, 2005; Harms, Cryer & Clifford, 2007) focus more broadly on global quality. Though there is some emerging evidence that more instructionally focused measures predict more precisely to early school outcomes (Burchinal et al., 2009; Mashburn et al., 2008), these measures of quality miss some key features of the environment (such as health and safety features) that may be quite important if the articulated goal is assuring an environment that supports development broadly. Further, measures focusing on instructional quality thus far have been developed only for three- and four-year-old children. They have only recently been extended to include home-based settings (Neuman, Dwyer & Koh, 2007).

Local vs. Universal Conceptualization of Quality

As states draw together key stakeholders and develop their own specific approaches to measuring quality, a question that is emerging is whether quality should be seen as context specific. Should different measurement strategies be used in different states, in keeping with state demographics, goals for school readiness, and existing oversight and support for quality improvement? Or should the conceptualization and measurement of quality be uniform across the country?

Examination of the content and measurement approaches in different state QRIS suggests that there are both similarities and differences. For example, many include observational measures of quality as a core component. However, states differ in terms of the number of quality indicators included in their rating systems, what is required for a rating of good or excellent quality, and how much weight is given to observational measures as opposed to measures that can be derived from administrative data or self report (Tout et al., 2010). It is important to know if states are showing minor differences with the same underlying conceptualization of quality, or whether there are significant substantive differences.

Financial Consequences and Strength of Prediction to Child Outcomes

Challenges are also emerging in terms of the evidence base that is necessary and sufficient for tying financial consequences to different quality ratings. A recent meta-analysis of early childhood studies and secondary analysis of major early childhood datasets concluded that while measures of observed quality are consistently related to child outcomes, the strength of the relationship is modest (Burchinal et al., 2009). The conclusion that observed quality and child outcomes show a modest relationship is not a new one, but one that has now been shown consistently across multiple data sources and analysis approaches (NICHD ECCRN & Duncan, 2003; NICHD ECCRN, 2006).

One interpretation of this finding concerns measurement quality: the possibility that measures of quality do not provide sufficient detail on the aspects of quality most closely related to particular child outcomes. The ORCE was carefully developed by experts in the field for the NICHD Study of Early Child Care and Youth Development to help address this issue, yet continues to show modest prediction to child outcomes (NICHD ECCRN & Duncan, 2003). However, a rapidly growing body of experimental evaluations of approaches to improving quality in early childhood settings (Zaslow, Tout, Halle, Vick, & Lavelle, 2010) is providing a foundation for new efforts to strengthen the measurement of quality, identifying which facets of quality can be changed and how they are related to child outcomes (Forry, Vick & Halle, 2009). Further, new evidence that quality in early childhood settings predicts outcomes all the way through adolescence (Vandell et al., 2010) raises the possibility that a small effect that is extremely consistent across studies and persistent over years of development may nevertheless be of importance. It is important to continue to focus on this issue, asking both whether the measurement of quality can be further strengthened, and at the same time, whether small but consistent and persistent effects suffice as a basis for quality ratings with financial incentives and consequences.

Evidence Underlying Distinctions Among Quality Levels

Mitchell (2005) has noted that QRIS first grew out of recognition of the limitations of “tiered reimbursement” policies. Tiered reimbursement is an approach within the child care subsidy system in which higher reimbursements are given to settings that have received recognition of high quality through an accreditation system. Mitchell notes that some states with tiered reimbursement policies found that fewer programs than expected qualified for the higher reimbursement rates because fewer than anticipated were accredited. These states identified a need for incentives that would apply to steps in quality between licensing and accreditation. In addition, the incentive provided through tiered reimbursement applied only to settings serving subsidy-receiving families. QRIS were designed to extend the availability of incentives further.

Most states continue to use accreditation in their QRIS, though the way in which accreditation is recognized in the systems varies (Tout et al., 2010). As states work to define rungs on the ladder of quality between licensing and accreditation, it will be important to use research approaches to examine the distinctions made (FPG, 2001; Tout et al, 2009). In addition, though the environment rating scales make distinctions in terms of the adequacy of care (inadequate, minimal etc), these distinctions are theoretically-based, especially using descriptions of developmentally appropriate practice as a foundation. The application of measures of quality to a policy context underscores the need for empirical approaches that explicitly examine thresholds. New work taking these approaches is now in process (Child Trends, 2006; 2007).

