

Abstract Title Page
Not included in page count.

Title: Efficacy Study of a Professional Development Intervention to Strengthen Organizational Conditions and Effective Teaching in Early Education Settings

Authors and Affiliations:

Debra M. Pacchiano: The Ounce of Prevention Fund

Samuel P. Whalen, Heather L. Horsley, Kathleen Parkinson: University of Illinois at Chicago

Abstract Body

Background/Context: Decades of evidence demonstrates that high-quality, well-implemented early education can positively impact the learning trajectories of vulnerable, high-needs young children. Yet, the majority of publicly-funded programs nationwide struggle to implement to quality standards with fidelity and fail to significantly advance children’s early achievement (e.g. Belfield, Nores, Barnett, & Schweinhart, 2006). Federal and state accountability structures hold program leaders responsible for particular elements of classroom quality—including learning environments and more recently teacher-child interactions, thereby incentivizing program leaders to focus quality improvement efforts on these classroom-level elements. (Burchinal, Vandergrift, Pianta, & Mashburn, 2010; Rohacek, Adams, Kisker, Danziger, Derrick-Mills, & Johnson, 2010). However, implementation science and school improvement research has greatly advanced our understanding, beyond these classroom-level elements, about the drivers of implementation and the levers of improvement. Key here is recognition that classroom processes do not occur in isolation from organizational processes. Implementation science identifies three core components of infrastructure –competency, data-systems, and facilitative leadership—as key drivers of fidelity (Fixsen, Blase, Duda, Naoom, & Wallace, 2009). And, longitudinal school improvement research in K-12 has empirically related five components of school organization to continuous improvement of teaching and learning: Effective Leaders, Collaborative Teachers, Supportive Environments, Ambitious Instruction, and Involved Families (Bryk, Sebring, Allensworth, Luppescu, & Easton, 2010). Undeniably, effective leadership and professional capacity building are fundamental to achieving *both* implementation fidelity and continuous quality improvement (Bryk, 2012; Metz & Bartley, 2012). The connection between instructional leadership and classroom instruction or children’s achievement is thus mediated by a large number of school processes (e.g., curriculum coordination; professional collaboration) and intermediate outcomes, such as high teacher expectations and climate focused on instruction coupled with supportive interactions. Schools that improve student achievement are more likely to have principals who are strong organizational managers that galvanize activity towards improvement by strategically hiring, professionally supporting, and retaining good teachers, than are schools with principals who spend time observing classrooms without using the information to structure professional development (Grissom, Loeb, & Master, 2013). Research in early education likewise indicates programs more successfully promote children’s learning when they have strong leadership and organizational structures in place to support educator’s professional learning and effectiveness (West-Olatunji, Behar-Horenstein, & Rank, 2008). In response, professional development (PD) purveyors are challenged with transforming designs and aims of PD models; that is, from externally-driven and delivered PD for individual capacity, to models that build administrator capacity to internally-drive continuous quality implementation and improvement by strengthening supports for teachers learning and effectiveness (Elmore, 2004; Hargreaves & Fullan, 2012).

Purpose/Objective/Research Question/Focus of Study: In Fall 2011 the first author was awarded a three-year Investing in Innovation (i3) development grant from the US Department of Education to implement and evaluate a Professional Development Intervention (PDI) for community-based early childhood teachers and administrators based on the five essentials supports framework (Figure 2). The PDI aimed to build (1) teachers’ capacities with ambitious instruction; that is, intentionally designing standards-aligned, data driven instruction, and deliberately implementing evidence-based *general* pedagogical practices such as those measured by the Classroom Assessment Scoring System (Pianta, LaParo, & Hamre, 2008); and (2)

administrators' systemic supports for teachers' effectiveness by strengthening the frequency and coherency of the instructional guidance and professional learning supports they provide teachers. The PDI used the framework of the five essential supports to distill indicated instructional leadership mindsets, functions, and corresponding behaviors that served as the PD content for administrators (Figure 1). The required external evaluation included studies of implementation and impact with the following research questions:

- (Implementation and fidelity) What challenges arise at each successive implementation and iteration of the PD cycles? How do iterations impact realized/intended implementation levels?
- (Adult Learning Outcomes) What features of implementation are most critical to realizing the targeted adult learning outcomes (i.e., specialized knowledge, skills, and dispositions)?
- (Child Impact) Does the PDI produce evidence of superior outcomes for high needs, low-income children in participating settings compared to the outcomes of children in matched non-participating programs? We expected to see the largest differences in social-emotional, language, and cognitive development.
- (Feasibility) How feasible is the implementation of the PDI in community-based settings and contexts? What installation phase characteristics are needed to engage the PDI productively?

