MEASURING THE QUALITY AND QUANTITY OF IMPLEMENTATION IN EARLY CHILDHOOD INTERVENTIONS

Research Brief OPRE 2013-12
April 2013

Disclaimer

The research reported herein was supported in part by the Institute of Education Sciences, U.S. Department of Education, through grant R305A060021 to the University of Virginia - funding the National Center for Research on Early Childhood Education (NCRECE). This paper was prepared under OPRE's **Child Care and Early Education Policy and Research Analysis and Technical Expertise Project** with Child Trends (contract #GS10F0030R). The opinions and any factual errors expressed in the paper are those of the authors and not of the University of Virginia, Frank Porter Graham Child Development Institute, or of Child Trends, nor does the content of this publication necessarily reflect the views or policies of the U.S. Department of Health and Human Services or the U.S. Department of Education.

Acknowledgements

Jason Downer, Ph.D. is Director of the Center for Advanced Study of Teaching and Learning at the University of Virginia. Noreen Yazejian, Ph.D., is a Scientist at the Frank Porter Graham Child Development Institute, University of North Carolina at Chapel Hill.

Measuring the Quality and Quantity of Implementation in Early Childhood Interventions

Research Brief OPRE 2013-12
April 2013

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:
Jason Downer, University of Virginia
Noreen Yazejian, Frank Porter Graham Child Development Institute, University of North Carolina at Chapel Hill

Contract Number: GS10F0030R
CONTRACTOR
Project Director: Tamara Halle
Child Trends
7315 Wisconsin Avenue
Suite 1200 West
Bethesda, MD 20814

Downer, J. & Yazejian, N. (2013). *Measuring the quality and quantity of implementation in early childhood interventions* (OPRE Research Brief OPRE 2013-12). Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Child TRENDS

OPRE

UNIVERSITY of VIRGINIA
CASTL
CURRY SCHOOL OF EDUCATION

UNC
FRANK PORTER GRAHAM
CHILD DEVELOPMENT INSTITUTE

Overview for OPRE Research Brief series on Applying Implementation Science to Early Care and Education Research and Evaluation

The "science of implementation" is the study of the process of implementing programs and practices that have some evidence from the research field to suggest they are worth replicating. Implementation science is the study of how a practice that is evidence-based or evidence-informed gets translated to different, more diverse contexts in the "real world." In this way, effective implementation bridges the gap between science and practice.

There is a growing body of research looking at the processes and core components of implementing evidence-based practices to different settings and, especially, at what it takes to move an evidence-based practice from the laboratory to the field (Berkel, Mauricio, Schoenfelder, & Sandler, 2010; Durlak & Dupre, 2008; Fixsen, Naoom, Blase, Friedman, & Wallace, 2005; Meyers, Durlak & Wandersman, 2012). However, historically much of this research has focused primarily on adult services (Simpson, 2002) rather than on services for young children and evidence-based practices that support young children's growth and development.

The salience of implementation has come to the fore within the early childhood field in recent years because, increasingly, early childhood program developers are being asked not only to prove their program's efficacy before being brought "to scale" or transported to other locations, but also asked to articulate what components of their model, or the contexts in which the model is deployed, are essential for making the intervention a success. This is true of individual programs, such as discrete language and literacy interventions, as well as for larger, systems-level interventions, such as statewide initiatives to improve early childhood educators' professional development, children's school readiness, or child care quality. However, up until now, the early childhood field has lacked a common framework and language with which to examine important implementation supports for successful initiatives.

This research brief series seeks to provide early childhood researchers, program developers, and funders with an introduction to implementation frameworks and promising practices in implementation science with the aim of facilitating their use in early care and education research and program evaluation.

- The first two briefs in this series lay the groundwork for understanding the principles and frameworks of implementation science and provide a common language for key terms and constructs used throughout the research brief series. Specifically, a brief by Allison Metz, Sandra Naoom, and Tamara Halle introduces key elements of effective implementation within an integrated, stage-based framework; and a brief by Eboni Howard, Lindsey Allard Agnamba, Julia Wessel and Victoria Rankin provides a review of the terminology used in implementation research in the early care and education literature.
- The third brief (by Jason Downer and Noreen Yazejian) defines two cross-cutting themes: the quality and quantity of implementation. A review of recent empirical work provides examples of how these constructs are assessed and examined in relation to early care and education program outcomes. The authors highlight implications for researchers, purveyors, and funders of early childhood programs.
- The fourth brief (by Barbara Wasik, Shira Kolnik Mattera, Chrishana Lloyd, and Kimberly Boller) uses an implementation science lens to help readers understand the effects that dosage of interventions can have on outcomes, as well as on general implementation factors such as training and program administration.