Weighting the Components of Quality

A further issue is that the environment rating scale score often represents just one piece of a quality assessment. In an effort to create a comprehensive portrait of quality, state QRIS typically assess a variety of quality dimensions that might include indicators of staff professional development, use of a curriculum, administrative practices, and child-adult ratio (Tout et al., 2010). The algorithm for scoring each of these component dimensions and combining them into a composite rating varies significantly across state systems. It is important to ask whether the components as well as the composite ratings in each system are related to differences in children's experiences and development (Weber & Wolfe, 2003).

Validating Measurement Across Types of Settings

States are also challenged to create and validate the systems of quality measurement they have developed for different types of care settings. Early care and education is a diverse and fragmented system with multiple types of settings available through the market (including center-based care, licensed family child care, and paid care by family, friends and neighbors) and settings available through public programs (such as state funded pre-kindergarten and Head Start) (Adams, Tout & Zaslow, 2007; Brandon & Scarpa, 2006). Many young children are also cared for in home-based settings where the care is not formally paid for.

In working to extend QRIS to apply across types of early care and education, a number of measurement issues emerge. First, there is currently a wider range of options for observing quality in center than in home-based care setting (Goodson & Layzer, 2010, this series; Halle & Vick, 2010). Further, it is only recently that any measures have been developed explicitly for family, friend and neighbor care provided in home-based settings (Porter, Rice & Rivera, 2006). And, even when measures of quality are available for home-based settings (regulated and unregulated), the alignment of the ratings for home-based settings and center-based settings is difficult. Finally, the majority of measures of quality focusing on instructional practice that have been developed thus far are only appropriate for center-based settings. Thus, states that have decided that a goal of the quality rating system is to support children's school readiness face few measurement options for home-based care.

Building an Appropriate Infrastructure for Data Collection

The measurement of quality in a policy context requires the development of an infrastructure to support reliable and valid data collection. For direct observation, decisions must be made about how often observations need to be conducted. For center-based settings, decisions must be made about the proportion of classrooms that need to be observed and the process that will be used to aggregate the ratings for a center-wide score (Stoney, 2004). In a study in Missouri, researchers found that using ratings from half of the classrooms in a center produced a better "match" with the average quality of all classrooms than selecting one-third of the classrooms (Mauzy & Thornburg, 2007). As more states begin to collect evidence on these issues, it will be important to share the results widely.

Decisions also need to be made about the training of observers and procedures for reaching and maintaining reliability (Mitchell, 2005; Stoney, 2004). Researchers do expect and accept some error in data collection. For example, inter-rater reliability at 85% agreement is acceptable in a research study. Yet, this level of error may not be acceptable in a high stakes context, where a difference of 1 point on an environment rating scale may mean a difference in the number of “stars” assigned to the setting and/or a level of compensation. A further question is whether reliability should be assessed only for overall ratings of quality, or whether reliability needs to be confirmed for specific subscales or components of observational measures. The answer may differ according to whether a QRIS uses the overall summary score or specific components in contributing to ratings. Many states have built appeals processes into their systems to accommodate complaints about scoring, but a longer-term, sustainable solution and recommendation for the issue of data collector reliability would be welcomed by researchers and policymakers alike.

Guidelines on verification of self-report data are also needed as part of state data collection systems. Most states face budget constraints in the implementation of their QRIS, and self-report data cost significantly less than observational data to collect (Brandon & Scarpa, 2006). For data collection relying on self-report by participants in the settings (directors, teachers, family child care providers), a system for review of the information is needed (Brandon & Scarpa, 2006; Weber & Wolfe, 2003) as there are questions about the reliability of self-report data. As one example, it may be difficult for providers/educators to recall accurately the number of hours of training they have completed within the past year. Further, there is concern about the potential for systematic self-report bias when consequences are attached to ratings. As with the issue of the frequency of direct observation, there is a need for methodological work examining which self-report data may be most important to collect and the practices that can be used to review and verify data.