Interim implementation study data was used to inform PDI implementation and iteration towards model fit, feasibility, and effectiveness during the grant period (Figure 2).

Setting: Implementation occurred from January 2012 to November 2014 in five publicly-funded, community-based, birth to five early learning programs in a large, urban, city in the Midwest.

Population / Participants / Subjects: Participants included 15 predominantly female administrators of color (i.e. center owners, directors, and direct supervisors); and 60 predominantly Black and Latino infant, toddler, and preschool teachers in the above five programs serving approximately 600 low-income children of color. Of the leaders, 65% hold a post-bachelor's degree, while 44% and 26% of teachers hold an associate's and bachelor's degree respectively, reflecting substantial educational asymmetry.

Research Design: *Implementation Study:* The implementation study was to document the fidelity of implementation of the PDI at coach, teacher, supervisor, leader, center, and PDI program levels, and explore fidelity issues. Three criteria for fidelity measurement served as framework (O'Donnell, 2008): (1) Adherence: Are PDI key components implemented as designed; (2) Duration: number, length, or frequency of PD implemented; (3) Participant responsiveness: extent of participants' engagement with PD. *Impact Study:* A quasi-experimental matched-sample study was conducted to evaluate the efficacy of the PDI. Tests of impacts for children examined change from baseline for the participating and comparison conditions, after establishment of baseline equivalence between treatment and comparison children for each impact measure. A total of 40 centers were matched to the five PDI participating centers based on the number of public funding streams comprising program budget; % free and reduced lunch; demographic composition of children including race, dual-language, and special education; and neighborhood census variables of unemployment and violence. The actual number of comparison centers and children used in specific analyses varied according to availability of data for different impact measures.

Data Collection and Analysis:

Implementation Study: Six key components were measured for fidelity of implementation, with each key component comprised of indicators specifying what is observable and able to be measured as "implemented as planned". Indicators related to fidelity of PDI implementation of

learning contexts by participant group and implementation phase were measured once per year for three years using multiple data sources. Indicators of adult learning were measured using (1) a KWLH graphical organizer was used as formative assessment of adult learning, collected every 4-8 weeks; and (2) individual, team, and focus group interviews with teachers and administrators were completed three times during the PDI. Interview and KWLH data were analyzed using constant comparison method which employs recursive open, axial, and selective coding to identify explanatory theories as data emerges from the study.

Impact Study: Teacher ratings of children's development and learning were collected quarterly (Spring 2012 through Fall 2014) in treatment and comparison programs using Teaching Strategies GOLD (GOLD). The impact study focused on children (N=208) who entered treatment centers in Fall 2012 (two years of treatment) or Fall 2013 (one year of treatment), and who were assessed at Time 2 in Spring 2014. Comparison children (N=924) entered at the same time (Fall 2012 or Fall 2013) and were also present and assessed in Spring 2014. Specific child N's varied by analysis according to whether children had baseline and Time 2 data for specific GOLD sub-scales. Direct kindergarten-readiness assessments of children transitioning in the fall to kindergarten were collected post intervention (Spr 2015) in the 5 treatment and 6 comparison centers using the Bracken School Readiness Assessment (BSRA), a standardized, norm-referenced measure. The 29 treatment and 29 comparison children who completed the BSRA also had GOLD Literacy sub-scale pretest data available which was used to establish baseline equivalency and qualified them for inclusion in the final impact analysis.

A two-level HLM model, with students nested within centers, was used to measure the impact of the PDI on students' average BSRA sub-scale scores and GOLD Social-Emotional, Language, Cognitive, Literacy, and Math sub-scale measures across centers. The model accounts for students' pretest scores, age at the time of pretest, and any statistically significant individual child characteristics impacting overall test scores. At the center level, the impact of similar clusters of centers on overall test scores were controlled for in measuring the Treatment effect, as well as the center-level variables on which treatment and control programs were matched (listed in the Research Design section).

