

**A STUDY OF THE EFFECTS OF USING
CLASSROOM ASSESSMENT FOR STUDENT LEARNING
(Study 2.1a)**

OMB Clearance Package Supporting Statement

Part A: Justification

Regional Educational Laboratory
for the
Central Region

Contract #ED-06-CO-0023

Submitted to:

**Institute of Education Sciences
U.S. Department of Education
555 New Jersey Ave., N.W.
Washington, DC 20208**

Submitted by:

**REL Central at
Mid-continent Research
for Education and Learning
4601 DTC Blvd., #500
Denver, CO 80237
Phone: 303-337-0990
Fax: 303-337-3005**

Project Officer

Sandra Garcia, Ph.D.

Project Director:

Louis F. Cicchinelli, Ph.D.

Deliverable 2.3/2.4

June 1, 2007



© 2007

This report was prepared for the Institute of Education Sciences under Contract #ED-06-CO-0023 by Regional Educational Laboratory Central Region, administered by Mid-continent Research for Education and Learning. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

TABLE OF CONTENTS

| | |
|--|----------|
| A. JUSTIFICATION..... | 1 |
| Introduction..... | 1 |
| 1. Circumstances That Make Data Collection Necessary..... | 1 |
| 2. How, by Whom, and for What Purpose the Information Is To Be Used..... | 2 |
| Study Purpose..... | 2 |
| Key Research Questions..... | 2 |
| Instruments and Data Collection..... | 3 |
| a. Teacher Background Information..... | 3 |
| b. Survey of Professional Development Activities..... | 3 |
| c. CASL Participation Log..... | 3 |
| d. Test of Assessment Knowledge..... | 5 |
| e. Teacher Work Samples..... | 5 |
| f. Teacher Survey of Student Involvement..... | 6 |
| g. Student Survey of Motivation..... | 6 |
| h. State Achievement Test..... | 6 |
| 3. Use of Information Technology to Reduce Burden..... | 7 |
| 4. Efforts to Identify and Avoid Duplication..... | 7 |
| 5. Impacts on Small Businesses and Other Small Entities..... | 7 |
| 6. Consequences to Federal Programs or Policies if Data Collection is Not Conducted..... | 8 |
| 7. Special Circumstances..... | 9 |
| 8. Solicitation of Public Comments and Consultation with People Outside the Agency..... | 9 |
| 9. Respondent Payments..... | 10 |
| 10. Confidentiality Assurances..... | 11 |
| 11. Justification for Questions of a Sensitive Nature..... | 13 |
| 12. Estimates of Hour Burden of Data Collection..... | 13 |
| 13. Estimate of Total Cost Burden to Respondents..... | 15 |
| 14. Estimate of Annualized Cost to the Federal Government..... | 15 |
| 15. Reasons for Changes or Adjustments in Burden..... | 15 |
| 16. Tabulation, Analysis, and Publication Plans and Schedule..... | 15 |
| Analysis..... | 16 |
| Pre-intervention Analyses..... | 16 |
| Descriptive Statistics..... | 16 |
| Assumptions/Outliers/Data Treatment..... | 17 |
| Attrition..... | 17 |
| Fidelity..... | 17 |
| Concurrent Events..... | 17 |
| Level of Analysis..... | 18 |
| Intent to Treat..... | 18 |
| Impact of CASL on Student Achievement..... | 18 |
| Impact of CASL on Student Achievement..... | 19 |
| Impact of CASL on Teacher Outcomes..... | 20 |
| Effect Sizes..... | 20 |

Reporting.....20
 Study Report Preparation.....20
 Dissemination.....21
 Public- or Restricted-Use Data Files.....21
17. OMB Expiration Date.....21
18. Exceptions to Certification Statement.....21

references.....22

A. JUSTIFICATION

INTRODUCTION

This submission requests approval for a data collection plan for a study of classroom assessment and student achievement. The project is sponsored by the Institute of Education Sciences within the U.S. Department of Education, and will be conducted by the Central Region Educational Laboratory (contract #ED-06-co-0023) administered by Mid-continent Research for Education and Learning (McREL).

The study will examine the impact of Classroom Assessment for Student Learning (CASL) — a professional development program in classroom assessment — on student achievement and other student and teacher outcomes. Schools participating in the study will be randomly assigned to either the intervention or control school groups. Each school in the intervention group will include a team of three to six Grade 4 and 5 mathematics teachers who will implement the CASL program. Teachers in the control schools will engage in their regular professional development activities. The study will begin in Fall 2007 with baseline data collection. Implementation of the intervention will begin during the 2007–2008 academic year, and the data collection will conclude at the end of the 2008–2009 academic year.

1. CIRCUMSTANCES THAT MAKE DATA COLLECTION NECESSARY

The Regional Educational Laboratories (REL) are authorized under the Education Sciences Reform Act of 2001 (Pub. L. 107-279) Part D, National Center for Education Evaluation and Regional Assistance, Section 174, (20 U.S.C. 9564) (see Appendix A), administered by the Institute of Education Sciences. The primary mission of the REL is to serve the educational needs of each region, using applied research, development, dissemination, and training and technical assistance to bring the latest and best scientifically valid research and proven practices into school improvement efforts.

The national priority for the current 2005–2010 REL contract is addressing the goals of the recently reauthorized Elementary and Secondary Education Act. The Elementary and Secondary Education Act—that is, the No Child Left Behind Act of 2001 (NCLB)—requires all students to be proficient in state content standards in reading and mathematics by the end of the 2013–2014 school year. In order to meet this goal, schools and teachers need effective and efficient methods to help students learn and attain proficiency on state standards. Recent research suggests that the application of the principles of quality formative assessment in the classroom can provide relatively large increases in student achievement for relatively small costs. Evidence of effectiveness of teacher professional development on classroom assessment, however, is limited.