Conclusion

In conclusion, as the measurement of quality goes to scale in quality initiatives like QRIS, there is a need to reexamine quality measures to ask how they are functioning in a policy context. We need information on what aspects of quality existing measures cover well and what they do not cover well, on the reliability of data collection at this scale, and the extent to which measures predict to the child outcomes that early childhood policies are focusing on. The system-wide implementation of quality measurement also increases the salience and importance of the availability of psychometrically strong measures of quality for all types of early care and education.

The authors thank Nancy Eisenberg and anonymous reviewers for their extremely helpful comments on earlier drafts of this research brief when under review by *Child Development Perspectives*. These comments were valuable in strengthening the brief.

References

- Adams, G., Tout, K., & Zaslow, M. (2007, May). *Early care and education for children in low-income families: Patterns of use, quality, and potential policy implications*. Prepared for the Urban Institute and Child Trends Roundtable on Children in Low-Income Families. Washington, DC.
- Barnard, W., Smith, W. E., Fiene, R., & Swanson, K. (2006). *Evaluation of Pennsylvania's Keystone STARS quality rating system in child care settings*. Pittsburgh, PA: University of Pittsburgh School of Education, Office of Child Development.
- Brandon, R. N., & Scarpa, J. P. (2006). *Supply, demand and accountability: Effective strategies to enhance the quality of early learning experiences through workforce improvement*. University of Washington, Human Services Policy Center.
- Burchinal, P., Kainz, K., Cai, K., Tout, K., Zaslow, M., Martinez-Beck, I. & Rathgeb, C. (2009). *Early care and education quality and child outcomes*, OPRE Policy Brief. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Busse, V., & Wesley, P. W. (2005). *Consultation practice in early childhood settings*. Baltimore: Paul H. Brookes.
- Child Trends. (2006). *Roundtable on Measuring Quality in Early Childhood and School-Age Settings: At the Junction of Research, Policy and Practice. Executive Summary*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Child Trends. (2008). *Meeting on the Evaluation of State Quality Rating Systems: Meeting Summary*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Forry, N., Vick, J., & Halle, T. (2009). *Evaluating, Developing, and Enhancing Domain-Specific Measures of Child Care Quality*, OPRE Policy Brief. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- FPG UNC-CH Smart Start Evaluation Team. (2001). *Validating North Carolina's 5-star child care licensing system. Smart Start*. Chapel Hill, NC: Frank Porter Graham Child Development Center.
- Frede, E., Jung, K., Barnett, W. S., Lamy, C. E., & Figueras, A. (2007). *The Abbott Preschool Program Longitudinal Effects Study (APPLES). Interim Report*. New Brunswick, NJ: National Institute for Early Education Research.
- Friedman, S. L. & Amadeo, J. A. (1999). Assessments of the child care environment and experience. In S. L. Friedman & T. D. Wachs (Eds.), *Assessment of the Environment Across the Lifespan*. Washington, DC: American Psychological Association Press.
- Goodson, B. D., & Layzer, J. I. (2010). *Defining and Measuring Quality in Home-Based Care Settings*, OPRE Research-to-Policy, Research-to-Practice Brief OPRE 2011-10a Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Gootman, J. A., & Smolensky, E. (Eds.) (2003). *Working families and growing kids: Caring for children and adolescents*. Committee on Family and Work Policies, Board on Children, Youth, and Families, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academies Press.

Halle, T., Vick, J. E. & Anderson, R. (2010). *Quality in Early Childhood Care and Education Settings: A Compendium of Measures (2nd Edition)*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Harms, T., Clifford, R. M., & Cryer, D. (2005). *Early Childhood Environment Rating Scale - Revised Edition*. New York, NY: Teachers College Press.

Harms, T., Cryer, D., & Clifford, R. M. (2007). *Family Child Care Environment Rating Scale - Revised Edition*. New York, NY: Teachers College Press.

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., & Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development, 79*(3), 732-749.