Findings/Results:

Implementation fidelity: More than 75% of centers implemented each key component of the PDI with fidelity across the years (see Table 1 for component fidelity in Years 2 and 3). When looking across all components and all years, every program met the overall PDI fidelity threshold of 75%. Fidelity rates of this magnitude indicate the PDI was implemented as intended, and that leaders, teachers, and coaches successfully engaged the PDI's intensive job-embedded methods. Variation in implementation and attendance was evident, occurring primarily in the initial phases of implementation, which resulting in fidelity not being met in year 2 for two of the six key components: component 2 (PDI Implementation), with indicators measuring the number of expected hours of PD for each group as well as implementation of those hours in a two month time frame; and component 5 (Direct Supervisor PD), with indicators measuring direct supervisor engagement with teacher PD. Meeting expected levels of fidelity on these two key components is extremely challenging, requiring leaders to swiftly develop systems that (1) embed routines for teacher collaboration into daily operations, and (2) protect time for and redeploy supervisors to facilitate teacher professional collaboration and improvement. Installing the PDI's routines raised challenges for leaders including coordinating these routines with other required regulatory meetings and providing classroom coverage to maintain required teacher-

child ratios while teachers meet together for job-embedded professional learning. The evidence suggests that once these factors of fit and feasibility were addressed, fidelity improved to meeting criteria during full implementation in year 3.

Adult learning outcomes: Qualitative analyses indicate that at baseline leaders did not have even declarative knowledge of the organizational conditions essential to continuous quality improvement. Post-PDI, leaders in all four centers shifted understandings, mindsets, and practices towards (a) a “systems-approach” to actively enabling practice and program improvement; that included (b) facilitating teams of educators in routine, structured practice examination and data inquiry; and (c) a disciplined use of PDI protocols by both supervisors and teachers to ensure structured, productive, collaborative reflection, examination, and planning. At baseline, teachers had declarative knowledge but incoherent conceptualization of how to design ambitious instruction. Post-PDI, supervisors had a schedule protecting weekly team lesson planning, and teaching teams were using the PDI lesson planning protocol with fidelity. Post-PDI, lesson plans were substantially goal-focused (vs. activity-focused) and evidenced the use of child progress data to individualize goals and instruction.

Impact study: Baseline equivalence was established for the BSRA (using the GOLD Literacy Sub-Scale) and for all five GOLD sub-scale measures between treatment and comparison groups at pretest. All impact estimates can be found in Table 2 in Appendix B. For the impact of PDI on children’s BSRA scores, the effect size was .13 ($p=.82$). Although no significant differences on the BSRA measure emerged between treatment and comparison children at post-test, data trends favored treatment centers. On the GOLD, effect sizes ranged from .09 to .26, with p-values ranging from .53 to .12. No significant differences on the GOLD sub-scale measures emerged between treatment and comparison children; however, modest effect sizes in favor of the treatment group emerged for social-emotional (.22), language (.16), and cognitive (.26), reflecting higher growth rates for treatment children from baseline. The effect sizes observed are in line with effect sizes of other classroom and school interventions in early education, including the Head Start Impact Study (U.S. DHHS, Administration for Children & Families 2010), the Head Start REDI Prevention Trial (Bierman et al., 2008), and the Chicago School Readiness Project (Raver, Jones, Li-Grining, Zhai, Bub, & Pressler, 2011).

Conclusions: The present work illustrates one approach to developing the instructional leadership capacities of community-based early education leaders and supervisors, leading to internally-driven implementation and support for continuous professional learning and improvement. When considering the complexity of the work involved in reorganizing a center to start up routines for teacher collaboration and data use, the ability to meet fidelity for the majority of key indicators is an impressive outcome. It suggests that the PDI is feasible in community-based settings, and that administrators are able to learn and cultivate specialized skills of instructional leadership and be willing to overcome related challenges of organizational development. The framework of the five essentials facilitated more inclusive and effective navigation by administrators of the required center reorganization and specialized skill development. The framework provided administrators with a coherent conceptualization of the levers and practices of improvement and raised their motivation to strengthen supportive organizational conditions for teaching. The PDI was a development project with a quasi-experimental design that was insufficiently powered due to data availability issues. However, the observed effect sizes suggest that a future adequately-powered randomized control trial may find significant effects for child outcomes of the program.

