Abstract Title Page
Not included in page count.

Title: The Development and Application of Fidelity Measures in a Preschool Curriculum Intervention

Author(s): Catherine L. Darrow, FPG Child Development Institute, University of North Carolina at Chapel Hill, David K. Dickinson, Vanderbilt University

Abstract Body

Limit 5 pages single spaced.

Background / Context:

Description of prior research and its intellectual context.

The goal of many preschool interventions is to better prepare children for the academic demands of school. Curriculum-based interventions endeavor to improve the ways in which teachers deliver instruction and developers of curriculum interventions often base curriculum design on assumptions about teaching and learning, build these into the content and structure of the activities, and provide teachers guidance that is designed to help foster learning in ways consistent with the curriculum. Program directors and practitioners often select curricular packages based on researchers' claims that they are effective in increasing children's academic development. Yet, in order for a curriculum intervention to claim success, researchers and developers of the curriculum must be able to identify the essential instructional and structural components and recognize the specific goals of the curriculum. They must also ensure adequate levels of delivery and provide evidence that the treatment condition differs significantly from the control condition through the use of well-developed measures of implementation fidelity.

Researchers often grapple with the multidimensional nature of teaching and learning and the inherent difficulties in implementing effective educational interventions. Models of interventions, as well as the fidelity measures developed to ensure teachers have adequately implemented the causal components, can be equally complex. Despite efforts to create effective measures of fidelity, the instruments produced by many curriculum developers and educational researchers and the methods in which they have been used have frequently fallen short. This paper focuses on the development and application of fidelity measures that represent the critical components of a preschool curriculum intervention. Data collected via these measures are presented through three approaches and provide evidence that fidelity of implementation is a complex and multi-dimensional concept.

A clear, accurate conceptual model acts as a guide to researchers and developers when creating fidelity measures that assess teachers on the degree to which they implement the causal components of the intervention. Use of the measures to assess teachers on the degree to which and quality with which they deliver the curriculum enables researchers to interpret the effectiveness of the intervention. Results from analysis of fidelity of implementation also reveal the ways in which specific elements of the intervention were and were not delivered. This information allows researchers to modify professional development and other in-service training of teachers. In addition, a better understanding of the difficulties teachers face when implementing the curriculum drives researchers and developers to rethink the feasibility of the intervention.

Purpose / Objective / Research Question / Focus of Study:

Description of the focus of the research.

This paper presents the development of reliable measures of fidelity that represent critical elements of the curriculum and its use in assessing teachers' delivering of the treatment curriculum and identifying the existence of program differentiation between the treatment and control conditions. In order to address several issues that arise when assessing levels of implementation while confirming differentiation between treatment and control conditions

through the use of well-developed measures, fidelity data collected from a preschool curriculum intervention have been analyzed to demonstrate an example of:

- Creating psychometrically strong, comprehensive fidelity measures that represent the critical components of an intervention
- Using those measures to view fidelity as something beyond a unitary construct by differentiating implementation of instructional constructs and activities across conditions

Setting:
Description of the research location.

Data in this study were collected as part of Teacher Enhanced Language and Literacy (TELL)[†] that involved a large Head Start program in a medium sized southern city. In this four-year intervention, teachers implemented the *Opening the World of Learning* (*OWL*) curriculum (Schickedanz & Dickinson, 2005) in conjunction with *The Creative Curriculum* (Dodge, Colker, & Heroman, 2002) in 36 treatment classrooms. Teachers in sixteen additional classrooms used *The Creative Curriculum* exclusively as the primary curriculum (i.e. business as usual).

Intervention / Program / Practice:
Description of the intervention, program or practice, including details of administration and duration.

OWL is a comprehensive curriculum with emphasis on developing language and literacy skills in preschool children. The curriculum spans a full year and is separated into six units focused on themes ranging from *Family* to *Things that Grow*. The preschool day is organized into six activities: Morning Meeting, Centers, Small Group (SG), Book Reading (READ), Group Literacy Instruction (GLI), and Let's Find Out About It/Let's Talk About It.