Mashburn, A. J., Pianta, R. C., Downer, J., & Hamre B. K. (2007, March). *My Teaching Partner: Effects of a Web-Based Intervention to Improve Teacher Quality*. Poster session presented at the biennial meeting of the Society for Research in Child Development, Boston, MA.

Mauzy, D., & Thornburg, K. R. (2007, March). *Developing a state Quality Rating System based on research*. Presentation at the NAEYC Public Policy Forum, Washington, DC.

Mitchell, A.W. (2005). *Stair steps to quality: A guide for states and communities developing quality rating systems for early care and education*. United Way Success by 6.

National Association for the Education of Young Children. (2005). *Early Childhood Program Standards and Accreditation Criteria*. Washington, DC: NAEYC.

National Child Care Information and Technical Assistance Center [NCCIC]. (2010). *Quality Rating and Improvement System Resource Guide*. Washington, DC: Administration for Children and Families, U.S. Department of Health and Human Services.

National Institute of Child Health and Human Development Early Child Care Research Network. (2006). Child care effect sizes for the NICHD Study of Early Child Care and Youth Development. *American Psychologist, 61*, 99-116.

National Institute of Child Health and Human Development Early Child Care Research Network, & Duncan, G. J. (2003). Modeling the impacts of child care quality on children's preschool cognitive development. *Child Development, 74*(5), 1454-1475.

Neuman, S., Dwyer, J., & Koh, S. (2007). *Child/Home Early Language & Literacy Observation Tool (CHELLO)*. Baltimore, MD: Brookes Publishing.

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2007). *Classroom Assessment Scoring System*. Baltimore, MD: Brookes Publishing.

Pittard, M., Zaslow, M., Lavelle, B., & Porter, T. (2006, April). *Investing in Quality: A Survey of Child Care and Development Fund Initiatives*. Washington, DC and New York, NY: National Association of State Child Care Administrators and Child Trends with the Bank Street College of Education.

Porter, T., Rice, R., & Rivera, E. (2006). *Assessing quality in family, friend, and neighbor care: The child care assessment tool for relatives*. New York: Bank Street College of Education.

Smith, M. W., Brady, J. P., & Anastasopoulos, L. (2008). *Early Language and Literacy Classroom Observation (ELLCO)*. Baltimore, MD: Brookes Publishing.

Smith, L. K., Sarkar, M., Perry-Manning, S., & Schmalzreid, B. (2006, November). *NACCRRRA's National Survey of Child Care Resource and Referral Training: Building a training system for the childcare workforce*. National Association of Child Care Resource and Referral Agencies: Washington, DC.

South Carolina ECERS Quality Study Team. (2006, March). *Analyses of ECERS-R Data and Teachers' Attitudes Toward the ECERS-R Process. Final Report*. The South Carolina Classroom Quality Research Project. Greensboro, NC: SERVE Center at the University of North Carolina at Greensboro.

Stoney, L. (2004). *Financing quality rating systems: Lessons learned*. United Way, Success by 6.

Tout, K., & Sherman, J. (2005). *Inside the preschool classroom: A snapshot of quality in Minnesota's child care centers*. Child Trends and the Minnesota Child Care Policy Research Partnership. St. Paul, MN: Minnesota Department of Human Services.

Tout, K., & Zaslow, M. (2004). *Tiered reimbursement in Minnesota child care settings*. Child Trends and the Minnesota Child Care Policy Research Partnership. St. Paul, MN: Minnesota Department of Human Services.

Tout, K., Zaslow, M., Halle, T., & Forry, N. (2009). *Issues for the Next Decade of Quality Rating and Improvement Systems*, OPRE Issue Brief. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

U.S. Department of Health and Human Services (2003, May). *Head Start FACES 2000: A whole-child perspective on program performance, Fourth Progress Report*. Washington DC: Administration for Children and Families, U.S. Department of Health and Human Services.

Vandell, D. L., Belsky, J., Burchinal, M., Steinberg, L., Vandergrift, N., and the NICHD Early Child Care Research Network (2010). Do effects of early child care extend to age 15? Results from the NICHD Study of early child care and youth development. *Child Development, 81*, 737-756.