Appendices

Appendix A. References

- Bierman, K.L., Domitrovich, C.E., Nix, R.L., Gest, S.D., Welsh, J.A., Greenberg, M.T., Blair, C., Nelson, K. & Gill, S. (2008). Promoting academic and social-emotional school readiness: The Head Start REDI program. *Child Development, 79*, 1802-1817.
- Belfield, C. R., Nores, M., Barnett, S., & Schweinhart, L. (2006). The High/Scope Perry Preschool Program: Cost-benefit analysis using data from the age-40 follow-up. *The Journal of Human Resources, 41*(1), 162-190.
- Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development, 78*(2), 647-663.
- Bryk, A.S. (2010). Organizing Schools for Improvement: Lessons from Chicago. *Phi Delta Kappan, V91, N7*, 23-30.
- Bryk, A. S., Sebring, P. B., Allensworth, E., Luppescu, S., & Easton, J. Q. (2010). *Organizing schools for improvement: Lessons from Chicago*. Chicago, IL: The University of Chicago Press.
- Burchinal, M., Vandergrift, N., Pianta, R. C., & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in prekindergarten programs. *Early Childhood Research Quarterly, 25*, 166-176.
- Elmore, R. (2004). *School reform from the inside out: Policy, Practice, and Performance*. Harvard Education Press.
- Fixsen, D. L., Blase, K., Duda, M., Naoom, S., & Wallace, F. (2009) Core implementation components. *Research on Social Work Practice, 19*, 531-540.
- Grissom, J.A., Loeb, S., & Master, B. (2013). Effective Instructional Time Use for School Leaders” Longitudinal Evidence from Observations of Principals. *Educational Researcher, Vol. 42 No. 8*, pp. 433-444.
- Hargeaves, A., & Fullan, M. (2012). *Professional Capital: Transforming Teaching in Every School*. New York: Teachers College Press.
- Metz, A. & Bartley, L. (2012). Active Implementation Frameworks for Program Success: How to Use Implementation Science to Improve Outcomes for Children. Zero to Three, March, 10-18.
- Pianta, R.C., La Paro, K.M., & Hamre, B.K. (2008). *The Classroom Assessment Scoring System Manual, PreK*. Paul H. Brookes Publishing: Baltimore, MD.
- Raver, C. C., Jones, S. M., Li-Grining, C. P., Zhai, F., Bub, K., & Pressler, E. (2011). CSRP’s impact on low-income preschoolers’ pre-academic skills: Self-regulation as a mediating mechanism. *Child Development, 82*(1), 362-378.
- Rohacek, M. Adams, G. C., Kisker, E. E., Danziger, A., Derrick-Mills, T., & Johnson, H. (2010). Understanding quality in context: Child care centers, communities, markets, and public policy. Retrieved from Urban Institute website: <http://urban.org/UploadedPDF/412191-understand-quality.pdf>
- U.S. Department of Health & Human Services, Administration for Children & Families (January 2010). *Head Start Impact Study Final Report*. Washington, DC: Author.
- West-Olatunji, C., Behar-Horenstein, L., & Rank, J. (2008). Mediated lesson study, collaborative learning and cultural competence among early childhood educators. *Journal*

of Research in Childhood Education, 23, 96-108.

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 Brief No. 2011-2010a). Retrieved from the U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research, and Evaluation website: http://www.acf.hhs.gov/programs/opre/cc/childcare_technical/reports