Of all the activities included in the *OWL* curriculum and implemented by teachers assigned to the treatment condition, this study includes analysis of data obtained from (a) Small Group, (b) Book Reading, and (c) Group Literacy Instruction. These three activities in particular are designed to increase children's exposure to concepts of language and literacy considered to be integral to their development. Because of this, it is important to understand how well and to what extent teachers are implementing lessons within these three activities.

Research Design:
Description of research design (e.g., qualitative case study, quasi-experimental design, secondary analysis, analytic essay, randomized field trial).

Analyses reported in this presentation come from a larger IES-funded Randomized Controlled Trial (RCT) in which clusters of Head Start programs were randomly assigned to treatment or control conditions.

Data Collection and Analysis:
Description of the methods for collecting and analyzing data.

[†] The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R324E060088A to Vanderbilt University.

Three activities were videotaped at two time points during the school year for each of the 52 classrooms. Observers coded two sessions of each videotaped activities for teachers' fidelity of implementation. Scores from these two sessions were then averaged to create a classroom score.

Instruments were created to assess teachers on the frequency with which they implemented essential elements of the curriculum across the day. A series of checklists was developed to represent three daily activities: Book Reading, Small Groups, and GLI. With the guidance of the co-author of *OWL*, instructional items identified as uniquely related to the objectives of *OWL* were identified. The number of *OWL*-specific items varied across activities with Book Reading, Small Group, and GLI checklists containing 9, 8, and 7 items respectively.

Findings / Results:

Description of the main findings with specific details.

Creating a Psychometrically Strong, Comprehensive Fidelity Measure

This fidelity measure (represented in Figure 1) accounts for three levels of the intervention: (1) curriculum-level, (2) activity-level, and (3) instructional- and structural-subconstruct level. The curriculum-level of the model includes the three primary daily activities: Small Group, GLI, and Book Reading. The activity-level of the model includes measures of process- and compliance directly related to each activity. Lastly, the instructional- and structural-subconstruct level contains groups of fidelity indicators contained in the observational checklist for that activity (e.g., S1 refers to the first item in Small Group Checklist). There are some commonalities in the subconstructs within each of these activities, yet the delivery of specific curricular elements is also unique to the different activities and settings.

Internal consistency reliability was assessed by calculating a Cronbach's coefficient alpha for all items within each activity and for all items grouped by instructional and structural subconstruct within each activity. Figure 1 shows that the overall scales for Small Group, GLI, and Book Reading activities were 0.77, 0.69, and 0.57 respectively. The calculated Cronbach's alphas for instructional constructs within activities ranged from 0.19 to 0.78. Any construct with an internal consistency less than 0.40 was dropped from the final analysis. The results of internal consistency tests confirm that the majority of fidelity indicators reliably represent the critical elements contained within the conceptual model at both the activity- and subconstruct-level. Therefore, the fidelity measures used in this study provide a valid picture of implementation levels demonstrated by teachers.

Using Measures to View Fidelity as Something beyond a Unitary Construct

The comprehensive set of fidelity measures used in this study provided a nuanced understanding of implementation, revealing differential levels of implementation between teachers assigned to the treatment and control groups. Fidelity measures were used to evaluate teachers' implementation through three different perspectives: overall fidelity across activities, fidelity by activity, and fidelity by structural and instructional subconstructs within activities.

A curriculum-level view: overall curriculum implementation by condition. As expected, teachers assigned to implement *OWL* had significantly higher scores than control teachers in overall fidelity to *OWL* across curricular activities. Treatment teachers implemented 50.2% (SD = 10.6) of the critical elements of *OWL*, averaged across two times and across three

curricular activities. This level of implementation was less than ideal, yet it was significantly higher than that of the control teachers (M = 27.5%; SD = 8.1).