- The fifth brief (by Diane Paulsell, Anne M. Berghout Austin, and Maegan Lokteff) introduces the importance of measuring implementation at multiple system levels and proposes tools for doing so. The benefits for practitioners, researchers, and policymakers of measuring implementation at multiple system levels are conveyed and suggestions and practical considerations are offered.
- The sixth brief (by Amy Susman-Stillman, Shannon B. Wanless, and Christina Weiland) reviews three theoretical frameworks of fidelity from the fields of prevention science, clinical psychology, and elementary education; highlights useful aspects of each framework; and offers early care and education researchers considerations for choosing a framework to use in their studies.

Implementation science offers a means by which to create a shared understanding of what it takes to have effective, replicable, and sustainable early childhood programs and systems in community-based settings. This research brief series aims to provide a useful overview of the current state of the field of implementation science research and its applications to the early care and education field. We hope that researchers, program developers, funders and other stakeholders will find this series helpful in facilitating the use of implementation science frameworks, methodologies, and analysis in early care and education research and program evaluation.

This research brief series may be found at http://www.acf.hhs.gov/programs/opre/research/project/child-care-and-early-education-policy-and-research-analysis-and-technical.

Berkel, C., Mauricio, A. M., Schoenfelder, E., & Sandler, I. N. (2010). Putting the pieces together: An integrated model of program implementation. *Prevention Science. 12*, 23-33.

Durlak, J. A. & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*, 327-350.

Fixsen, D., Naoom, S., Blase, K., Friedman, R., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, National Implementation Research Network. (FMHI Publication No. 231).

Meyers, D. C., Durlak, J. A., & Wandersman, A. (2012). The quality implementation framework: A synthesis of critical steps to the implementation process. *American Journal of Community Psychology.* Advanced Online Publication. doi 10.1007/s10464-012-9522-x

Simpson, D. D. (2002). A conceptual framework for transferring research to practice. *Journal of Substance Abuse Treatment, 22*(4), 171-182.

MEASURING THE QUALITY AND QUANTITY OF IMPLEMENTATION IN EARLY CHILDHOOD INTERVENTIONS

Overview

In this brief, we define quantity and quality of implementation and embed these terms in conceptual frameworks and theory on how variation in implementation is linked with key program outcomes; provide examples from a brief review of recent empirical work of how these constructs are examined in published work; and convey the benefits of measuring both the quantity and quality of implementation to researchers, practitioners, and policymakers.

Quantity of implementation measures focus on capturing straightforward, objective counts of interventionist or participant behavior. Examples of quantity measures include dosage (amount of an intervention), intensity (how much of an intervention is delivered during a session), frequency (how often intervention is delivered), and adherence (proportion of intervention components delivered). Alternatively, implementation quality measures examine the level of skill shown by an interventionist (e.g., coach/mentor, supervisor, teacher) in delivering an intervention. Examples of quality measures include how well an interventionist delivers the intervention (e.g., ability to engage participants, pacing, developmental appropriateness, ability to individualize, generalization to other types of tasks), as well as indirect measures of quality through participants' engagement in the intervention.

Our brief review of recent literature published in three peer-reviewed early childhood journals found that implementation data were gathered by most intervention studies (about 75 percent) and that implementation quantity measures were far more frequently used than quality measures. The most prevalent use of implementation data was to describe the intervention rather than to account for intervention effects. In addition, we did not find any studies that examined the interaction between quality and quantity of implementation as predictive of child outcomes. We did, however, find (and we highlight here) studies that offer the field ideas about how to use both types of measures more comprehensively in the quest to understand implementation and intervention effectiveness.

While it is clear that both quality and quantity of implementation measures are being used in early childhood intervention research, quantity measures are far more prevalent, suggesting that greater efforts are needed to incorporate quality measures into the implementation evaluation process. These measures provide data that have the potential to guide feedback to interventionists in ways that quantity measures, like dosage and intensity, cannot. Also important is that few research studies have fully explored the interplay between quality and quantity of implementation, or how the two operate in tandem to ensure optimal program outcomes. Much can be learned by adopting an implementation evaluation strategy that explicitly balances measures that tap into both quality and quantity.

Relevance

Significance of the Problem

In recent years, a remarkable number of early childhood interventions have been developed and tested in highly controlled studies with consistently positive results on caregiver behavior and child outcomes (e.g., Domitrovich et al., 2009; Landry, Swank, Smith, Assel, & Gunnewig, 2006; Neuman & Wright, 2010; Raver et al., 2008). Yet as John Easton, Director of the Institute for Education Sciences, said in a recent keynote about new priorities for intervention research (2010), there is a need to move beyond simply examining "what works and what doesn't" and start understanding "How?," "Why?," "For whom?," and "Under what conditions?" interventions are effective. Toward this end, in 2010, the Administration for Children and Families, along with other federal partners (the Office of the Assistant Secretary for Planning and Evaluation, the Centers for Disease Control's Division of Violence Prevention, the National Institute for Drug Abuse and the National Institute for Mental Health, all at the Department of Health and Human Services, and the Institute for Education Sciences at the Department of Education), convened a meeting to examine methods in implementation research in social and behavioral sciences. The goal of the meeting was to understand the state of knowledge around methods in implementation research and how to build a research agenda to address any identified gaps.