This study examines the effects of *Classroom Assessment for Student Learning (CASL)* (Stiggins, Arter, Chappuis, & Chappuis, 2004), a teacher professional development program, on student achievement. *CASL* is a widely-used, stable program for improving classroom assessment that has been used for over ten years with a number of testimonial successes. The first edition of *CASL* was published in 1994; the current and fourth edition, published in 2004,

has sold approximately 75,000 copies¹. However, despite its wide usage, no direct causal evidence of the effectiveness of *CASL* is available. This study therefore responds to the legislative intent of the Education Sciences Reform Act of 2001 regarding activities of the REL; activities include identifying successful educational programs and making such information available so that such programs may be considered for inclusion in the national education dissemination system (Pub. L. 107-279, Section 174(g)(5)) (see Appendix A).

2. HOW, BY WHOM, AND FOR WHAT PURPOSE THE INFORMATION IS TO BE USED

Study Purpose. The data, collected as part of an experimental study conducted by the Central Region Educational Laboratory (Contract #ED-06-co-0023), will be used by IES to determine the impact of the *Classroom Assessment for Student Learning* program on students' achievement in mathematics and reading in Grades 4 and 5. The study is intended to determine if *CASL* is effective in raising student achievement as measured by the state assessment system; raising student achievement is an essential component of meeting the Adequate Yearly Progress provision in NCLB. The purpose is to provide educators with rigorous evidence regarding the effectiveness of *CASL* for raising students' achievement. Evidence regarding the effectiveness of *CASL* will provide practical utility by helping guide decisions about selecting practices and programs for increasing student achievement. Below is a brief discussion of the research questions, instrumentation, and data collection for the study.

Key Research Questions. Our primary research question concerns the effect of *CASL* on student achievement. Additionally, we plan to examine the effects of *CASL* on teacher outcomes and student involvement and motivation. The research questions are as follows:

1. Does teacher participation in *CASL* have a significant impact on student achievement?
2. Does teacher participation in *CASL* have a significant impact on teacher knowledge of classroom assessment practices?
3. Does teacher participation in *CASL* have a significant impact on the quality of classroom assessment practices?
4. Does teacher participation in *CASL* have a significant impact on the extent to which students are involved in formative assessment?
5. Does teacher participation in *CASL* have a significant impact on the extent to which students are motivated to learn?

Instruments and Data Collection. Eight instruments will be used to measure implementation fidelity, teacher outcomes, and student outcomes. Table 1 provides a list of the instruments to be used in this study, their focus, associated research question (if applicable), and the schedule for collection.

¹ Stephen Chappuis, personal communication, January 2, 2007.

All instruments except the state achievement test are being pilot-tested by the study team with a small group of teachers (less than 10) to ensure that the directions are clear, the requests for data are unambiguous, the process is efficient, the time required is known in advance, and the online instruments function properly and are easy to use. The pilot test is ongoing and a report is not yet available.

a. Teacher Background Information (see Appendix B). Information on teacher characteristics will be collected at the beginning of the study. Teachers will be asked to complete an information sheet requesting general demographic information, such as gender, race, ethnicity, and years of teaching experience. Category definitions for race and ethnicity will be consistent with OMB statistical classifications.

b. Survey of Professional Development Activities (see Appendix C). The primary purpose of this survey will be to gather information regarding all teacher professional development activities that occur outside the *CASL* program. Data from this survey will be used to describe and examine any confounding events occurring during the course of the study, such as professional development in mathematics pedagogy or alignment with state standards and assessment. Data from this survey also will be used to identify any cross-over of the control teachers who may receive all or part of the intervention or a similar professional development emphasizing classroom assessment. Teachers will be asked to record the types of professional development activities, their frequency, duration, subject area, emphasis, perceived quality, and perceived impact on classroom practice.

All teachers, intervention and control, will be asked to complete the survey at the end of each semester. Survey responses will provide information on activities that occurred during that semester. Completing the survey once each semester will likely provide more accurate information than would be obtained from a single survey administration at the end of the year.

c. CASL Participation Log (see Appendix D). Teachers in the intervention group will complete a brief log regarding the fidelity of implementation of the *CASL* program. Data from the log will be used to: a) capture the degree to which teachers in the intervention group are implementing the *CASL* program as prescribed by the developer, b) describe variations in the implementation of the *CASL* program, and c) to monitor possible treatment attrition—teachers assigned to the intervention group who do not complete the intervention program as defined. Items in the log are based on characteristics of effective professional learning communities identified in previous research (DuFour, 2005; Schmoker, 1999, 2005, 2006).

Table 1. Summary of Instruments

| Instrument | Focus | Research Question | Year 1 | | Year 2 | |
|--|---------------------------------------|-------------------|-----------|-------------|-----------|-------------|
| | | | Fall 2007 | Spring 2008 | Fall 2008 | Spring 2009 |
| a. Teacher Background Information | Background information | | X | | | |
| b. Survey of Professional Development Activities | Implementation fidelity | | X | X | X | X |
| c. CASL Participant Log* | | | I | I | | |
| d. Test of Assessment Knowledge | Teacher knowledge | #2 | X | X | | X |
| e. Teacher Work Sample | Teacher classroom assessment practice | #3 | FT | | | X |
| f. Survey of Student Involvement | Student involvement | #4 | FT | X | X | X |
| g. Student Survey of Motivation | Student motivation | #5 | | X | | X |
| h. State Achievement Test | Student achievement | #1 | X** | X | | X |

Note. X = Intervention and control school teachers. FT = Field Test. I = Intervention teachers only.

* = The CASL participant log will be completed each time the intervention teachers complete a chapter of the CASL textbook.

** = The CSAP is administered in the spring of each year; Spring 2007 CSAP results will be collected by researchers from participating schools' districts as soon as they become available.

Teachers will be asked to complete an initial log regarding the establishment of the learning team and one log after completing each of the 13 chapters of the *CASL* textbook. In this manner, teachers will be reporting on each segment of their work with *CASL* while it is still fresh in their minds, thereby improving the quality of data collected. A log is a theoretically sound and common practice for measuring teacher activities as part of a professional development program (York-Barr, Sommers, Ghere, & Montie, 2001).

d. Test of Assessment Knowledge (see Appendix E). The Test of Assessment Knowledge, developed by the study team, will measure teachers' knowledge of classroom assessment. This instrument will be used as a pretest to measure teacher knowledge at the beginning of the study and as an outcome to estimate the effects of the intervention. An original instrument was developed in order to ensure that the measure is well-aligned with the intervention and sensitive to its effects. An established and proven instrument sufficiently well-aligned with the intervention does not exist. The lack of alignment between existing instruments and the intervention was determined to be a threat to construct validity.

