Evaluation Studies and Methods

Continuous Evaluation and Quality Improvement: A Head Start and University Partnership

Christopher L. Smith, David P. Lindeman, Linda Broyles

PRESENTERS: Christopher L. Smith, Linda Broyles

Beginning in October of 2001, the Southeast Kansas Community Action Program (SEK-CAP, Inc.) and the Kansas University Center on Developmental Disabilities (KUCDD) in Parsons began a collaboration to develop a comprehensive system-level assessment, as well as an in-depth evaluation of the organization's early childhood programs in particular. Each year, SEK-CAP, Inc. provides services to over 950 children in Early Head Start and Head Start. These children are served in a combination of home and center-based programs in 12 counties. The centers number 14 locations, 16 sites (some towns have more than one site), and 29 classes (many sites have both morning and afternoon classes). In addition, the organization provides other services and programs such as family services, transportation, and housing. Data from consumer surveys indicate that as many as 65% of all consumers actually use two or more of the agency's services, while as many as 30% may use three or more. This data suggests that evaluating the effect of a Head Start program that is embedded in a larger organizational service context must include a more integrated approach to evaluation. Continuous systems-level assessment (CSLA) such as the evaluation system being employed by these evaluation partners is conceptualized as an ongoing, comprehensive, integrated, and staff development oriented process. A continuous evaluation process creates feedback loops within the organization, allowing staff and consumers to utilize both existing and new data to improve service delivery on a longitudinal basis (Morrill, 1996). In this sense, the CSLA model is much like Stufflebeam and Shinkfield's (1985) Context-Input-Process-Product (CIPP) model, the Continuous Quality Improvement model of Deming (1986), and Senge's (1990) learning organization. In addition, in order to make changes in their practices, teachers (like those in Head Start centers) need ongoing support and feedback on the effects of their work (Guskey, 1995). Planning and developing staff improvement activities without considering organizational issues, policies, and other systems-related variables can limit the success of new programs (Guskey, 1995; Joyce & Showers, 1995). For these reasons, the CSLA process is conceptualized in a way that facilitates an ongoing assessment of staff development issues as they relate to the organization and community concerns identified in the initial and ongoing needs assessment processes. SEK-CAP, Inc. has undertaken significant organizational change in order to accommodate the needs to continuously collect, analyze, and reflect on its performance data. Consistent with the evaluator's previous work on continuous systems level assessment (Smith & Freeman, 2002) and the priorities of the organization, the partnership evaluation strategies have included three key types of relevant evaluation including descriptive (through community needs assessment), process/formative (continuous monitoring of services), and outcomes evaluation. The poster session provides graphic and other representations of the continuous systems level assessment process, as well as examples of the types of data used by the organization. These data offer key insights into the use of specific evaluation methods that are highly relevant to Head Start programs, and especially those that are embedded as a part of a larger community-based organization.

References

Deming, W. E. (1986). Out of crisis. Massachusetts: MIT Press.

Guskey, T. (1995). Professional development in education: In search of the optimal mix. In T. R. Guskey & M. Huberman (Eds.), Professional development in education (pp. 113-131). New York: Teachers College Press.

Joyce, B., & Showers, B. (1995). Student achievement through staff development: Fundamentals of school renewal (2nd ed.). White Plains, NY: Longman.

Morrill, W. A. (1996). Implications for the future of service delivery reform. In J. M. Marquart & E. L. Konrad (Eds.), Evaluating initiatives to integrate human services (pp. 85-95). San Francisco, CA: Jossey-Bass.

Senge, P. M. (1990). The fifth discipline: The art and science of the learning organization (pp. 139-173). New York: Doubleday.

Smith, C. L., & Freeman, R. L. (2002). Using Continuous System Level Assessment Methods to Build School Capacity. American Journal of Evaluation, 23(3), 307-319.

Stufflebeam, D. L., & Shinkfield, A. J. (1985). Systematic evaluation (pp. 151-206). Boston: Kluwer-Nijhoff.