Recent large-scale research efforts, such as the evaluation of the Maternal, Infant, and Early Childhood Home Visiting program, have been guided by implementation science and the quest for understanding how intervention components relate to one another and influence outcomes (Michalopoulos, 2011). Smaller-scale research endeavors have also attempted to move beyond the question of whether interventions work to understanding how they work. For example, the *Classroom Links to Early Literacy* intervention found that a workshop followed by literacy coaching for early childhood educators resulted in positive impacts on teacher practices and children's emergent literacy skills (Powell, Diamond, Burchinal, & Koehler, 2010). However, important questions remain about how many coaching sessions are necessary to result in these impacts, as well as how the quality of the coaching interactions may contribute to variation in uptake of content by teachers. These types of questions are particularly important as the field of early care and education (ECE) moves from relatively small, university-initiated studies of programs to community-based implementation of promising intervention approaches on a large scale. What are the key components of an intervention, how do they lead to more effective practice, and what supports are needed to ensure the implementation of these components? A critical next step toward answering these questions is improving measurement of implementation during intervention research trials as well as during scale-up of model programs.

Though there are a handful of frameworks for defining and measuring implementation (Berkel, Mauricio, Schoenfelder, & Sandler, 2011; Dane & Schneider, 1998; Durlak & DuPre, 2008; Dusenbury, Brannigan, Falco, & Hansen, 2003; Metz, Naoom, & Halle, 2013, this series; Wandersman et al., 2008), this brief will focus in particular on two cross-cutting themes: the quality and quantity of implementation. By considering these constructs together, this brief complements the intervention dosage brief in this series (Wasik, Mattera, Lloyd, & Boller, 2013), which concentrates solely on the quantity of implementation. It is often easier to measure the number or frequency with which facilitators implement certain intervention components or the number of times participants are exposed to an intervention (quantity), rather than assess what actually happens during the implementation of the intervention and whether components were implemented well (quality). This leaves the field with a lack of knowledge about *how* early childhood interventions are implemented on the ground, and to what extent quality and quantity of implementation may operate in combination to ensure positive program outcomes.

Goals of the Brief

In this brief, we will describe conceptual frameworks used to define and measure implementation, define quality and quantity of implementation, embed these terms in theory on how variation in implementation is linked with key program outcomes (e.g., Berkel et al., 2011), and use a brief review of recent empirical work to provide examples of how these constructs are assessed and examined in relation to program outcomes. We will then convey the benefits of measuring both the quantity and quality of implementation to researchers, practitioners, and policymakers. Specifically, we will make recommendations about how to best: 1) document implementation using multiple measures from diverse perspectives, to understand program effectiveness and conditions and requirements for replication; 2) obtain more-thorough information for use in program improvement feedback loops to maximize benefits of early childhood interventions; and 3) obtain a greater understanding of the resources and time needed to examine implementation fully and to promote program effectiveness.

Conceptual Frameworks and Definitions

Conceptual Frameworks

Interest in measuring fidelity of implementation began to increase in the 1970s (see Dusenbury et al., 2003, for a concise historical review). Since then, several frameworks from a wide variety of disciplines have been proposed for defining and measuring implementation (see Metz, Naoom, & Halle, 2013, this series, for a general review of implementation frameworks and Susman-Stillman, Wanless, & Weiland, 2013, this series, for a review of fidelity frameworks). These conceptual frameworks vary in terms of breadth and the specific dimensions included in the models. The broadest framework has been proposed by Fixsen, Naoom, Blase, Friedman, and Wallace (2005), which examines implementation for programs at scale and considers elements supportive of implementation at all levels within a system (see Paulsell, Austin and Lokteff, 2013, this series, for a discussion about levels of implementation). Other conceptual frameworks for measuring implementation focus on smaller intervention research trials aimed at establishing initial program impacts (or replication of impacts in a different setting or with different populations) and examine more narrowly the degree to which programs are implemented as planned (e.g., Berkel et al., 2011; Dane & Schneider, 1998; Daro, 2010). Sanetti and Kratochwill (2009) reviewed conceptual models of implementation and concluded that despite the use of different labels and organizational schemes for describing dimensions of implementation, there exists much conceptual overlap among the various models. For example, many frameworks include structural and process components, as well as aspects of participant responsiveness. Two cross-cutting aspects of implementation are included in all conceptual frameworks: the quality and quantity of implementation. These broad concepts map onto more specific facets of implementation that are often defined differently by different theorists or researchers.

Definitions

As Fixsen et al. (2005) describe, “implementation is defined as a specified set of activities designed to put into practice an activity or program of known dimensions” (p. 5). Important to note, however, is that the study of implementation can be applied to at least two very different situations: intervention research trials and scale-up of evidence-based practices or programs. In the first case, implementation is measured and studied when practices and programs are tested in experimental and quasi-experimental trials used to establish scientific evidence of their efficacy (i.e., positive impacts on desired outcomes under favorable conditions). Under the second scenario, implementation is considered when evidence-based practices or programs (i.e., proven efficacious as described above) are then taken to scale within community settings. Given that the study of implementation in the early childhood field is only beginning to gain momentum, this brief focuses on implementation of intervention research trials, including quasi-experimental and randomized controlled studies.