Appendix B. Tables and Figures

Figure 1. PDI Theory of Intervention and Impact

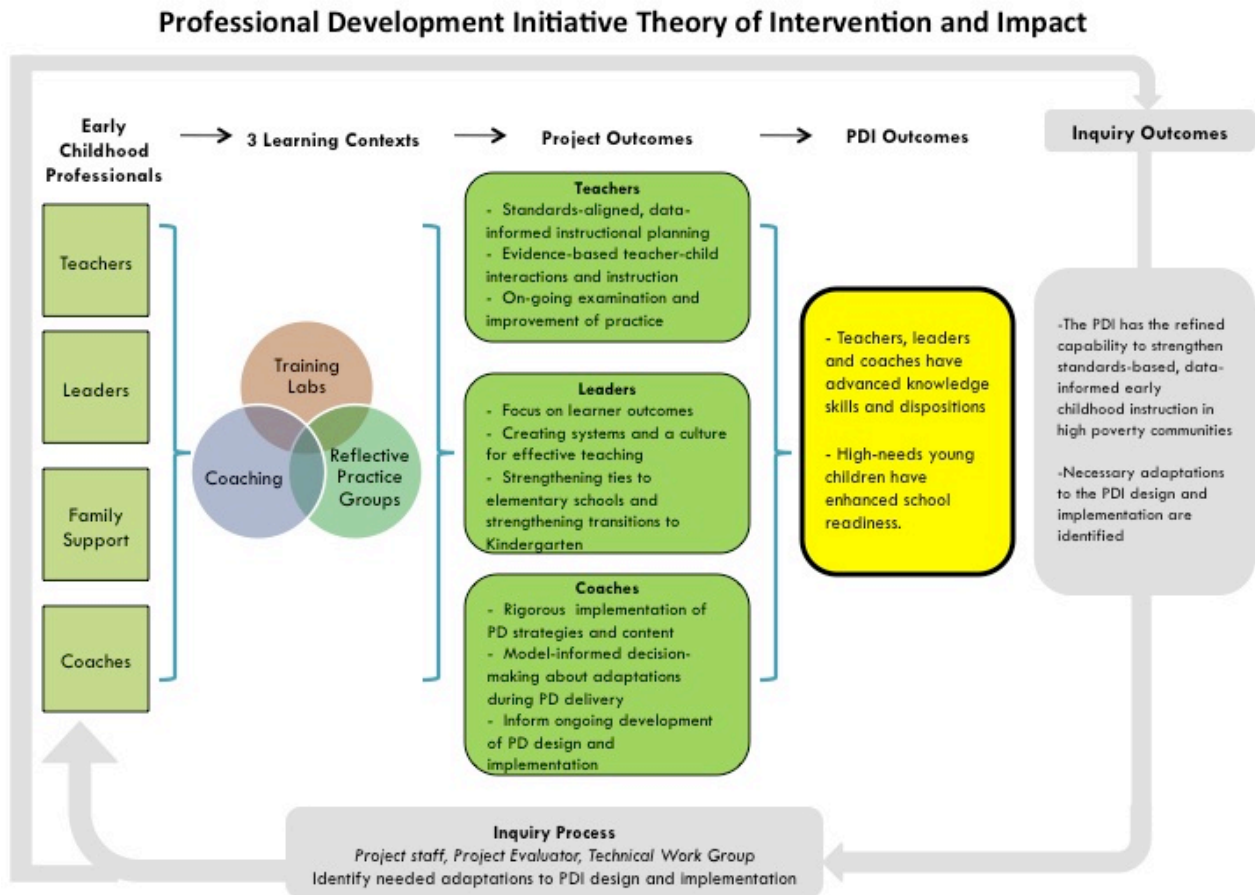
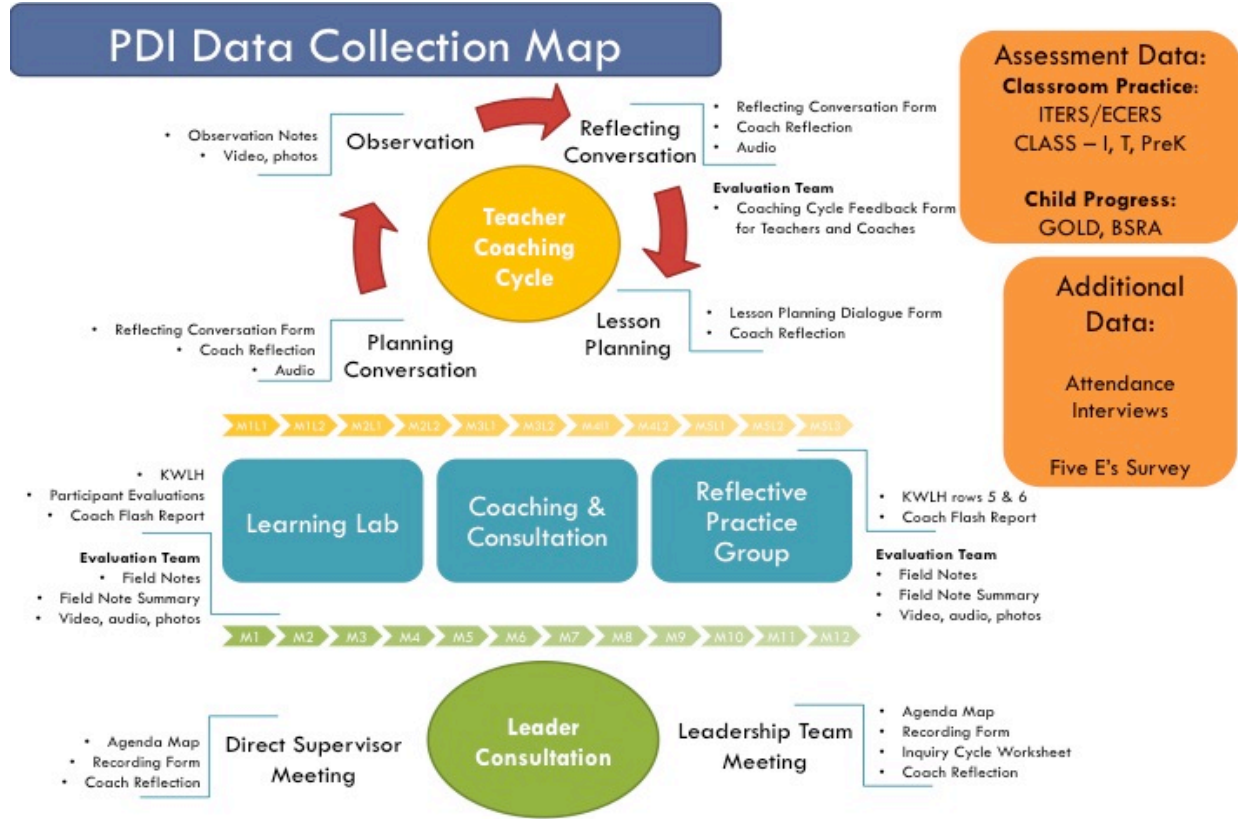


