

Part B

**Supporting Justification Request for
OMB Clearance of Information Collection Forms for:**

**AN IMPACT EVALUATION
OF EARLY LITERACY PROGRAMS**

July 16, 2007

Submitted to:

U.S. Department of Education
Institute of Education Sciences
555 New Jersey Ave., NW, Rm. 506
Washington, DC 20208
202-219-1597

Project Officer:
Sandra Garcia

Submitted by:

REL Appalachia / CNAC
4825 Mark Center Drive
Alexandria, VA 22311
703-824-2828

Principal Investigator:
Steven Ross

TABLE OF CONTENTS

B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS

1. Respondent Universe and Sampling Procedures.....	3
2. Statistical Methods for Sample Selection and Degree of Accuracy Needed.....	5
3. Methods to Maximize Response Rates and Deal with Nonresponse.....	9
4. Test of Procedures and Methods to be Undertaken.....	10
5. Individuals Consulted on Statistical Aspects.....	10
REFERENCES.....	11
APPENDIX: Early Literacy Study Consent Form.....	12
Teacher and Paraprofessional Consent Form.....	13
ATTACHMENTS.....	14
Exhibit A: Memphis City Schools OWL Program Teacher Questionnaire.....	15
Exhibit B: Memphis City Schools Pre-K Program Teacher Questionnaire.....	16
Exhibit C: Memphis City Schools OWL Program Paraprofessional Questionnaire.....	17
Exhibit D: Memphis City Schools Pre-K Program Paraprofessional Questionnaire.....	18

B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

1. Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection method to be used. Data on the number of entities (e.g. establishments, State and local governmental units, households, or persons) in the universe and the corresponding sample are to be provided in tabular form. The tabulation must also include expected response rates for the collection as a whole. If the collection has been conducted before, provide the actual response rate achieved.

Memphis City Schools (MCS) is the largest school system in Tennessee and is comprised of 191 schools in grades K-12. The district serves more than 115,000 students in 112 elementary schools, 25 middle schools, and 31 high schools. The student population is comprised of 85% African American students, 9% White, 4% Hispanic, and 2%, Other. Over 71% of students qualify for free or reduced-price lunch, and 14.4% are enrolled in special education programs. The setting for the study will be MCS preschool classrooms funded by the district or Title I. Children served will be four-to-five year-olds who are predominantly from low-income families. The 50 study preschool classrooms are located in 41 Title I schools. The sample of children will include high proportions of both educationally at-risk and African American children. The race/ethnicity categories of children (treatment and control) are supplied by the school district. For Cohort I, 97% are African American, 1.5% are Hispanic, .9% are White and .6% are Other.

Random assignment to groups was made using the Social Psychology Network's web-based randomization tool, Research Randomizer (www.randomizer.org). The 41 schools are the randomization unit, and the 50 classrooms are units within the schools. The program generates random numbers through a complex algorithm seeded through timekeeping software. The schools agreed to random assignment on request of the district, independent of (but in consonance with) the needs of the randomized study. The program schools received *Opening the World of Learning (OWL)* curriculum training and materials in August. A district representative observed the randomization process, and results were shared with principals and teachers and approved by the executive leadership in MCS. Memoranda of Understanding and MCS research approval have been received. The Education Innovations (EI) research team made the random assignments using a stratification based on whether the preschools were (a) district funded (n = 19) or (b) funded through Title I (n = 22). The randomization procedure assigned 26 classrooms in 21 schools to the treatment group and 24 classrooms in 20 schools to the control group.

Within each classroom, all student enrollees (having parent permission) will be assessed, for a total of approximately 1000 children (with the expectation of 15-20% attrition by the end of the school year). In Year 2, a second cohort in the same schools and classrooms will be added to the study. Because the primary focus of the study is on the impact of the Early Literacy program intervention on children's school readiness, all

analyses will be at the individual student level, with school variables controlled via the hierarchical linear model (HLM) analyses to be employed. For this new collection, the expected response rate is 85% (post-test) for students, 85% for teachers, and 85% for paraprofessionals. The expected response rate for the collection as a whole is 85%. These figures are based on the 2007 collection by the school district: 92% for students, 88% for teachers, and 82% for paraprofessionals. To minimize attrition among teachers and paraprofessionals, project staff will make reminder phone calls and will make visits to schools to encourage survey participation. See Attachments A, B, C, and D for copies of the survey instruments and the Appendix for the district-administered consent form.