Implementation can be broken down into subcomponents that differ somewhat across conceptual frameworks, but often include one or more of the following: adherence, dosage, exposure, differentiation, participant responsiveness, and others (Howard, Agnamba, Wessel, and Rankin, 2013, this series). Most of these subcomponents focus on quantity of implementation, which tends to be simpler and easier to collect. Examples of quantity measures include dosage (amount of an intervention), intensity (how much of an intervention is delivered during a session), frequency (how often intervention is delivered), exposure (duration of services received), session duration (length of a session), adherence (proportion of intervention components delivered), etc. All of these are focused on capturing straightforward, objective counts of interventionist or participant behavior. Alternatively, quality measures examine the level of skill shown by an interventionist (e.g., coach/mentor, supervisor, teacher) in delivering an intervention, and tend to be far more difficult (i.e., time consuming, expensive, or possibly problematic in terms of reliability or validity) to gather. However, quality measures also provide more nuanced information on what promotes effective implementation and may have stronger associations with targeted outcomes. Examples of quality measures include how well an interventionist delivers the intervention (e.g., ability to engage participants, pacing, developmental appropriateness, ability to individualize, generalization to other types of tasks), as well as participants' engagement in the intervention. Please refer to Table 1 for some specific examples of how select subcomponents of implementation frameworks are organized and measured along the themes of quality and quantity.

Table 1

Quantity and Quality Examples of Implementation Measures Used in Early Childhood Intervention Research Trials

	MEASURE EXAMPLES
QUANTITY	
Adherence	Teacher report of how often they implemented intervention components as intended (e.g., never to always) Independent observation of a checklist of intervention components delivered as intended
Dosage	Home visitor records of the number of home visits completed Teacher report of hours spent in professional development focused on classroom behavior management strategies
Exposure	Group leader records of the days a child attended a social skills group
Differentiation	Teacher report in both intervention and control groups of hours spent in professional development outside of the intervention trial
QUALITY	
Quality of Delivery	Independent coding of how well a coach provided feedback to a teacher from the transcript of a conference Coaches' observation of how well teachers implement a dialogic book reading practices
Participant Responsiveness	Independent observation of children's engagement during a whole-group literacy lesson Parents' report of their interest in home visits

Literature Review

Purpose

Now that we have established how quality and quantity measures map onto some of the common subcomponents of implementation, it is important to examine how these measures are used across intervention trials in early childhood. We conducted a limited review of recent work to identify current practices in assessing these features of implementation and how these data are utilized to understand what happens in intervention studies. We examined articles published from 2006 to 2011 in three peer-reviewed journals: *Child Development*, *Early Childhood Research Quarterly*, and *Early Education and Development*. These journals were chosen because the first one has the highest impact factor (a measure reflecting the average number of citations to recent articles published in the journal; a proxy for the relative importance of a journal within its field) of any child development journal that publishes early childhood intervention research, and the next two have the highest impact factors of any early childhood-specific journals. Our review identified 57 articles that described research and evaluation of early childhood intervention trials (including experimental and quasi-experimental designs). Our goals were to understand the types of implementation data reported in different intervention studies and to ascertain how those data were gathered and used. We then selected two articles that highlight innovative new directions in the measurement and utilization of implementation quality and quantity data.

Methodology

All 57 articles were coded independently by the two authors of this brief, followed by discussion of the codes to reach consensus in cases of discrepancies. The articles were coded along the following dimensions and reconciled to 100 percent agreement:

- Focus of intervention (classroom, family, both)
- Age range of target children (0-3, 0-5, 3-6)
- Implementation data collected (quantity, quality, both)
- Source of implementation data (self-report, other/observation, both)
- Use of implementation data (descriptively, as predicted by participant characteristics, as an outcome of intervention activities, as a predictor of outcomes).

When interpreting these results, it is important to keep in mind that we made several assumptions in conducting the coding. First, we narrowly defined "intervention" and "outcomes." In their implied theories of change, all reviewed interventions expressed the ultimate goal of improving outcomes for young children. However, many of the interventions were directed at teacher or parent behaviors as mediating variables of effects on children. In line with recent theory on educational and behavioral interventions that endorse this mediational theory of change (see Hulleman, Rimm-Kaufman, & Abry, 2013 and Nelson, Cordray, Hulleman, Darrow & Sommer, 2012), we therefore considered any assessments of adult behaviors or activities that were directly targeted and supported by the intervention to be both proximal intervention outcomes (mediators of impact on the ultimate outcome – children's development) as well as reflections of implementation. This is well-illustrated by the Literacy Environment Enrichment Program (LEEP), a four-credit, in-service course for Head Start teachers that aims to improve supports for children's language and literacy development in the classroom. A study by Dickinson and Caswell (2007) examines the impacts of LEEP on teachers' use of high quality language and literacy supports, as measured by the Language, Literacy, and Curriculum subscale of the *Early Language and Literacy Classroom Observation* (ELLCO). Given that this part of the ELLCO assesses the same language and literacy supports that are emphasized in the LEEP course, in this study we considered this ELLCO subscale to be both an outcome of the intervention as well as a measure of how well the teachers