The test includes multiple-choice and true-false items. Items were developed to sample the knowledge and reasoning skills represented in the *CASL* program. The test will give more weight to topics that are described in depth, comprise a large domain of information, and are critically important to the *CASL* program. Although the test will be aligned to the *CASL* program, the test will measure generally accepted principles and practices of classroom and formative assessment. Terminology specific only to the *CASL* program will be avoided in the items. Appendix E includes the seventy-two items developed and pilot-tested. From these, approximately 50 items will be chosen to be field-tested in the fall of 2007.

The test will be administered three times over the course of the study. The first administration will serve as the field test and will provide baseline data on teachers' assessment knowledge, pre-intervention. The second administration, in spring 2008, will measure any gains that teachers have made in their assessment knowledge immediately after completing the professional development. The third administration, in spring 2009, is necessary to measure teachers' post-intervention level of knowledge.

e. Teacher Work Samples (see Appendix F). Data from the Teacher Work Samples will be used to measure the effects of the *CASL* professional development on the quality of teachers' classroom assessment practices. Prior research suggests that systematically collecting samples of graded student work is an efficient, reliable, and valid way to find out what is happening in classrooms when teachers engage in assessment. Although teachers participating in *CASL* will learn and apply the content to all content areas and the *CASL* professional development is expected to help raise student achievement in all content areas, we are focusing data collection on mathematics only to reduce data burden.

Teacher work samples will be collected twice during the study. The first collection will occur in the fall of 2007; these data will be used to examine the psychometric functioning of the instrument and collect work samples for training anchor papers. Teachers also will be asked to submit samples of their work in the spring of 2009 to examine the impact of *CASL* training on teacher assessment practices.

f. Teacher Survey of Student Involvement (see Appendix G). The data on frequency of student involvement will be used to determine the impact of *CASL* on the extent to which students themselves are involved in identifying/discussing learning targets, using scoring guides, and explaining what they need to do to improve their mathematics achievement. Inclusion of this student outcome measure is necessary to fully assess the impact of *CASL*. Teachers will be asked via the survey to report activities related to classroom assessment. The survey will include 15 to 20 items asking teachers to record the number of different days during the previous two-week period that they involved their students in activities such as discussing the learning objectives, evaluating their own work using scoring guides or rubrics, and revising work to correct errors. The sum of the number of days from each item will represent the score for each administration of the survey.

Teachers in both the intervention and control schools will be asked to complete a survey once each semester. The first data collection will provide data to examine the psychometric functioning of each item and the survey as a whole. The second data collection, at the end of Year 1, will provide an estimate of the impact of the Year 1 *CASL* training on student involvement. The average of the two surveys administered during Year 2 will be used to estimate the difference in student involvement between the treatment and control schools. The two data collections will provide a more reliable estimate of student involvement during Year 2 than would only one collection. The average of the scores from the two collections will help reduce event sampling error that may be present in any individual data collection.

g. Student Survey of Motivation (see Appendix H). This survey, to be completed by students, will be used to measure the effects of the *CASL* program on student motivation in mathematics. This survey consists of items from the Ongoing Engagement and Perceived Autonomy (Self-Regulation) subscales of the elementary student Research Assessment Package for Schools (RAPS-SE) (IRRE, 1998) and the Academic Efficacy subscale of the Patterns of Adapted Learning Scales (PALS) (Midgley et al., 2000). Both the RAPS-SE and PALS have been used extensively in education research, with results published in peer-reviewed journals. The items taken from these extant instruments have been adapted for our study to refer specifically to mathematics. This adaptation was necessary to focus data collection on mathematics to reduce the respondent burden, and also to be consistent with the theory that self-efficacy (Bandura, 1997; Pajares, 1996) and engagement and perceived autonomy (Ryan & Deci, 2000) are domain-specific. The Student Survey of Motivation will be administered at the end of each year as a post-intervention measure only, in order to keep data collection to a minimum.

h. State Achievement Test. Students' scale scores from Colorado's state assessment system will be used as the measure of impact of *CASL* on student achievement. This assessment, the Colorado Student Assessment Program (CSAP), provides scores on a vertical scale from Grades 3 through 8 and has four performance levels that define proficiency on the state content standards. Using state assessment results as the measure of the ultimate outcome—student achievement—will allow the researchers to determine the extent to which implementation of the *CASL* program impacts student achievement in relation to the goals of NCLB. Data will be collected for Reading, Writing, and Mathematics. Other demographic data on students, including eligibility for free or reduced-price lunch and disability status, will be requested for inclusion in the achievement data files. Advance copies of this test are not available and so do not appear in an appendix (Colorado Department of Education, no date).

3. USE OF INFORMATION TECHNOLOGY TO REDUCE BURDEN

Data collection activities have been planned to maximize efficiency and minimize burden while still providing the information necessary to conduct a rigorous study of the effectiveness of CASL. Data collection will occur online whenever possible to minimize the reporting burden for participants. The online data collection will reduce participant burden because they will not have to manage paper documents or mailing activities. Teachers will receive an e-mail message providing them with a link to the data collection instrument and a requested timeline for completion. Acknowledgements of receipt, reminders, and other communication can be received without addition to the current paperwork burden for teachers, and implementation fidelity can be monitored in an ongoing manner. Researchers will be able to tailor distribution of reminders so that only non-responders will be contacted.

4. EFFORTS TO IDENTIFY AND AVOID DUPLICATION

The purpose of this data collection is to determine if *CASL* has a significant effect on student achievement. Although *Classroom Assessment for Student Learning* is a widely-used program, there have been no previous studies that systematically and rigorously evaluated the impact of *CASL* using randomized field trials—the preferred method for answering causal questions about the effectiveness of programs. In addition, no other regional educational laboratories are studying this intervention.