Evaluating the Impact of a State Early Childhood Education Policy Initiative on Children's School Readiness: Lessons Learned from a Community Based Collaboration

Kimberly G. Browning, Marijata Daniel-Echols, Susan DeVenny, Dan Wuori

PRESENTERS: Kimberly G. Browning, Marijata Daniel-Echols, Susan DeVenny, Dan Wuori

This poster describes results from an evaluation of the South Carolina First Steps to School Readiness initiative. First Steps is a comprehensive education initiative to help poor and at-risk children reach kindergarten and first grade healthy and ready to succeed in school. The initiative combines' public and private support through county partnerships to enable individual communities to address unmet needs of young children and their families with the goal of improving school readiness.   Community leaders collaborate in each county to conduct needs/resource assessments to develop strategic plans that address the unmet early childhood needs specific to each community. The program then provides funds to fill gaps where services are needed by allowing a community to begin to offer new or previously unavailable services (e.g. child care vouchers, teacher training, child care quality improvement, and parenting and adult literacy classes) or to expand existing services such as public preschool (4K), Head Start, or Success by 6, among others.

Secondary analysis was conducted on existing data over four cohorts of children that had received some sort of initiative investment in themselves or their family between 2000 and 2005. Each cohort is a data set of all children who entered kindergarten in public schools during a school year beginning in 2001 through 2004. Child assessment data was collected statewide by the South Carolina Department of Education in kindergarten, first and third grade.

Analysis of cohort data clearly indicates that children who receive First Steps funds experience more risk factors for school failure than those who do not. They are more likely to be poor and are more often minorities compared to those who do not attend 4K at all. Among First Steps-funded with 4K children, those enrolled in full-day 4K are the most disadvantaged. Overall, however, First Steps-funded children who do not attend 4K are the poorest and most at-risk.

There is evidence that all children who attend 4K increase their language, math, and learning approaches over the course of one preschool year. There is also evidence that 4K produces better long-term outcomes for all children who attend. There is a trend among First Steps children who attend 4K of improvement in their odds of grade retention and they generally scored higher on the South Carolina Readiness Assessment (SCRA) compared to non-4K First Steps children. Children who participated in First Steps-funded programs without 4K are most at risk for special needs placement and lower academic achievement compared to both their First Steps-funded with 4K and non–First Steps 4K peers.

There appears to be consistent evidence that enrollment in full-day 4K has positive effects on child outcomes particularly during the kindergarten year. It is clear from these analyses that while, in general, 4K experiences reduce the odds of being retained, full-day 4K has a greater impact in reducing the odds of being retained. A consistent finding within the SCRA and the

state mandated Palmetto Achievement Test scores in 3rd grade is that 4K effects are stronger for minority children than for White children.

Keeping It Positive in the Head Start Classroom: Preliminary Reliability and Validity of the Checklist of Behavior Management Style

Molly Metzger, C. Cybele Raver, Ta-Tanisha Young, Bonnie Solomon

PRESENTERS: Molly Metzger, Bonnie Solomon

To measure prekindergarten classroom quality, researchers are equipped with well-established tools measuring didactic aspects of the early learning environment (e.g., ECERS-R, Harms, Clifford, & Cryer, 1998). Though measures like these may include items regarding social-emotional aspects of the classroom, there remains a need for a measure that focuses explicitly on teachers' ways of handling children's disruptive behavior. Such a measure would provide clearer evidence as to whether teachers successfully implement the classroom management strategies that they have been exposed to in teacher training sessions or in-service training. This study aims to fill that void by offering a new observation-based measure of teachers' behavior management strategies. This measure has been developed to be aligned with some of the curricular components of evidence-based teacher training used in a number of early intervention settings (Webster-Stratton, 2004).

The present sample consists of 35 classrooms within 18 Head Start sites partnering with the Chicago School Readiness Project. From 2004 through 2006, the study obtained parental consent for 604 children in the participating classrooms. These children ranged from ages 3 to 5 and represented roughly equal numbers of girls and boys, with 29% being Latino/a, 64% African American, and 7% Caucasian/other. A total of 98 teachers also participated in the project.

Observational classroom-quality data were collected in each classroom at three points throughout the school-year (January, March, and May), using two measures of teacher behavior. The Classroom Assessment Scoring System (CLASS; LaParo, Pianta, & Stuhlman, 2004) provided global, 7-point Likert scores in 5 aspects of emotional support (positive climate, teacher sensitivity, behavior management, negative climate, and over control), as well as 1 item concerning children's level of engagement ($\alpha = .85$ for emotional support items, LaParo, Pianta, & Stuhlman, 2004). The new, 10-item Checklist of Behavior Management Style (CBMS) provided specific information regarding strategies that teachers employ to keep children "on track" with classroom activities. Both of these measures were collected during three periods (breakfast, "circle time," and lunch) of each observation day, and the resulting data utilize the mean scores across these observation periods.