implemented instructional strategies (implementation quality) that they learned in the LEEP course. In contrast, another ELLCO subscale, General Classroom Environment, emphasizes elements of teacher practice that were not a part of the LEEP syllabus, and therefore would only be considered an outcome measure and not a measure of implementation quality. Therefore, measures of teacher (or parent) practice may be discussed in an article only as outcomes of professional development, but for the purposes of this review were also considered to be measures of implementation quality when they reflected how well teachers were applying strategies in the classroom that were learned as part of the professional development intervention.

Second, we noticed that there was often a fine line of difference between measures of quality and quantity that required careful attention. Specifically, participant and interventionist reports of implementation might be sorted into the quality or quantity bin solely because of the reporting scale. For example, a study might use a teacher's report of how often she implemented a set of best practices during book readings. Given that the emphasis is on the amount or frequency of use, we would code this as a quantity measure. However, if the exact same items were used but a teacher was asked to rate how well she implemented each of the practices, then we would code this as a quality instrument. These types of subtle distinctions might or might not be meaningful when using these data to evaluate an intervention, but are certainly worthy of note when constructing new scales or making decisions about which pre-existing scales to choose and for what purposes.

Summary of Results

A summary of the coding results is provided in Table 2. First, we describe the focus of these intervention studies and the ages of children who were involved. There are many more intervention studies that address the preschool and early elementary school years (*n*=45) than infant-toddler ages (*n*=5) or the full zero-age 6 continuum (*n*=7). Also, the interventions were more likely to have a sole focus on classrooms (*n*=38) compared to families (*n*=6), though there was also a sizeable number (*n*=13) with a dual home-school emphasis.

Next, we note general patterns of using implementation measures. Two-thirds of the studies included quantity measures, while less than half reported assessments of implementation quality (though this percentage of quality measures would be considerably lower if we had not been inclusive of teacher and parent practices that also served as proximal intervention outcomes). Though a third of the studies addressed *both* quality and quantity, almost a quarter of the intervention studies from the past five years did not mention implementation measures at all (they might or might not have collected them). When implementation data were reported within a study, researchers were more likely to use third-party reports or observations (alone or in tandem with self-reports) than to use self-report ratings alone. Such use of objective approaches to measurement and triangulation using multiple methods aligns well with cited best practices in measuring intervention implementation (O'Donnell, 2008). We also set forth to learn how data on the quality and quantity of implementation are being utilized analytically. It is clear that these measures are most often used to provide a description of what happened during the intervention implementation (*n*=30), often in the "Methods" section. More often than not, these descriptions rely heavily on quantity measures. For example, Farver, Lonigan, and Eppe (2009) report children's attendance rates in their intervention—small-group activities from the Literacy Express Preschool Curriculum—in the "Procedures" section.

The second-most-common use of these data is as outcomes to examine intervention effects (*n*=16), followed by implementation measures serving as predictors of child outcomes (*n*=10) or being predicted by participant or setting characteristics (*n*=10). Examples include Koh and Neuman (2009) testing the effects of coursework and coaching on how well family child care providers implemented literacy instructional practices; Driscoll and Pianta (2010) examining whether or not a child's exposure to Banking Time (i.e., one-on-one, child-directed

sessions between a preschool teacher and child) contributed to increased closeness with the teacher; and Baker, Kupersmidt, Voegler-Lee, Arnold, and Willoughby (2010) evaluating the extent to which preschool teacher demographics, job satisfaction, and work environment predicted variation in teachers' participation during a preventative intervention designed to promote social, emotional, academic, and behavioral school readiness.

It remains exceedingly rare for intervention studies to examine quality of implementation as a mediator of the association between intervention condition and child outcomes, or to examine quality and quantity of implementation in combination as predictors of child outcomes. In a random assignment study examining the effectiveness of a social-emotional curriculum, an analysis might go beyond a comparison of the social skills of children in the treatment and control groups to include implementation quality as a mediator of intervention effects. Similarly, an analysis in that same study might include an interaction between children's attendance during the intervention, a quantity measure, and how well teachers taught the whole-group lessons, a quality measure, in predicting social skill outcomes.