To the extent possible, the study team is using existing data to avoid duplicating data collection efforts; for example, we are using state achievement test scores rather than administering an additional achievement test solely for the purposes of this study. However, the study team must also collect other new data targeting the research questions in order to perform the study. The information to be collected from students and teachers is not available elsewhere. The information to be collected will represent the knowledge and practice of teachers in the area of classroom assessment as well as student motivation.

Data collection is repeated at several points during the study. Each instance of data collection serves an important purpose. Baseline data collection will be used for checking the results of random assignment and controlling for pre-intervention characteristics of participants in the data analyses. Data collected at the end of Year 1 will allow for the examination of the effects of the *CASL* program after one year of training. The final data collection at the end of Year 2 will be used to determine effects of *CASL* on after one year of training and after an additional year of implementation in the classroom.

5. IMPACTS ON SMALL BUSINESSES AND OTHER SMALL ENTITIES

No small businesses will be included in our sample. The primary respondents in our study will be teachers and students, resulting in some burden on schools, which are small entities. This burden will be minimized by using online forms of data collection whenever possible, and limiting requests for data to only those data to be included in pre-specified analyses, with survey

and log length kept to a minimum. In addition, participating school districts will be asked to provide state assessment data on the students of the participating teachers. The burden placed on district staff in retrieving these data will be minimized by requesting existing assessment results and carefully specifying the study information needs.

6. CONSEQUENCES TO FEDERAL PROGRAMS OR POLICIES IF DATA COLLECTION IS NOT CONDUCTED

Through the Regional Educational Laboratory system, the Institute of Education Sciences seeks to provide policymakers and practitioners with training and technical assistance on scientifically valid, research-based ways to meet the goals of NCLB. NCLB requires that schools and districts measure academic performance in reading and mathematics in Grades 3 through 8, identify weaknesses, and make appropriate changes. The Education Sciences Reform Act of 2001 requires that REL training, technical assistance, and dissemination of information relevant to planning changes be based on the highest-quality evidence as defined by principles of scientifically valid research. Where such evidence is not available, the RELs are expected to fill the void with applied research and development. Without the data and findings from this study, REL Central will be unable to disseminate scientifically valid research on the CASL program’s effectiveness for improving student academic performance. Without the data and findings, the RELs will be limited in their ability to promote the use and application of scientifically valid research to improve classroom practice.

The following table summarizes the frequency with which each collection will be conducted for this study, as well as the consequences of collecting data less frequently (where applicable):

Table 2. Frequency of Data Collection

| Focus | Instrument | Number of collections | Consequences if administered less frequently |
|--------------------------------|--|-----------------------|---|
| Background information | 1. Teacher Background Information | 1 | Not applicable |
| Implementation fidelity | 2. Survey of Professional Development Activities | 4 | <ul style="list-style-type: none"> ▪ Less accurate data on teachers’ professional development activities |
| | 3. CASL Participant Log | 13 | <ul style="list-style-type: none"> ▪ Less accurate data on teachers’ CASL activities |
| Teacher outcomes | 4. Test of Assessment Knowledge | 3 | <ul style="list-style-type: none"> ▪ An inability to measure changes in teacher knowledge after completing the training year (Year 1) ▪ An inability to measure changes in teacher knowledge after a full year of implementation (Year 2) |
| | 5. Teacher Work Sample | 2 | <ul style="list-style-type: none"> ▪ An inability to field test the instrument ▪ An inability to control for teachers’ prior knowledge of classroom assessment |

Table 2. Frequency of Data Collection (cont'd)

| Focus | Instrument | Number of collections | Consequences if administered less frequently |
|------------------|----------------------------------|-----------------------|--|
| Student outcomes | 6. Survey of Student Involvement | 4 | <ul style="list-style-type: none"> ▪ An inability to field test the instrument ▪ An inability to assess the impact of CASL on student involvement at end of training year (Year 1) ▪ A higher probability of an event sampling error, if only one data collection from Year 2 is used to estimate the difference in student involvement between the treatment and control schools |
| | 7. Student Survey of Motivation | 2 | <ul style="list-style-type: none"> ▪ The inability to measure changes in student motivation |
| | 8. State Achievement Test | 3 | Not applicable – Administered by state |

7. SPECIAL CIRCUMSTANCES

The study requires respondents to 1) report certain information more often than quarterly and 2) complete instruments within 14 days of their receipt.

Only one of the eight data sources utilized in this study, the CASL Participant Log (see Item A2 for more detail), requires respondents to report the requested information more often than quarterly. The anticipated frequency for completing the logs is approximately once every three weeks. Treatment teachers will be asked to complete the log after completing each of the 13 chapters in the CASL program. Teachers will be reporting on each segment of their work with CASL while it is still fresh in their minds, thereby improving the quality of data collected. Although teachers will be completing a log entry more often than quarterly, they will not be completing a log entry for the same CASL chapter more than once.

Except for submission of student achievement data, respondents are asked to complete instruments within 14 days of their receipt. Many of the surveys ask respondents to report on their activities; asking respondents to complete the survey within 14 days of their receipt should improve the accuracy of the data provided. In addition, survey methodology research suggests that giving respondents a longer window, such as a month, to respond to a survey makes it more likely that they will forget or lose the request (Dillman, 2000).

No other special circumstances apply to this study.

8. SOLICITATION OF PUBLIC COMMENTS AND CONSULTATION WITH PEOPLE OUTSIDE THE AGENCY

A 60-day notice was published in the Federal Register on April 24, 2007, with an end date of June 24, 2007 to provide the opportunity for public comment. Comments were received in

regards to _____, _____, and _____. Revisions made to the study to address these comments include _____, _____, and _____. See Appendix I for copies of the Federal Register notices pertaining to this study, numbers _____ and _____.