Participants will be children enrolled in 50 preschool classrooms in Memphis City Schools. They will have become four years of age by September 30th of the year they enrolled in the preschool program. The children typically live in that school’s community and are anticipated to enroll in kindergarten at that school in August of the next year. Teachers are state certified in Early Childhood Education and are full-time employees of Memphis City Schools (MCS). Paraprofessionals meet state requirements to be teaching assistants in preschool classrooms and are employees of MCS. Parents are the parents or legal guardians of children participating in the information collection. The majority of families are living at the poverty level, since poverty is one criterion for acceptance into the preschool program. Table 1 represents the tabulation of the potential respondent universe. OMB approval is sought by EI for the teacher and paraprofessional surveys highlighted in bold font. The school district, as stipulated by state program requirements, administers all parent surveys. The school district also committed to conduct all Cohort I (Year II of study) assessments of teachers and paraprofessionals. Table 1 displays all assessment to take place in the study. OMB approval is required only for Year III (Spring 2008) to permit EI to conduct teacher and paraprofessional surveys for Cohort II.

Table 1. Tabulation of the potential respondent universe

Year	Period	E-LOT	ELLCO	Teacher	Para-professionals	Parent Assessments	Child assessments
1	Spring '06						
1	Fall '06	26	26				1,000
2	Spring '07	50	50	50	50	400-425 (50% non-respondents)	800-850 (15-20% attrition)
2	Fall '07	50	50				1,800-1,850
3	Spring '08	50	50	50*	50*	400-425 (50% non-respondents)	1,440-1,572 (15-20% attrition)
3	Fall '08						1,440-1,572
4	Spring '09						1,152-1,336 (15-20% attrition)
4	Fall '09						1,152-1,336
5	Spring '10						
5	Fall '10						

***OMB approval sought for these survey assessments**

2. Describe the procedures for the collection, including: the statistical methodology

for stratification and sample selection; the estimation procedure; the degree of accuracy needed for the purpose described in the justification; any unusual problems requiring specialized sampling procedures; and any use of periodic (less frequent than annual) data collection cycles to reduce burden.

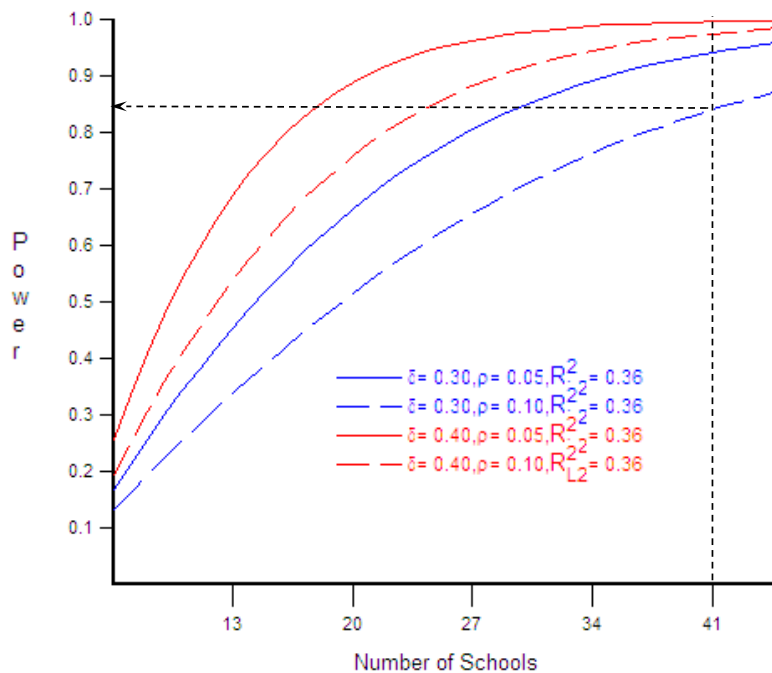
We randomly assigned 41 schools (21 treatment and 20 control schools) to the treatment groups. The total number of classrooms involved is 50. So for most of the schools, there will be only one classroom, making classroom and school almost indistinguishable for analysis purpose. Because of this, our analysis will be 2-level modeling analysis. Within each school, we project a median of 20 children, for an approximate total of 1000 children.