Zooming in: curriculum implementation at the activity-level. Unique patterns of implementation emerged, however, when examining levels of fidelity to *OWL* within specific curricular activities. Treatment teachers implemented *OWL*-specific elements of Small Group (M = 62.67) at higher rates than in any other activity. However, control teachers were also successful at implementing elements of Small Group instruction that were endorsed by *OWL* (M = 28.91) (see Table 1 for complete data). Despite delivering more than half of the curriculum-specific items in Small Group, treatment teachers did not differ much from the teachers assigned to the control condition, in which no *OWL*-specific instruction was expected. Thus, the apparently stronger levels of fidelity in the *OWL* classrooms were negated by the fact that the instruction delivered by treatment and control teachers in Small Groups was not as different as researchers predicted at the start of the study.

Overall levels of fidelity by treatment teachers during GLI were similar to Small Group. On average, treatment teachers implemented roughly 53% of *OWL*-specific items, yet control teachers delivered 21.4% of the OWL curriculum. Although this difference between treatment and control is statistically significant (t = -5.90, p = .000, DF = 50), these levels of implementation do not differ to the extent that researchers expected. Data also suggest that treatment teachers had the most difficulty in implementing *OWL* during Book Reading (M = 32.1). Differences between the *OWL*-specific instruction delivered in the treatment classrooms as compared to the control classrooms were also at their lowest during Book Reading.

The analysis of fidelity at the activity-level provides a more specific picture of differentiation between treatment and control classrooms in how teachers delivered instruction within each of the three curricular activities. Although rates of *OWL* implementation were significantly higher for treatment teachers in all three activities, data revealed more similarities between conditions than did the curriculum-level view of fidelity discussed earlier.

A nuanced view: analysis of structural and instructional subconstructs. A third and final analysis shows that patterns of implementation also varied with regard to specific instructional and structural elements highlighted in *OWL*. Because the same set of validated tools was used to observe and assess both treatment and control teachers on the degree to which they implemented particular instructional methods and structural components, it was possible to focus on nuanced differences between the conditions. Treatment teachers showed higher rates of fidelity in most of the instructional subconstructs related to Small Group and GLI activities. In particular, they were more supportive of children's language development and analytical thinking in Small Group activities. In addition, treatment teachers delivered code-focused instruction during Group Literacy Instruction with higher fidelity than did control teachers.

The set of fidelity measures produced unexpected results when data collected from the Book Reading activity were analyzed. There were no statistical differences in the ways that treatment and control teachers supported children's language development and analytical thinking while reading storybooks.

Data collected on teacher fidelity also represented structural elements of implementation, referred to as compliance in this study. In a similar way to teachers' implementation patterns of instructional elements, teachers' levels of compliance to the treatment curriculum followed a mixed pattern. Treatment teachers demonstrated greater compliance than control teachers during GLI activities in that they presented activities that were specified by lesson plans for the recommended length of time. On the other hand, treatment teachers were statistically equivalent

to control teachers when adhering to structural elements of Small Group activities. Treatment teachers did not consistently select the most challenging lessons scheduled for that day and engage children for the recommended length of time for Small Group activities. Lower rates of implementation may point to a lack of change in teacher behavior, but it may also indicate weakness in the measure itself.

Conclusions:

Description of conclusions, recommendations, and limitations based on findings.

The analyses presented in this paper provide evidence that psychometrically valid set of measures may indeed be used to account for the multi-dimensional nature of implementation fidelity. Given the results of the internal consistency reliability testing, it is possible to conclude that the creation of accurate and valid models and associated fidelity indicators is complex and challenging. Although this model does represent key components of the curriculum, it has some limitations which will be discussed during the conference presentation.

The comprehensive and multidimensional set of fidelity measures used in this study provided researchers the opportunity to examine implementation through a variety of perspectives. When defining fidelity as a unitary construct by assigning each teacher a single value of fidelity across activities (i.e., at the curriculum-level), results suggest that treatment teachers implemented the intervention curriculum at significantly higher rates than did control teachers. When examining levels of fidelity within particular curricular activities, data revealed that the treatment teachers were the better implementers in all three of the activities but levels of differentiation between conditions varied by activity. Furthermore, this perspective provided evidence that treatment teachers had more difficulty implementing Book Reading than any other activity. The final and most nuanced analysis which reported fidelity at the instructional- and structural-subconstruct level revealed a different picture of fidelity by exposing instances in which treatment teachers did not implement instructional and structural elements unique to the intervention curriculum at higher rates than the control teachers.