In addition, throughout the course of this study, we will draw on the experience and expertise of a technical working group (TWG) that will provide a diverse range of experience and perspectives as well as expertise in relevant methodological and content areas. The first meeting of the TWG was held from May 31 through June 2, 2006. The second meeting of the TWG was held from September 5 through September 7, 2006. The members of this group are:

Dr. Geoffrey Borman, Associate Professor, University of Wisconsin-Madison

Dr. Susan Brookhart, Coordinator of Assessment & Evaluation, School of Education, Duquesne University

Dr. Robert D. (Robin) Morris, Vice President of Research, Georgia State University

Dr. Barbara Plake, Professor Emeritus, University of Nebraska

Dr. Andrew Porter, Director, Learning Sciences Institute, Vanderbilt University

Dr. Robert St. Pierre, President, STP Associates

9. RESPONDENT PAYMENTS

We understand that our study will place burdens of participation and data collection on teachers in the study. Intervention-group teachers will engage in several activities that are not part of their regular duties, including attending team meetings, reading the *CASL* textbook, completing logs and surveys, submitting samples of their assessment work, and administering surveys to students. Control-group teachers will also participate in these activities, with the exception of reading the textbook. In addition, teachers commit to the study for a period of two years, and it is important to retain them in the study for the entire time period so they can provide the needed data. We have therefore decided to provide payments to the teachers to encourage them to volunteer initially, to encourage them to stay in the study, and as partial compensation for their time for the efforts to provide data; these efforts are above and beyond what they are required to do as part of their jobs as teachers. There will be no remuneration for students involved in the study.

Reduced response rates, coupled with teacher attrition, may weaken the causal inferences of the proposed study (What Works Clearinghouse, 2006). Severe overall attrition would reduce statistical power such that a potentially important and substantial effect would go undetected. We believe that providing honoraria to encourage teachers to respond to and remain in the study will lessen attrition and improve compliance.

Prior research indicates that payments to survey respondents increase response rates and reduce nonresponse bias, even in government-sponsored research (James & Bolstein, 1992; King & Vaughan, 2004; Shettle & Mooney, 1999; Teisl, Roe, & Vayda, 2006). For example, in their study of the effects of monetary compensation on subjects' willingness to participate in both a market research interview and a follow-up survey, Wiseman, Schafer, & Schafer (1983)

concluded that “monetary (compensation) has both an immediate and a carryover effect which increases the likelihood of respondent cooperation.” Whiteman et al. (2003) reported that the use of monetary compensation significantly improved response rates in a randomized trial of compensation for completion of a mailed questionnaire. Payments have also been shown to increase the number and length of written comments in surveys that included open-ended items (James & Bolstein, 1990), like those included in the present study. The payment helps the research by putting the data collection effort into a social context in which the respondent is inclined to reciprocate for the payment, or it may indicate to the respondent how important the data collection is to the study sponsor (Groves, Cialdini, & Couper, 1992).

Compensation will be provided only under OMB guidelines as stated in the “Standards and Guidelines for Statistical Surveys” document released in September, 2006, under Guideline 2.3.2 (Office of Management and Budget, 2006).

The compensation plan is summarized in Table 3. Treatment teachers will be compensated approximately \$10 for each of the 28 data collection responses they will provide during the two years of the study. Control teachers will receive an equivalent amount to encourage their continued participation in the study. The amount of total compensation also approximates the hourly payment rate for teachers (\$30) times the number of hours required to respond to requests for data. The compensation will be given in four payments of increasingly larger increments in order to encourage participants to provide data throughout the entire study.

Table 3. Compensation Schedule

| | Fall 07 End of semester | Spring 08 End of semester | Fall 08 End of semester | Spring 09 End of semester | |
|-----------|-----------------------------------|-------------------------------------|-----------------------------------|-------------------------------------|-------|
| Activity | Baseline data collection | Spring data collection | Fall data collection | Spring data collection | Total |
| Treatment | \$50 | \$75 | \$75 | \$100 | \$300 |
| Control | \$50 | \$75 | \$75 | \$100 | \$300 |

10. CONFIDENTIALITY ASSURANCES

Because this study requires associating student-level data with teacher-level data, we have a system for identifying these units and managing these linkages. In accordance with Title 34, Code of Federal Regulations, Part 97, Protection of Human Subjects, which includes Subpart A, Basic Policy, and Subpart D, Additional Protections for Children, all districts, schools, classroom teachers, and parents of students who provide data for this study will be assured, in writing, that the information provided will not be released in a form that identifies individual students, teachers, schools, or districts, except as required by law.

McREL follows the confidentiality and data protection requirements of IES (The Education Sciences Reform Act of 2002, Title I, Part E, Section 183). McREL will protect the confidentiality of all information collected for the study and will use it for research purposes only. No information that identifies any study participant will be released. Information from participating institutions and respondents will be presented at aggregate levels in reports.

Information on respondents will be linked to their institution but not to any individually identifiable information. No individually identifiable information will be maintained by the study team. All institution-level identifiable information will be kept in secured locations and identifiers will be destroyed as soon as they are no longer required. McREL obtains signed NCEE Affidavits of Nondisclosure from all employees, subcontractors, and consultants that may have access to this data and submits them to our NCEE COR.

This study has been designed to protect against risks to participants' confidentiality. Procedures to protect confidentiality include the following:

- Responses to online surveys are stored on McREL's server, where access by non-McREL employees is not possible. Access to online survey data is password protected for access by authorized research team personnel only.
- Data collected by participating teachers will be submitted directly to McREL researchers via the postal service in sealed stamped envelopes.
- Each school, teacher and student participant will be assigned an ID number and all identifying information will be stripped from data files. No individually-identifiable information will be kept in the data files.
- Participant ID number lists will be kept in locked cabinets or in password-protected computer files and not released by authorized research team personnel.

These procedures will be codified in a memorandum of understanding (MOU) to be signed with each participating district and school. Further, we will inform the parents of each participating child about the study and the procedures that will be followed to ensure data confidentiality.

The MOU between the Study PI and authorized personnel from each participating district and school will be used to help ensure clarity of expectations, roles, and responsibilities (see Appendix J for a copy of the MOU). The MOU identifies the responsibilities of McREL and each participating school. For example, the MOU states that McREL will provide treatment schools with *CASL* materials. The MOU further states that McREL will use responses to data collection only for statistical purposes, will summarize findings across the sample, and will neither associate responses with individuals, nor provide information that identifies districts, schools or individuals to anyone outside the study team, except as required by law. The MOU states that schools will comply with random assignment to treatment or control groups and adhere to study procedures according to assignment. The MOU states that in treatment schools, all 4th and 5th grade teachers will use the *CASL* materials in a learning team and complete all data collection activities; in control schools, all 4th and 5th grade teachers will continue their usual professional development practice, complete data collection activities, and refrain from using *CASL* until after study completion.