Table 2

Summary of Intervention Research Trial Articles Reviewed (N = 57)

	Implementation Data Collected			
	None (*n*=14)	Quality (*n*=5)	Quantity (*n*=18)	Both Quality and Quantity (*n*=20)
Focus of the Intervention				
Classroom	10	4	11	13
Family	2	0	1	3
Both Classroom and Family	2	1	6	4
Age of Target Children				
0-3	2	0	2	1
0-6	3	3	0	1
3-6	9	2	16	18
Implementation Data Source				
Self-Report	na	2	8	2
Other Report/Observation	na	3	10	8
Mixed Methods (Self and Other)	na	0	0	10
How Implementation Data were Used *(Categories not mutually exclusive)*				
Descriptively	na	1	14	15
As Predicted by Other Variables	na	1	3	6
As Predictor of Child Outcomes	na	1	5	4
As an Outcome	na	4	3	9

Model Uses of Quality and Quantity Implementation Measures

The majority of intervention studies in early childhood are not measuring quality and quantity of implementation in tandem, nor using them to understand how interventions lead to change in the ultimate outcomes for children. We therefore chose two studies to illustrate the potential of applying these measures during initial research trials. The first study highlights taking full advantage of measuring both quality and quantity of implementation, while the second provides an example of how these measures can be used to explain variation in child outcomes.

In a study of Getting Ready, an intervention for children from birth to age five that promotes school readiness through parent engagement, Knoche, Sheridan, Edwards, and Osborn (2010) took an approach to documenting implementation that stands as a best practices model for thoroughly examining both the quality and quantity of implementation within the context of an intervention study. Recognizing that implementation is a complex, multidimensional construct, their innovative approach to measurement included: (a) using a comprehensive framework for guiding the measurement of intervention implementation that included an emphasis on quality and quantity; (b) measuring implementation equally across both intervention and comparison groups, when possible; (c) using objective coders who were naïve to group assignment, rather than relying on self-report; and (d) measuring multiple levels of implementation, which in their case meant examining both home visitors and parents as intervention agents. Measures of quantity of implementation included how often specific Getting Ready intervention strategies were used and how often during home visits interactions occurred between home visitors and parents, and parents and children. Measures of quality of implementation included ratings of how well home visitors initiated parental interest and how engaged parents and children were with each other. In addition to examining differences between treatment and comparison groups on these measures, analyses explored the relationships between (a) rate of strategy use and rating of effectiveness, and (b) rate of interactions and rating of engagement. This examination of implementation using both types of measures (quality, quantity) recognized the complexity of the implementation process and the inter-relatedness of different aspects of implementation.

Whereas the Knoche et al. (2010) study provides a clear illustration of how to be thorough in measuring the complexities of quality and quantity of implementation, another study by Domitrovich et al. (2010) offers several novel applications of such measures to understand better the course of implementation over time and the role it may play in producing targeted outcomes (i.e., children's school readiness). Domitrovich, Gest, Jones, Gill, and DeRousie (2010) examined Head Start REDI (Research-based, Developmentally Informed), a comprehensive preschool curriculum focused on children's social-emotional, language, and literacy skills that was implemented by teachers receiving regular coaching support. Both quality and quantity of implementation measures were collected via REDI Trainers, who made weekly visits to teachers' classrooms throughout the year. Trainers rated the quality of implementation by observing children's level of interest as a measure of teachers' ability to engage children during REDI lessons and activities, whereas quantity of implementation was assessed by trainers' tracking the degree to which teachers delivered lessons and activities as intended, and teachers' noting the number of intervention units delivered each week. These data served dual purposes. First, the longitudinal nature of the data collection allowed for plotting of both quality and quantity of implementation across the year, leading to the identification of patterns such as a rise in children's engagement (a quality measure) over time during certain activities. This type of information could be used to guide trainers' feedback to teachers and inform timing/emphasis of REDI implementation supports to teachers during the year. Second, the association of these two implementation components was examined in relation to child outcomes. For example, it was noted that variability in teacher fidelity (a quantity measure) and children's engagement (a quality measure) during social-emotional lessons were consistently linked with social-emotional outcomes, whereas dosage (a quantity measure) was not. Such an approach to analysis can illuminate which aspects of implementation might be contributing to positive intervention effects, thus offering insight into intervention components that drive the effect or suggesting points of emphasis for bolstering implementation with consequences for targeted outcomes.

Implications and Recommendations

This brief review of recent literature published in three peer-reviewed early childhood journals found that implementation data were gathered by most intervention studies (about 75 percent). While roughly a third of the studies reviewed used measures of both quantity and quality of implementation, quantity measures were more frequently used. The most prevalent use of implementation data was to describe the intervention rather than to account for intervention effects. In addition, we did not find any examples of studies in the three journals reviewed that examined the interaction between quality and quantity of implementation as predictive of child outcomes. For example, while the REDI evaluation described above included both quality and quantity measures, associations between these measures and child outcomes were examined separately; interactions between implementation measures were not examined. We did, however, highlight two innovative studies that offer the field ideas about how to utilize both types of measures more comprehensively in the quest to understand implementation and intervention effectiveness. We offer the following implications and recommendations based on our review:

For researchers and funders. All early childhood intervention research must include the collection of implementation data to describe the key components of the intervention and how those components lead to better outcomes for children. Researchers measure quantity of implementation more often for good reasons: compared to quality measures, quantity measures are typically more objective (e.g., counts of instances or records of time intervals) and less time-intensive to collect, and therefore less expensive. However, quantity measures often do not answer important questions about how and why an intervention is effective with particular participants. Therefore, these measures are less useful in identifying the critical components of an intervention. Quality measures, on the other hand, are often less useful in answering important policy and practice questions about thresholds of intervention that produce effects (e.g., how many components or for how long). By including *both* quantity and quality measures of implementation from multiple perspectives, researchers will be better able to describe what happens, at what levels, and how well.