Dr. Xitao Fan from the Curry School of Education at the University of Virginia did the power analysis. Dr. Fan is a quantitative methodologist and a consultant to the Regional Educational Laboratory Appalachia.

Power Curves for Cluster Randomized Trials:

(Note: Power curves presented in Figure below were based on the procedures implemented in “Optimal Design for Longitudinal and Multilevel Research-V1.55”, by Raudenbush, S. W., Spybrook, J., Liu, X.F., & Congdon, R., 2005. The details of the statistical derivations for these power analysis procedures are provided in the User Guide Manual of this software.)

Figure 1. Power curves for Early Literacy Study



Major Assumptions for the Power Analysis Results:

1. Cluster Size = 25 (average number of students in each randomized school)

2. Intraclass correlation (ρ):

We assume two different levels of intraclass correlations: $\rho=0.05$ and $\rho=0.10$. The intraclass correlation represents the “clustering” effect. More specifically, these two levels of the intraclass correlation represent the situations in which 5 percent and 10 percent of the total variation in the outcome lies between clusters (schools).

3. Cluster Level Covariate:

We plan to use at least one aggregated cluster level covariate in the model. The projected cluster level covariate is the pretest score aggregated at the school level (i.e., the average pretest scores for all the participating students in a school). Research on academic achievement has overwhelmingly shown that pretest scores are highly correlated with posttest scores, and thus substantially correlated with cluster means of the posttest scores. We assume that the relationship is 0.60 ($R^2_{L2}=0.36$).

4. Effect Size Magnitude:

We assume effect size of $\delta=0.30$ and $\delta=0.40$. These two levels of effect size represent intervention effects ranging from small to medium.

5. Number of Clusters:

As previously described, we randomized at the school level. We had planned for a total of 41 schools (with about 50 classrooms in these 41 schools), 21 in the treatment condition, and the other 20 in the control condition.

Power Analysis Conclusions:

Given these parameters and assumptions detailed above, we can see that the proposed design and analysis should have sufficient power for detecting effect sizes ranging from small to medium. For the planned 41 schools to be randomized in this study, the conservative estimate of the power level is about 0.84 (for the lower level of effect size, $\delta=0.30$, and higher level of ICC, $\rho=0.10$), as indicated by the dashed horizontal arrow imposed on the power curves. In statistical analysis, statistical power of 0.80 and above is considered sufficient (e.g., Cohen, 1988). Our estimated power level of 0.84 would allow us to tolerate some degree of attrition. For each one-year cohort, we do not anticipate substantial attrition, because attrition typically occurs across academic years (e.g., teacher moving, changing jobs, students moving, etc.). We expect that, for this one-year cohort, the attrition will primarily occur at the children’s level. Even for about 15% attrition (1000 children at the beginning decreases to 850 children at the end of the one-year experiment), we will still maintain an adequate level of statistical power level at about 0.81.

Describe the assumptions made regarding fixed effects estimates (findings limited to the sites) as opposed to random effects estimates (findings have generalizability).

In our design, schools are randomly assigned to either the control or treatment group conditions. To be consistent with research being conducted at other RELs, schools will be treated as fixed effects. Because we will use multi-level analysis (HLM) as our analytic tool, correct standard errors should be obtained to account for possible school-level clustering. Furthermore, we believe treating school as a fixed effect is appropriate because we expect little or no spillover effects due to the location of treatment and control teachers in different schools.

Describe the assumptions made regarding intra-class correlations.

We will use Schochet's (2005) definition of intraclass correlation (ICC), namely the amount of variability of children's outcome measure(s) across classrooms relative to the total variation of children's outcome measure(s). As noted by Schochet (2005), ICCs vary by both type of data source and grade level. However, they do typically become smaller when adjusted for district fixed effects. Finally, Schochet (2005) notes that when standardized tests are used as outcomes, ICCs are typically between .10-.20.