Weber, R. J., & Wolfe, J. (2003). Improving child care: Providing comparative information on child care facilities to parents and the community. *Child Care Policy Research Issue Brief*. Oregon Child Care Research Partnership, Oregon State University:

Zaslow, M., Tout, K., Halle, T., & Forry, N. (2009). *Multiple Purposes for Measuring Quality in Early Childhood Settings: Implications for Collecting and Communicating Information on Quality*, OPRE Issue Brief. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Zaslow, M., Tout, K., Halle, T., Vick, J., & Lavelle, B. (2010). *Towards the identification of features of effective professional development for early childhood educators: A review of the literature*. Submitted to U.S. Department of Education. Washington, DC, Child Trends.

Tout, K., Starr, R., Sol, M., Moodie, S., Kirby, G., Boller, K. (2010). *Compendium of Quality Rating Systems and Evaluations*, OPRE Report. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Overview for OPRE Research Brief Series on Measuring Quality in Early Care and Education Settings

Measures to assess the quality of early care and education environments, originally developed as research tools and, in some cases, as guides for improving practice, now play a prominent role in the early childhood policy arena. Many states use information from on-site observations and environmental rating scales to make decisions about inclusion of programs in publicly funded initiatives and interventions, to target quality improvement dollars and to target incentives when programs meet higher quality standards. To date, the majority of states that have developed statewide Quality Rating Systems combine scores on observational measures of quality with other quality indicators to provide a rating that is available to the public. The intent is to provide better information to parents, and to provide a framework within which quality benchmarks, financial support, technical assistance, and monitoring create leverage for quality improvements in early care and education.

Yet the use of quality measures in “high-stakes” policy and programmatic decisions raises important new questions about their content, reliability, validity, and applicability with diverse populations across a broad range of settings. To address these questions, the Office of Planning, Research and Evaluation in the Administration for Children and Families of the U.S. Department of Health and Human Services and other federal partners convened a meeting of researchers, state policymakers, practitioners and other key stakeholders. The meeting provided a forum for analyzing current quality measures, engaging in critical discussion about the use of quality measures in the policy arena, and outlining the steps needed to improve measurement strategies.

The four coordinated research briefs in this series were developed based on presentations made at the meeting, with the intent of informing policymakers, researchers and practitioners about new developments in quality measurement being generated at the intersection of child development research and early childhood policy.

- The first paper (by Martha Zaslow, Kathryn Tout and Ivelisse Martinez-Beck) describes why and how quality measures are currently used in policy and practice contexts and the issues and concerns that arise as a result of this widespread use.
- The second paper (by Margaret Burchinal) reviews the literature on the dimensions of quality that have been measured in early care and education settings and identifies the quality dimensions that have received a more thorough treatment in the literature compared to those that have not been studied as extensively.
- The third and fourth papers review the quality measures that have been developed for use in center-based early care and education programs (paper by Donna Bryant) and home-based settings (paper by Barbara Goodson and Jean Layzer). In addition to highlighting the types of measures used, their psychometric properties, and their value in predicting child outcomes, the authors discuss the importance of the findings for policymakers and practitioners.

Overall, we hope that the four papers provide a useful review of the current state of the field of quality measurement and suggest important next steps that policymakers, researchers, and practitioners can take to assure the integrity of measurement strategies and the appropriate use of data on the quality of early care and education settings especially when measures are widely implemented in policy and practice initiatives.

Those interested in the issue of the measurement of quality in early childhood settings may also want to read these OPRE briefs:

Burchinal, P., Kainz, K., Cai, K., Tout, K., Zaslow, M., Martinez-Beck, I. & Rathgeb, C. (2009). Early care and education quality and child outcomes. OPRE Research-to-Policy Brief. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, and Child Trends.

Tout, K., Zaslow, M., Halle, T., & Forry, N. (2009). Issues for the next decade of quality rating and improvement systems. OPRE Issue Brief. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, and Child Trends.

Zaslow, M., Tout, K., Halle, T., & Forry, N. (2009). Multiple purposes for measuring quality in early childhood settings: Implications for collecting and communicating information on quality. OPRE Issue Brief. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, and Child Trends.