Classroom observations were double-coded for 66% of the data collected. Inter-rater reliability for each of the 10 CBMS items averaged at $\alpha = .67$. Exploratory factor analyses, conducted separately for each of the three months of data collection, demonstrated that the CBMS cohered around two factors: proactive strategies (such as positive redirection and the use of non-verbal support) and harsh/reactive strategies (such as using a loud or harsh tone of voice and shaming or humiliating children). Preliminary evidence for external validity of the CBMS was demonstrated, with proactive strategies of the CBMS positively correlated with CLASS items such as teacher sensitivity and behavior management. Likewise, the harsh/reactive CBMS

strategies were significantly associated with CLASS items such as negative climate and over-control. These results provide support for the continued piloting and development of the CBMS.

References
Harms, T., Clifford, R. M., & Cryer, D. (1998). Early childhood environmental rating scale-Revised edition. New York: Teachers College Press.

Fantuzzo, J.W., & Hampton, V.R. (2000). Penn interactive peer play scale: A parent and teacher rating system for young children. In K. Gitlin-Weiner, A. Sandgrund, & C. Schaefer (Eds.), Play diagnosis and assessment (2nd Ed.). New York: Wiley.

LaParo, K. M., Pianta, R. C., & Stuhlman, M. (2004). The classroom assessment scoring system: Findings from the prekindergarten year. The Elementary School Journal, 104(5), 409–426.

Webster-Stratton, C. (2004). How to Promote Children's Social and Emotional Competence, Thousand Oaks, CA: Sage Publications Inc.

Diversity in Disadvantage: Tracking the Developmental Trajectories of Head Start Graduates in the Early Childhood Longitudinal Study – Kindergarten Cohort (ECLS-K)

Jerry West, Louisa Banks Tarullo, Nikki Aikens

PRESENTERS: Jerry West, Louisa Banks Tarullo

Successive cohorts of a large, longitudinal descriptive study of Head Start children, the Head Start Family and Child Experiences Survey (FACES), show that most Head Start children have below-average literacy skills both when they enter and exit from the program. However, findings from FACES also suggest that the population of entering Head Start children is actually more diverse than might be expected (ACF, 2003). The highest score quartile of Head Start entrants scored at or above national averages in vocabulary, early math, and early writing. In contrast, mean standard scores for the lowest quartile of Head Start entrants were at least two standard deviations below national averages in these domains. By the spring of kindergarten, the average child was approaching national norms in vocabulary, math, and early writing, but children who started Head Start with scores in the highest quartile tended to show the least gains. These data lead to two questions: What are the learning trajectories of these different groups of children as they move further through the early school grades? What factors may influence variations in their learning trajectories during the early school years?

Data from the Early Childhood Longitudinal Study - Kindergarten Cohort (ECLS-K) are used to help answer these two questions (Tourangeau et al., 2005). The ECLS-K included a sample of children for whom Head Start attendance was verified by contacting the programs that children had been presumed to attend the year before kindergarten based on parental reports, and asking about the child's attendance (Manship & West, 2004; West & Bose, 2002). This methodology yielded a sample of 2,200 children.

The ECLS-K assessed children's early reading and mathematics knowledge and skills using a two-stage adaptive test (Pollack et al., 2005). For the analyses presented in this poster, the sample of Head Start-verified children is divided into three groups, characterized by low, middle and high scores on the ECLS-K reading and mathematics achievement tests at kindergarten entry. The top third of Head Start graduates entered kindergarten with reading and mathematics scores that were several points above the national average, and maintained this trajectory into the spring of third grade. Children in the high group were more likely to be White and were more advantaged (e.g., higher household income and more home literacy resources) than children in the other two groups.

Hierarchical linear models (HLM) estimating children's reading trajectories from the fall of kindergarten to the spring of third grade indicated that Head Start children in the high-achievement group not only had higher reading scores in the fall of kindergarten but that they gained reading skills at faster rates through third grade in comparison to the low- and middle-level groups. Additional models suggest that differences in gains between low- and high-achievement groups are greater for children whose household income is below the federal

poverty level than for other children. Differences in the rates of growth between the three achievement groups were smaller for minority children than for White children.

References
Administration for Children and Families (ACF). Head Start FACES 2000: A Whole-Child Perspective on Program Performance. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, 2003.

Manship, K. and West, J. "Assessing the Quality of Parent Reports of Their Children's Head Start Attendance: Findings from the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99. Paper presented at the Head Start 7th National Research Conference, Washington, DC, 2004.