We assume two different levels of intraclass correlations: $\rho=0.05$ and $\rho=0.10$. The intraclass correlation represents the "clustering" effect. More specifically, these two intraclass levels represent the situations in which 5 percent and 10 percent of the total variation in the outcome lies between clusters (teachers).

To make sure that our power estimates are correct, we carefully examined our assumptions for calculating the intraclass correlation (ICC). In our proposal, we assumed that the ICC to be 0.10 (i.e., 10% of the variation in the outcome variable is accounted for by clustering at the school level). A concern has been expressed that this assumption may be optimistic, and the actual ICC may be higher than 0.10. This may appear at the lower range of observed ICCs in educational data (e.g., 0.10-0.20). After our careful consideration, we believe that the assumption of 0.10 ICC is realistic for the schools we plan to target in this research project.

As discussed in our proposal, our targeted schools are Title I schools overwhelmingly comprised of disadvantaged minority students. As such, these schools are typically at the lower spectrum of academic performance. As such, there is considerably less variability in academic performance across the schools. In other words, for the schools we are targeting, there is more uniformity across these targeted schools than, say, if we involve a representative sample of schools from the full spectrum of academic performance. Because the targeted schools share considerable similarity (e.g., ethnic group composition, lower level of academic performance), the variation among these schools should be much smaller than that among all possible schools (e.g., including schools with high academic performance levels as well as those with low academic performance levels). So compared with a situation where all schools are considered, the variation among our targeted high-need schools should be smaller than that among schools in

general. Based on this consideration, we consider the hypothesized ICC of 0.10 (10% of variation in the data due to clustering, i.e., variation among the schools) reasonable, and it should not be an underestimate of the clustering effect.

Within-Study Replication of the Experiment

The power analysis described above is for the first cohort in Year 1. In the second year, we will repeat the experiment for the second cohort under the same conditions. Because the two cohorts will be similar in terms of sample size and other conditions, the power analysis results presented above are applicable for the second cohort.

We will treat the two cohorts as replications, and data analyses will be conducted separately for the two cohorts. In educational research, replication has not received the kind of attention that it deserves. In science, the ultimate test for the validity of any research results is replication; any experiments that cannot be reliably replicated are considered very suspicious. A well-known example is the cold fusion “breakthrough” reported from University of Utah in the late 80s. Subsequent replication experiments throughout the world largely failed to produce the results at the level as reported by the original researchers. As a result, this once-promising line of research has largely been abandoned. In educational research, unfortunately, replication is the exception, rather the norm. This state of affairs is unfortunate, and is likely detrimental to the reputation of educational research.

Our design of including a replication in the second cohort adds a design mechanism for checking the internal validity and generalizability of the results from our first cohort. We believe that this replication mechanism in our design contributes substantially to the validity and generalizability of any outcomes from our experiment. If the results from both cohorts converge, that will speak volumes about the validity of results of the study. If the convergence does not occur, that will alert us of potential issues/problems in the effectiveness of the treatment. Finding stronger effects in the second cohort than in the first would not be surprising, however, given the extensive literature demonstrating implementation challenges, often associated with actual dips in performance (i.e., negative effects), the first year an intervention is put into practice by schools. In either case, our conclusions from the study will be considerably strengthened by the addition of this replication mechanism in the study design and implementation.

A variety of factors other than the intervention itself may impact a child’s acquisition of early literacy skills. As part of our design, both in terms of sampling and subsequent analysis, we propose assessing covariates in order to refine our assessment of potential program impact. At the program level, programs will be stratified on the following criteria: percent minority and economically disadvantaged. All of our analytic models will control for the possible influence of these covariates. The purpose of the study requires baseline assessment and end-of-school-year assessment to gauge children’s progress in the development of early literacy. Every effort is being made to reduce the paperwork burden yet meet the needs of the study.

3. Describe the methods used to maximize response rates and to deal with nonresponse. The accuracy and reliability of the information collected must be shown to be adequate for the intended uses. For collections based on sampling, a special justification must be provided if they will not yield "reliable" data that can be generalized to the universe studied.