Informed consent will be sought and obtained from participating teachers and their students' parents. The teacher informed consent is active and the parent informed consent is passive (see Appendices K and L for copies of the teacher and parent/student informed consent letters). This will take place after school principals have approved the conduct of the study in their schools and have communicated expectations about study participation (including study duration, data

collection procedures, treatment and control groups, random assignment, and the fact that study participation is voluntary) to teachers.

The Principal Investigator will monitor research team members' compliance with the procedures identified in the above bullets and report the status of compliance annually to McREL's Institutional Review Board (IRB).

A System of Records Notification is not required for this study. According to the Privacy Act of 1974, Public Law 93-579 (5 USC Sec. 552a), this data collection does not meet the definition of a "systems of records" because information will not be retrieved from it by the names of individuals or any other identifying information (see Appendix M, highlighted text).

11. JUSTIFICATION FOR QUESTIONS OF A SENSITIVE NATURE

No questions of a sensitive nature will be included in the study.

12. ESTIMATES OF HOUR BURDEN OF DATA COLLECTION

The study design calls for collecting data through eight separate protocols, two of which will be field tested in fall 2007. In all of these activities there is some burden placed upon individuals—teachers, students, or district administrative staff. Table 4 below presents the anticipated respondent burden associated with these activities. The average total number of responses for this study is 8,800 per year (17,600 for the entire 2-year study). The average annual hour burden for respondents is 1,952 hours per year (a total of 3,905 hours spread over two years). The average annual cost to respondents is \$31,069 per year (a total of \$62,139 spread over two years).

Table 4. Estimated Respondent Burden Over Both Years²

| Data Collection Activity | Number of Respondents per Data Collection | Number of Data Collections | Total Number of Responses | Time per Response (in minutes) | Total Hour Burden | Hourly Rate (CDE, 2004; 2005) | Total Monetary Burden |
|---|---|----------------------------|---------------------------|--------------------------------|-------------------|-------------------------------|-----------------------|
| Teacher background information sheet | 256 | 1 | 256 | 5 | 21 | \$30.04 ³ | \$641 |
| Survey of professional development activities | 256 | 4 | 1,024 | 10 | 171 | \$30.04 | \$5,126 |
| CASL participant log | 128 | 14 | 1,792 | 10 | 299 | \$30.04 | \$8,971 |
| Test of assessment knowledge | 256 | 3 | 768 | 30 | 384 | \$30.04 | \$11,534 |
| Teacher work samples | 256 | 1 | 256 | 60 | 256 | \$30.04 | \$7,690 |
| Field test of teacher work samples | 256 | 1 | 256 | 60 | 256 | \$30.04 | \$7,690 |
| Teacher survey of student involvement | 256 | 3 | 768 | 10 | 128 | \$30.04 | \$3,845 |
| Field test of student involvement survey | 256 | 1 | 256 | 10 | 43 | \$30.04 | \$1,282 |
| Student survey of motivation | 3,840 | 2 | 7,680 | 10 | 1,280 | NA | NA |
| Memorandum of Understanding/State wide achievement test score reporting | 64 | 3 | 192 | 120 | 384 | \$25.22 ⁴ | \$9,684 |
| Teacher Informed Consent | 256 | 1 | 256 | 10 | 43 | \$30.04 | \$1,292 |
| Parent/Student Informed Consent | 3,840 | 1 | 3,840 | 10 | 640 | \$6.85 ⁵ | \$4,384 |
| Total | | 35 | 17,600 | | 3,905 | | \$62,139 |

² Using the values from the table to calculate the Total Monetary Burden may result in slightly different totals than appear in the table due to rounding differences.

³ Teacher hourly rate calculated by dividing the daily rate of Regular Classroom Teachers by eight.

⁴ Hourly rate calculated by dividing by eight the average of the daily salaries of the three types of school staff judged most likely to be tasked with retrieving student achievement results (Instructional Program Coordinators/Supervisors, Computer Technology Staff, and General Office Support).

⁵ Parent hourly rate based on June 27, 2007 Colorado minimum wage of \$6.85.

13. ESTIMATE OF TOTAL COST BURDEN TO RESPONDENTS

There are no respondent costs associated with this data collection other than the hour and monetary burden estimated in item A.12 above (see Table 4).

14. ESTIMATE OF ANNUALIZED COST TO THE FEDERAL GOVERNMENT

The estimated cost to the federal government of conducting the Study of the Effects of *Classroom Assessment for Student Learning* is approximately \$2 million across the entire course of the study. The annual cost is approximately \$500,000 per year with larger budgets for the middle years of study when data collection and analysis occur. The total amount includes the following amounts:

| | |
|--|-----------|
| Semi-annual meetings with REL Directors and Department of Education, planning, development, document review and revision, and consultations with Mathematica and IES | \$164,000 |
| Consultation with Technical Working Group | \$175,000 |
| Recruitment of sites | \$94,000 |
| IRB and OMB approval processes | \$44,000 |
| Baseline data collection and random assignment | \$17,000 |
| Intervention materials, training, and implementation | \$127,000 |
| Data collection in Years 1 and 2 | \$509,000 |
| Data analysis | \$567,000 |
| Report preparation | \$369,000 |

15. REASONS FOR CHANGES OR ADJUSTMENTS IN BURDEN

This request is for a new information collection.

16. TABULATION, ANALYSIS, AND PUBLICATION PLANS AND SCHEDULE

The study aims to begin baseline data collection in the fall of 2007. Training in the *CASL* program will occur during the 2007–2008 school year in participating schools randomly assigned to the intervention. Data collection will continue during the 2008–2009 school year. Final outcome data for the examination of the effects of *CASL* will be collected in the spring of 2009. Table 5 below presents the schedule for the major activities that will occur over the course of the study. Descriptions of data analysis are provided below.