Pollack, J., Atkins-Burnett, S., Rock, D., and Weiss, M. Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Psychometric Report for the Third Grade. Washington, DC: National Center for Education Statistics, 2005.

Tourangeau, K. Brick, M., Byrne, L., Le, T., Nord, C., West, J., Germino Hausken, E. Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Third Grade Methodology Report. Washington, DC: National Center for Education Statistics, 2005.

West, J. and Bose, J. "Using Administrative Records and Follow Up Interviews to Verify Household Survey Responses." Paper presented at the International Conference on Improving Surveys, Copenhagen, Denmark, 2002.

Implementation and Social Competence Outcomes of a Prevention/Readiness Curriculum: Children's School Success

Samuel L. Odom, Marci J. Hanson, Shana Cohen, Janese Kerr Daniels, Ruth Schneider

PRESENTERS: Samuel L. Odom, Marci J. Hanson, Shana Cohen, Janese Kerr Daniels, Ruth Schneider

Success in school depends, in part, on children's social competence. Young children from low-income communities often exhibit higher incidence of problem behavior and lower social performance than children from other social strata (Webster-Stratton & Hammond, 1997). To promote children social competence, reduce problem behavior and promote readiness for kindergarten, the Children's School Success (CSS) Project developed an integrated curriculum that address social and academic performance. As part of a larger randomized experimental study of curriculum efficacy, researchers trained teachers to implement the curriculum in typical Head Start classrooms, assessed the level and quality of implementation, and examined the relationship between degree of implementation and children's performances. The primary research questions were: a) to what degree and with what quality did the teachers implement the social component of the experimental curriculum? b) Is there a relationship between implementation and child outcomes?

Teachers and children from 15 classrooms located in CA, KS, IN, WV, and MD participated in the study. All classrooms were from Head Start programs, except for the KS site, which were state pre-kindergarten and private preschools serving low SES children. In each classroom, data were collected on 6-13 four-year old children. Although three social outcome variables were assessed (i.e., direct observation of children, children's social problem assessment, teacher ratings), for this presentation only data from the Social Skills and Problem Behavior subtests of the Social Skills Rating System (Gresham & Elliott, 1990) were analyzed. Teachers completed the questionnaires at the beginning and end of the school year.

The social component of the CSS curriculum was based on the Incredible Years Dinosaur School curriculum, developed by Webster-Stratton and colleagues (2003; in press). Although the social and academic components were integrated, the primary emphasis of the curriculum alternated between the two, so that the social skills activities (which occurred in large and small groups) were implemented every other day. Before the year began, the experimental classroom teachers were trained to implement the CSS curriculum in their classrooms. Site supervisors visited the classroom teachers weekly to observe teachers' implementation, model teaching techniques, and collaboratively plan with teachers the next week set of activities.

Analyses of the implementation indicated variability across classrooms in implementation. On average, about 70% (range = 15%-95%) of the social competence curriculum was implemented and the mean quality item rating was 3.63 (range 1.57 - 4.99). The implementation metric average was 2.67 (range = .23 – 4.74).

A multilevel (HLM) analysis examined the relationship between implementation and teacher ratings of social competence. A positive relationship existed between implementation and teacher ratings, with the strongest relationship occurring for children who were the lowest

performing at pretest. For the problem behavior subtest, a negative relationship between implementation and standardized score occurred for children who had the lowest scores at pretest (i.e., children with the fewest negative behaviors), although this relationship did not occur for children with average and relative high levels of problem behavior.

References

Gresham, F. M., & Elliott, S. N. (1990). *Social skills rating system.* Circle Pines, MN: American Guidance Service.

Webster-Stratton, C., & Hammond, M. (1998). Conduct problems and level of social competence in Head Start children: Prevalence, pervasiveness and associated risk factors. *Clinical Child Psychology and Family Psychology Review, 1*(2), 101-124.

Webster-Stratton, C., & Reid, M. J. (2003). Treating conduct problems and strengthening social emotional competence in young children (ages 4-8 years): The Dina Dinosaur treatment program. *Journal of Emotional and Behavioral Disorders, 11*(3), 130-143.

Webster-Stratton, C., & Reid, M. J. (in press). Strengthening social and emotional Competence in socio-economically disadvantaged young children: Preschool and kindergarten school-based curricula. In W. Brown, S. Odom, & S. McConnell (Eds.), *Social competence of young children: Risk, disability, and evidence-based practices* (2nd Edition), Baltimore: Paul Brookes.