Even when quality and quantity data are being collected, their full potential is not always realized. They are often used only to describe what happened during the intervention. But variability in these measures holds great potential for identifying active ingredients or thresholds of implementation that contribute to positive intervention effects on target outcomes. In fact, it could be that the interactions among these quantity and quality variables offer the most explanatory power in terms of intervention effectiveness. For example, there is evidence in the early education literature that the number of days exposed (quantity) to well-implemented child care (quality) is associated with cognitive and/or social-emotional child outcomes (see Zaslow et al., 2010, for a review). Analyses that examine the nuances in quality-quantity associations, including how the two variables interact to predict outcomes differently for different populations, should be pursued.

For purveyors. Data describing aspects of quantity and quality of implementation provide a rich resource for purveyors that can be used to make continuous improvements to programs through alignment of appropriate implementation supports based on ongoing data collection and analysis. This continuous quality feedback loop was rarely, if ever, mentioned in the reviewed articles, but constitutes one of the central reasons for purveyors to collect quality and quantity of implementation data. Collection of these data on an ongoing basis provides real-time evidence that program benchmarks are being met, such as basic levels of exposure (e.g., attendance at workshop sessions). Perhaps even more importantly, capturing quality data provides an opportunity to give feedback about how well program components are being implemented and to activate supports needed to align these quality indicators with program expectations. Given that data are often already collected during program implementation, this brief simply underscores the need to review existing measures to ensure that both quantity and quality are well represented and then utilized to inform program improvement and replication efforts.

Conclusion

It is clear that both quality and quantity of implementation measures are being used in early childhood intervention research. However, quantity measures are far more prevalent, suggesting that greater efforts are needed to incorporate quality measures into the implementation evaluation process. This is particularly important, because quality measures provide data that have the potential to guide feedback to interventionists in ways that quantity measures, like dosage and exposure, cannot. Also important is that few research studies have explored fully the interplay between implementation quality and quantity, and how the two operate in tandem to ensure optimal program outcomes. There remains much to be learned by adopting an implementation evaluation strategy that explicitly balances measures that tap into both quality and quantity.

Glossary of Terms

Adherence is the degree to which an intervention is delivered as prescribed.

Dosage refers to the duration and frequency of administration of the intervention.

A program is considered to be *efficacious* when it has demonstrated positive impacts on desired outcomes during a scientifically rigorous evaluation.

Exposure refers to the duration and/or frequency of receipt of the intervention.

Participant responsiveness assesses the degree to which participants are supportive of, engaged in, and find value in the intervention.

Program differentiation refers to the extent to which a given intervention differs from other generalized practices.

A ***purveyor*** is an individual or group of individuals, representing a program or practice, who actively work(s) to implement that practice or program with fidelity and good effect.

Randomized controlled trials involve a rigorous, experimental test of a program's impact on desired outcomes, involving participants who were randomly assigned to either the intervention or control groups.

Scale up is defined as the deliberate expansion of an externally-developed program that has been previously proven efficacious in one or a small number of settings, to many settings.

References

Baker, C. N., Kupersmidt, J. B., Voegler-Lee, M. E., Arnold, D. H., & Willoughby, M. T. (2010). Predicting teacher participation in a classroom-based, integrated preventive intervention for preschoolers. *Early Childhood Research Quarterly, 25,* 270-283.

Berkel, C., Mauricio, A.M., Schoenfelder, E., & Sandler, I.N. (2011). Putting the pieces together: An integrated model of program implementation. *Prevention Science*, *12*, 23-33.

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review, 18*, 23–45.

Daro, D. (2010). *Replicating evidence-based home visiting models: A framework for assessing fidelity*. Princeton, NJ: Mathematica Policy Research.

Dickinson, D. K., & Caswell, L. (2007). Building support for language and early literacy in preschool classrooms through in-service professional development: Effects of the Literacy Environment Enrichment Program (LEEP). *Early Childhood Research Quarterly*, *22*, 243-260.

Domitrovich, C. E., Gest, S. D., Gill, S., Bierman, K. L., Welsh, J., & Jones, D. (2009). Fostering high quality teaching with an enriched curriculum and professional development support: The Head Start REDI program. *American Educational Research Journal*, *46*, 567-597.

Domitrovich, C. E., Gest, S. D., Jones, D., Gill, S., DeRousie, R. S. (2010). Implementation quality: Lessons learned in the context of the Head Start REDI trial. *Early Childhood Research Quarterly, 25,* 284-298.