The anticipated response rate is very high for this collection. The primary reason is that the respondents are readily available. Teachers and paraprofessionals are in classrooms daily. Written and verbal reminders to these school district staff to complete the on-line surveys will be provided as needed.

Direct child assessments will be administered by trained evaluators to preschool children. Surveys will be conducted by trained evaluators as well. Individually administered reading tests that are valid, reliable, and age-appropriate will be administered to preschool students, providing profiles of early language and literacy skills. For assessing achievement of preschool children, pre- and post –assessment results from the Peabody Picture Vocabulary Test – Third Edition and The Phonological Awareness Literacy Screening for Preschool will be used. These assessments are approved by the US Department of Education as outcome measures for the ERF Grant. For assessing achievement of kindergarten through second grade children, pre- and post –assessment results from the DIBELS and Woodcock-Johnson III will be used. Student level variables will be analyzed and used to provide measurement of differences between student sub-groups, including age, race, gender, socioeconomic and English as Second Language (ESL) status, and children with disabilities/special needs.

The validity and time requirements of all surveys are based on their application and psychometric studies in prior Early Reading First and Reading First evaluation studies in Tennessee, Oklahoma, Wisconsin, Louisiana, and Texas. Teachers will be administered a survey of approximately 20 closed-ended and three open-ended questions focusing on perceptions of professional development, resources, pedagogical change, outcomes, and support pertaining to the Early Literacy program. In addition, we will collect information on teacher training, level of education, length of employment, and other professional development activities. The survey administered to paraprofessionals will be approximately 10 closed-ended and three open-ended questions focusing on perceptions of professional development, resources, pedagogical change, outcomes, and support pertaining to the Early Literacy program. In addition, we will collect information on teacher training, level of education, length of employment, and other professional development activities. The school district will administer to parents a survey of approximately 20 closed-ended and three open-ended questions to assess family perceptions of the literacy and preschool program.

Classroom Observations will utilize published instruments with validity and reliability documentation. The Early Language and Literacy Classroom Observation (ELLCO) contains three assessments tools: a Literacy Environment Checklist, a protocol to conduct classroom observation and administer teacher interviews, and a Literacy Activities Rating Scale. It is used for research purposes in over 150 preschool classrooms; reliability was 90 percent or better (ELLCO, 2004). The Early Literacy Observation Tool (E-LOT) was designed to assist schools in evaluating the effectiveness of teacher training and implementation of research-based strategies. It has been aligned to the National Reading Panel and National Research Council findings (Smith, Ross, & Grehan, 2003) and captures all essential components of the Early Reading First program. A reliability study using Kappa statistic and intraclass correlation (icc) examined interrater reliability for the Early Literacy Observation Tool (E-LOT) was conducted by Huang, Franceschetti, and Ross (2006).

4. Describe any tests of procedures or methods to be undertaken. Tests are encouraged as effective means to refine collections, but if ten or more test respondents are involved OMB must give prior approval.

Instruments for child assessment are published tests that have documented validity and reliability studies. Teacher and paraprofessional surveys were validated when used for prior studies and were approved by MCS. The average time required for completion of each survey was determined at previous administrations.

5. Provide the name and telephone number of individuals consulted on the statistical aspects of the design, and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.

Project Officer: Sandra Garcia, IES	sandra.garcia@ed.gov
TWG: Johannes M. Bos, Ph.D.	hans@bpacal.com
Laura M. Desimone, Ph.D.	l.desimone@vanderbilt.edu
Barbara D. Goodson, Ph.D.	Barbara_Goodson@abtassoc.com
Rebecca A. Maynard, Ph.D.	rmaynard@gse.upenn.edu
Samuel C. Stringfield, Ph.D.	sam.stringfield@louisville.edu

Consultants:

Dr. Steven M. Ross, President, Education Innovations, LLC
(901) 678-3413
Dr. Anna Grehan, Research Associate Professor, University of Memphis
(901) 678-4222
Dr. Sarah L. Friedman, Director, REL Appalachia
Michael Puma, Chesapeake Research Associates LLC

Name of Agency to collect and analyze the information:
Education Innovations, LLC 3161 Campus Postal Station