Table 5. Schedule of Activities

| Activity | Schedule |
|---|--------------------------------|
| Approval of research study design | January 2007 |
| Recruitment and selection of sites | August 2006-August 2007 |
| Final list of Participating Sites | August 2007 |
| Baseline data collection | September 2007 |
| Training in CASL program in intervention sites and Year 1 data collection | October 2007 – June 2008 |
| Implementation of CASL program in intervention sites and Year 2 data collection | September 2008 – June 2009 |
| Retrieval of student test score data (Year 2 outcome) | August 2009 |
| Complete data files for analysis | September 2009 |
| Analyses | September 2009 – December 2009 |
| Draft Technical Reports | March 2010 |
| Final Technical Report and data file with accompanying codebook | May 2010 |
| Draft Non-Technical Report | April 2010 |
| Final Non-Technical Report | June 2010 |

Analysis. The data collected as part of this study will be analyzed to estimate the impact of the CASL program. The primary emphasis of the data analysis will be to estimate the effects of the CASL program on the academic achievement of the students in Year 2 classrooms. Additional analyses will be conducted to examine the effects of CASL on the Year 2 teacher outcomes. Analysis of the effect of CASL on both teacher and student outcomes at the end of Year 1 will be used to gauge progress of the intervention and the effects after one year of training. Prior to data analysis, data files will be examined to ensure data quality. Data management and analyses procedures will be documented for quality control and reporting.

Pre-intervention Analyses. Data analyses will be conducted prior to the implementation of the intervention in order to compare the intervention and control schools on teacher knowledge of classroom assessment and student achievement. Identifying pre-intervention differences will be important to accurately estimating the effect of the intervention. Any large differences between the two groups will warrant a review of the random assignment procedure and possible use of statistical methods to adjust for pre-intervention differences. The distribution of intervention and control schools within each district will also be checked prior to intervention.

Descriptive Statistics. Descriptive statistics will be produced for both groups on all instruments included in the study; descriptive statistics are useful for examining the properties of the data collected and informing interpretations of results. Descriptive statistics (e.g., means, standard deviations, frequency distribution, item-total correlations, and internal consistency) will be used to examine the psychometric characteristics of all instruments used in the study to determine whether the instruments are capturing variability in response, whether there are floor

or ceiling effects, whether the items within a particular instrument are tapping into the same underlying variable, and whether the items are effective in discriminating between respondents who differ on the underlying variable. The mean and standard deviation of the study sample students' scale scores on the achievement tests will be compared to the mean and standard deviation of the entire state in order to examine the degree to which the study sample is similar to the target population in terms of average academic achievement and variations in academic achievement. Comparisons of means and standard deviations will be conducted separately by grade and content area, such as Grade 4 mathematics, Grade 5 mathematics, etc.

Assumptions/Outliers/Data Treatment. Data will be examined for relevant statistical assumptions; violations of statistical assumptions can lead to inaccurate inferences. The data will also be examined for outliers. Any treatment of the data to deal with violations of assumptions or outliers will be reported.

Attrition. A number of analyses will be conducted in order to determine the rate of attrition. Attrition is a threat to internal validity and analysis of attrition will help monitor this threat. Of primary concern is attrition that results in the sample being below the minimum necessary to permit a sufficiently precise estimate of the effect size, differential attrition; and attrition that results in differences between the two groups. Analyses will be conducted and reported to examine the possibility of differential attrition where there is unequal loss of participants across the two groups. The extent of attrition from each group will be calculated, compared, and reported. Attrition can create differences between the treatment group and the control group; failure to account for any differences in groups due to attrition can bias results or lead to inaccurate estimate of effects. The sample of schools in each experimental group that completes the study will be compared to the initial sample of schools in each experimental group to determine if there are any systematic attrition-related differences between the two groups. Particular attention will be paid to the number of low-performing schools that are lost from each group.

Fidelity. Data will be analyzed and presented from the CASL Participant Logs in order to determine the number of individuals assigned to the intervention group who actually participated in the intervention. Descriptive statistics will be produced regarding the key components of the intervention program; that is, the number of participants in learning teams, meetings per learning team, activities completed, and hours spent in self-study and in classroom application, and teachers' perceptions of the usefulness of the learning team meetings. Descriptive data gathered from fidelity measures will provide information on the variations of implementation that occurred in the sample. A description of these variations will be useful for interpreting results. Analysis of the fidelity data will also help determine if the intervention was implemented in a manner consistent with its design.

Concurrent Events. Analysis of the professional development activities data from both intervention and control teachers will be used to monitor concurrent events that could pose a threat to internal validity. Descriptive statistics from the Survey of Professional Development activities will be used to monitor the duration, academic subject area, and topics emphasize to monitor the activities of participating teachers. Reports from the control teachers of professional development activities emphasizing assessment-related topics among the control group teachers would be evidence of a concurrent event that would threaten internal validity. Reports from control group teachers of participation in professional development

activities originating from the intervention developer or including intervention materials would represent a threat to internal validity. Reports from the intervention group teachers of participation in long-term and many hours of professional development activities that emphasize activities directly related to student achievement (e.g., curricula, instructional methods, or content standards) could represent a concurrent event that threatens internal validity.

Any professional development concerning classroom assessment or assessment in general will be addressed specifically. These data will be useful in examining and describing any confounding events or contamination occurring among control teachers. These data will also be useful for determining any cross-over of the control teachers who may have received all or part of the intervention.

Level of Analysis. Consistent with the random assignment of schools to either the intervention or control group, main effects will be analyzed at the school level. Outcome data will be collected at the level of the student and the teacher. Intervention effects will be estimated at the school level using hierarchical linear modeling to account for the sources of variability in the data resulting from the structure of the school environment.

Intent to Treat. The researchers will attempt to collect outcome data from all of the schools participating in the study. If schools drop out of the study, complete data from surveys, work samples, and logs will likely not be available. An effort will be made, however, to acquire student achievement data from all schools that begin the study in order to estimate the intervention's effect (i.e., intention to treat). Using data from all schools that began the study will also prevent the intervention from appearing more effective than the control condition if schools that have difficulty implementing the training drop out of the study.