Driscoll, K. C., & Pianta, R. C. (2010). Banking time in Head Start: Early efficacy of an intervention designed to promote supportive teacher-child relationships. *Early Education and Development, 21,* 38-64.

Durlak, J., & DuPre, E. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, *41*, 327-350.

Dusenbury, L.A., Brannigan, R., Falco, M., & Hansen, W.B. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research*, *18*, 237-256.

Easton, J. Q. (2010). *New research initiatives for the Institute of Education Sciences*. IES Research Conference Keynote Address on June 29, 2010.

Farver, J. M., Lonigan, C. J., & Eppe, S. (2009). Effective early literacy skill development for young Spanish-speaking English language learners: An experimental study of two methods. *Child Development, 80,* 703-719.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature* (FMHI Publication No. 231). Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, National Implementation Research Network.

Howard, E. C., Agnamba, L. A., Wessel, J., & Rankin, V. (2013). *Uses and definitions of implementation terms in early care and education research* (OPRE Research Brief). Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Hulleman, C. S., Rimm-Kaufman, S. E., & Abry, T. (2013). Innovative methodologies to explore implementation: Whole-part-whole – Construct validity, measurement, and analytical issues for intervention fidelity assessment in education research. In T. Halle, A. Metz, & I. Martinez-Beck (Eds.), *Applying implementation science to early care and education programs and systems: Exploring a new frontier* (pp. 65-93). Baltimore, MD: Brookes Publishing.

Knoche, L. L., Sheridan, S. M., Edwards, C. P., & Osborn, A. Q. (2010). Implementation of a relationship-based school readiness intervention: A multidimensional approach to fidelity measurement for early childhood. *Early Childhood Research Quarterly, 25,* 299-313.

Koh, S., & Neuman, S. B. (2009). The impact of professional development in family child care: A practice-based approach. *Early Education and Development, 20,* 537-562.

Landry, S.H., Swank, P.R., Smith, K.E., Assel, M.A., & Gunnewig, S.B. (2006). Enhancing early literacy skills for preschool children: Bringing a professional development model to scale. *Journal of Learning Disabilities*, *39*, 306–324.

Metz, A., Naoom, S., & Halle, T. (2013). *A stage-based approach for the implementation of early childhood programs and systems: An integrated framework for moving research to practice* (OPRE Research Brief). Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Michalopoulos, C., Duggan, A., Knox, V., Filene, J. H., Lundquist, E., Snell, E. K., Corso, P. S., Ingels, J. B., Kim, S., & Mello, M. (2011.) *Design options for the home visiting evaluation: Draft final report* (ACF OPRE Report 2011-16). Washington, DC: U.S. Department of Health and Human Services.

Nelson, M. C., Cordray, D. S., Hulleman, C. S., Darrow, C. L., & Sommer, E. C. (2012). A procedure for assessing intervention fidelity in experiments testing educational and behavioral interventions. *Journal of Behavioral Health Services & Research, 39*(4), 374-396.

Neuman, S. B., & Wright, T. S. (2010). Promoting language and literacy development for early childhood educators: A mixed-methods study of coursework and coaching. *The Elementary School Journal, 111,* 63-86.

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research, 78,* 33-84.

Paulsell, D., Austin, A. M. B., & Lokteff, M. (2013). *Measuring implementation of early childhood interventions at multiple system levels* (OPRE Research Brief OPRE 2013-16). Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Powell, D. R., Diamond, K. E., Burchinal, M. R., & Koehler, M. J. (2010). Effects of an early literacy professional development intervention on Head Start teachers and children. *Journal of Educational Psychology, 102*, 299-312.

Raver, C. C., Jones, A. S., Li-Grining, C. P., Metzger, M., Smallwood, K., & Sardin, L. (2008). Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in head Start settings. *Early Childhood Research Quarterly, 23*, 10-26.

Sanetti, L. M. H., & Kratochwill, T. R. (2009). Toward developing a science of treatment integrity: Introduction to the special series. *School Psychology Review*, *38*, 445-459.

Susman-Stillman, A., Wanless, S. B., & Weiland, C. (2013). *Considerations of applying theoretical frameworks of fidelity in early care and education research* (OPRE Research Brief). Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Wandersman, A., Duffy, J., Flaspohler, P., Noonan, R., Lubell, K., Stillman, L., Blachman, M., Dunville, R., & Saul, J. (2008). Bridging the gap between prevention research and practice: The interactive systems framework for dissemination and implementation. *American Journal of Community Psychology, 41*, 171-81.

Wasik, B.A., Mattera, S. K., Lloyd, C. M., & Boller, K. (2013). *Intervention dosage in early childhood care and education: It's complicated* (OPRE Research Brief OPRE 2013-15). Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Zaslow, M., Anderson R., Redd, Z., Wessel, J., Tarullo, L., & Burchinal, M. (2010). *Quality dosage, thresholds, and features in early childhood settings: A review of the literature.* Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.