Views of implementation fidelity as a unitary construct may obscure details about the specific ways in which treatment and control teachers vary and may limit how researchers understand curriculum interventions. The addition of fine-grained analyses made possible by multi-dimensional fidelity measures that examine the degree to which teachers from different experimental conditions deliver specific structural and instructional elements unique to the treatment condition is necessary to provide a robust and accurate representation of implementation. This analytical approach has implications for the ways in which educational researchers perceive the complexity of implementation fidelity, interpret intervention effectiveness, and provide valuable professional development.

Appendices

Not included in page count.

Appendix A. References

References are to be in APA version 6 format.

Schickedanz, J., & Dickinson, D. K. (2005). *Opening the World of Learning*. Parsippany, NY: Pearson Education, Inc.

Dodge, D. T., Colker, L. J., & Heroman, C. (2002). *The Creative Curriculum for Preschool*. Washington, DC: Teaching Strategies, Inc.

Appendix B. Tables and Figures

Not included in page count.

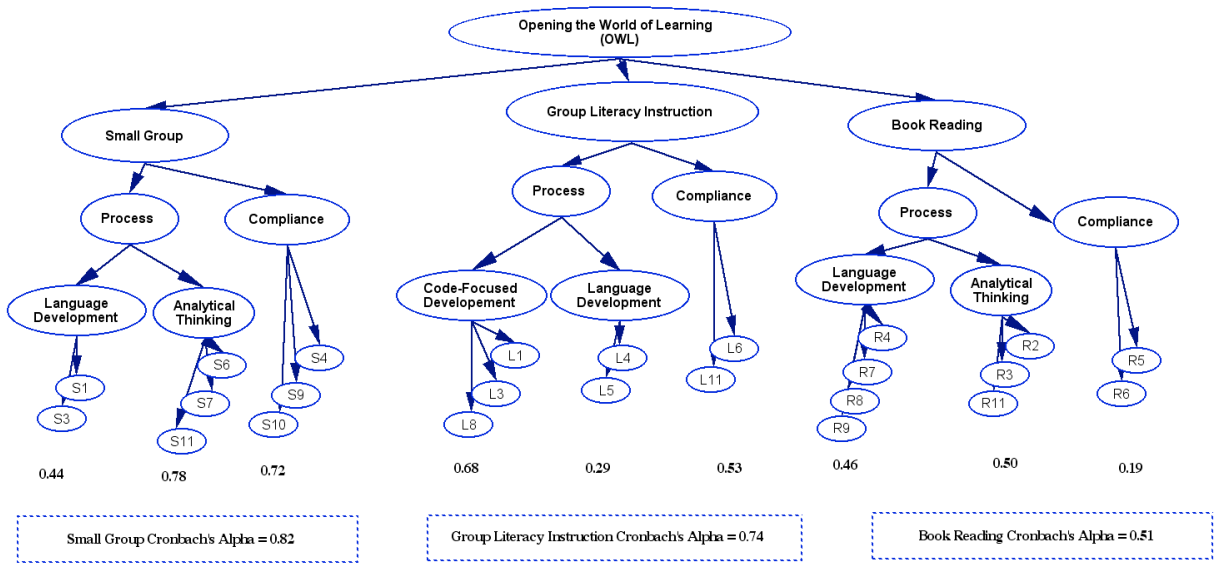

Figure 1. Change model for *OWL* curriculum intervention.

Table 1

Percentage Scores for Fidelity for each Curricular Activity by Condition

	Small Group		GLI		Book Reading	
	Treatment	Control	Treatment	Control	Treatment	Control
	Mean (SD)		Mean (SD)		Mean (SD)	
OWL FOI	62.7 (20.6)	28.9 (14.2)	52.6 (18.5)	21.4 (15.2)	32.1 (11.4)	18.5 (9.0)
CC FOI	52.8 (15.5)	54.5 (20.5)	94.8 (9.6)	80.5 (21.4)	42.3 (8.8)	51.3 (13.6)