Memphis, TN 38152-3830
REL Appalachia, CNA Corporation

4825 Mark Center Drive
Alexandria, VA 22311

REFERENCES

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.
- Early Language and Literacy Classroom Observation (ELLCO). Technical Report. Retrieved December 7, 2004 from http://www.brookespublishing.com/store/books/smith-ellco/ELLCO_TechnicalReport.pdf
- Grehan, A. W. & Smith, L., & Ross, S.M. (2004). *The Early Reading First Literacy Observation Tool*. Memphis, TN: The University of Memphis, Center for Research in Educational Policy
- Huang, Y., Franceschetti, L., & Ross, S. (2007). *Inter-rater Reliability Analysis of the Early Literacy Observation Tool*. Memphis, TN: The University of Memphis. Center for Research in Educational Policy.
- Raudenbush, SW, Liu, X-F, Spybrook, J, Martinez, A, & Congdon, R. (2006). *Optimal Design software for multi-level and longitudinal research (Version 1.77)* [Computer software]. Available at <http://sitemaker.umich.edu/group-based>
- Schochet, PZ. (2005). *Statistical power for random assignment evaluations of educational programs*. Princeton, NJ: Mathematica Policy Research
- Smith, M.W. & Dickinson, D.K. (2002). *User's Guide to the ELLCO Toolkit: Research Edition*. Educational Development Center, INC
- Sterbinsky, A., & Ross, S. M. (2003). *Literacy Observation Tool Reliability Study*. Memphis, TN: The University of Memphis, Center for Research in Educational Policy.

Appendix

Early Literacy Study Consent Form



**Early Literacy Study
Consent Form for Teachers and Paraprofessionals**

As part of a preschool program research study conducted in Memphis City Schools by Education Innovations, LLC, you are being asked to participate in an evaluation of the program at your school this year. The focus of the study will be to determine the effectiveness of preschool programs in raising student achievement. We expect to have 50 classrooms and approximately 1,000 children participating in the study. Approximately half of the classrooms will use the *Opening the World of Learning* curriculum as part of their instructional program as a way of measuring the effectiveness of that curriculum.

If you agree to participate in the study, you will be involved in the following ways:

- (1) We will ask you to complete a brief survey about your perception of the professional development, resources, pedagogical change, outcomes, and support of the preschool program at your school.
- (2) We will send trained staff to individually assess children in your classroom on language skills to determine progress made during the year.
- (3) We will also send site researchers to observe your classroom twice during the year.

All survey responses and observation data will be considered confidential, within the limits allowed by law. When findings are described or quoted in technical reports or journal articles related to this study, neither the identity of your school nor the identity of any individual will be revealed. There are no perceived risks or costs to you associated with participating in the study. The benefits are the additional information available to you regarding each student's current skill levels and progress made by the end of the year.

If you have any questions regarding this study or the use of data collected, you may contact Dr. Steven Ross, at (901) 678-2310 or Dr. Anna Grehan, at (901) 678-4222. Questions regarding the rights of research subjects may be directed to the Chair of the University of Memphis Institutional Review Board (901) 678-2533. (Note: The University of Memphis does not have any funds budgeted for compensation for injury, damages, or other related expenses.)

I have read the information in the consent form. Any questions that I may have had have been answered. I may withdraw my participation at any time by notifying Dr. Ross or Dr. Grehan. Refusal to participate will involve no penalty or loss of benefits to which I am otherwise entitled. Similarly, the researcher may choose to terminate my participation if the study requirements recommend participation changes. I have not waived any rights by agreeing to participate.

Signature

Date

Responses to this data collection will be used only for statistical purposes. The reports prepared for this study will summarize findings across the sample and will not associate responses with a specific district or individual. We will not provide information that identifies you or your district to anyone outside the study team, except as required by law.

Attachments

Exhibit A

Memphis City Schools OWL Program Teacher Questionnaire

Please see separate file for a copy of this measure.

Exhibit B

Memphis City Schools Pre-K Program Teacher Questionnaire

Please see separate file for a copy of this measure.

Exhibit C

Memphis City Schools OWL Program Paraprofessional Questionnaire

Please see separate file for a copy of this measure.

Exhibit D

Memphis City Schools Pre-K Program Paraprofessional Questionnaire

Please see separate file for a copy of this measure.