Figure 2. Data Collection Map



V.4 11-22-2014

Table 1: Component Level Fidelity of Implementation Findings for Year 2 and Year 3

Key Component	Year 2		Year 3	
	<i>Component Score</i>	<i>Implemented with Fidelity*</i>	<i>Component Score</i>	<i>Implemented with Fidelity*</i>
1. Coach Community of Practice	100%	Yes	100%	Yes
2. PDI Implementation	50%	No	75%	Yes
3. Coach PD	100%	Yes	100%	Yes
4. Teacher PD	100%	Yes	100%	Yes
5. Direct Supervisor PD	0%	No	100%	Yes
6. Center Administrator PD	75%	Yes	100%	Yes

** More than 75% of centers implemented each component with fidelity in order to meet overall program fidelity*

Table 2. Impact Estimates from HLM Output for GOLD and BSRA at Time 2

Measure Name	Impact Treatment Group N	Impact Comparison Group N	Impact Unadjusted Treatment Group Mean (SD)	Impact Unadjusted Comparison Group Mean (SD)	Standardized Adjusted Impact Estimate	Impact Standard Error	p-value
GOLD Social Emotional	199	907	584.83(85.02)	586.51(85.25)	0.22	16.50	0.27
GOLD Language	198	908	579.39(92.66)	588.91(91.85)	0.16	16.91	0.40
GOLD Cognitive	195	897	587.77(101.16)	597.29(99.48)	0.26	16.06	0.12
GOLD Literacy	192	863	589.45(86.76)	601.94(88.43)	0.10	12.95	0.53
GOLD Mathematics	194	852	592.68(92.79)	610.34(84.79)	0.09	13.10	0.52
Bracken School Readiness Assessment* (BSRA)							
Bracken	29	29	97.55(11.18)	96.93(14.30)	0.13	6.89	.82

*The BSRA included only a sub-sample of treatment (n=29) and comparison (n=29) children who also had GOLD Literacy sub-scale pretest data in order for baseline equivalency to be established which qualified them for inclusion in the final impact analysis.