Impact of CASL on Student Achievement. Effects of *CASL* on student achievement will be analyzed via a two-level hierarchical model to estimate the school level effects of the random assignment to the intervention. The Level 1 model will nest students within schools and will include the students' grade level as a predictor. Student grade level will be coded as -1.0 for Grade 4 or 1.0 for Grade 5. This coding will control for school level differences in the proportion of students in each grade level and will allow the Level 2 intercept for overall school level performance to be interpreted as the average performance of the fourth and fifth graders combined. The Level 1 model is specified as

$$Y_{ij} = \beta_{0j} + \beta_{1j}(\text{Grade})_{ij} + e_{ij}.$$

The Level 2 model will include an indicator for assignment to intervention or control school as a predictor of mean school achievement to estimate the effect of the intervention on student achievement. The Level 2 model is specified as:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{mean CSAP})_j + \gamma_{02}(\text{Treatment})_j + u_{0j},$$

where $\beta_{1j} = \gamma_{01}$.

Teacher-level effects are not included in the analysis of student achievement because simulation studies have found that clustering within intermediate units has little effect on Type I error (Murray, Hannan, & Baker, 1996). Student achievement outcome data, in both the Level 1 and Level 2 models, will include data from all students for whom data are available at the time of the outcome measure. For the analysis of the effects of the training year, the outcome measure is the state math test administered in the spring of 2008. The state test administered in the spring of 2009 will be the outcome measure for estimating the impact of *CASL* after two years.

The Level 2 model includes mean CSAP, a cluster level covariate used to explain additional between school variance not explained in the Level 1 model, to control for prior achievement, and to improve the power of the estimation of the intervention's effect (Raudenbush, Spybrook, Liu, & Congdon, 2006). The cluster level covariate represents each school's average level of achievement. For the analyses of the impact of *CASL* after Year 1, the cluster level covariate will be mean achievement scores from all the third and fourth grades students tested in the spring of 2007 at each school, Cohorts 1 and 2, respectively. The cluster level covariate for the examination of effects after two years will be mean achievement of all the third graders tested in the spring of 2007 (Cohort 1 students) and all the scores from the third graders tested in the spring of 2008 (Cohort 3 students).

Impact of CASL on Student Achievement. Student motivation is considered a secondary outcome for this study. As such, a post-test only design is considered sufficient for estimating the effects of the *CASL* program on student motivation. To reduce the data collection burden, data are not collected in the fall of 2007 or the fall of 2008 to provide a measure of prior motivation. Motivation outcomes will be examined in the spring of 2008 as well as the spring of 2009. Although one year of training is not expected to have a large impact on student motivation, the motivation of students in the two groups will be compared to estimate this effect after the training is completed at the end of Year 1. The effect of *CASL* after one year of teacher training and one year of full implementation in the classroom will be examined at the end of Year 2. As with the analysis of student achievement, the outcome data will include all students for whom outcome data is available. The covariate in the Level 2 model will be calculated at the cluster (school) level.

Analysis of the effects on student motivation will be conducted separately for Year 1 data and Year 2 data. The statistical model for both the Year 1 and the Year 2 analyses will be the same. A two-level hierarchical model will be used to estimate the school level effects of the random assignment to the intervention on student motivation. The Level 1 model will nest students within schools and will include the students' grade level as a predictor. Student grade level will be coded as described above. The Level 1 model is specified as:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(\text{Grade})_{ij} + r_{ij},$$

The Level 2 model will include an indicator for assignment to intervention or control school as a predictor of mean school achievement to estimate the effect of the intervention on student achievement. The Level 2 model will not include a covariate as this is a post-test only design. The Level 2 model is specified as:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{Treatment})_j + u_{0j},$$

where $\beta_{1j} = \gamma_{01}$

Motivation data from all students who were administered the survey in the spring of 2008 will be used in the analysis to estimate the effects of one year of *CASL* training. Data from all students who are administered the motivation survey in the spring of 2009 will be used to estimate the effects of *CASL* after two years.

Impact of CASL on Teacher Outcomes. Each teacher outcome will be compared across the intervention and control schools using separate two-level hierarchical models in order to estimate the effects of *CASL* on teacher outcomes. The Level 1 model will include teachers nested within schools. Teachers' grade level will be included in the Level 1 model to control for the effect of grade level on teacher outcomes. The Level 1 model is specified as:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(\text{Grade})_{ij} + r_{ij},$$

The Level 2 model will include a cluster-level covariate at the teacher level from the baseline measure of the test of teacher knowledge as well as dummy variable to indicate group assignment. The group assignment indicator will be used to estimate the effects of random assignment to intervention or control schools on teacher outcomes. The Level 2 model is specified as:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{mean teacher knowledge})_j + \gamma_{02}(\text{Treatment})_j + u_{0j},$$

where $\beta_{1j} = \gamma_{01}$.

The effects of the *CASL* program on each teacher outcome will be examined using data from spring 2008 to estimate the effects of one year of *CASL* study. Separate analysis of spring 2009 teacher outcome data will estimate the effects of *CASL* after two years. Although the teachers in the study are expected to be relatively stable, some teacher mobility is inevitable. Teachers who enter participating schools in Year 2 of the study will be excluded from the analysis.

Effect Sizes. In addition to the above analyses, effect sizes will be calculated for all the outcome variables in the study, regardless of the direction of effect. Means and standard deviations for each group and any appropriate subgroups will be provided in final reports to allow for the calculation of effect sizes and effect directions by the reader.

Reporting. The following sections provide an overview of the reporting plan for this study.

Study Report Preparation. Researchers will prepare a technical report that is consistent with IES technical standards. The report will fully explain the rationale, study questions, research design, method, and results. Findings for each of the analysis questions will be presented and interpreted in light of the fidelity of implementation. In addition to standardized effect sizes, results will convey the educational importance of an intervention in understandable, real-world terms, such as cost, percentiles, and percent above cut-points. Threats to validity will be considered and ruled out as appropriate, and conclusions about the research questions will be

drawn based on these considerations. The report will be prepared such that it is appropriate for a peer-reviewed scholarly journal.

Dissemination. The final report will be posted on the REL website for access by researchers and practitioners. The researchers will submit a version of the technical report for publication in a scholarly journal and for presentation at one or more research conferences appropriate to the topic. A non-technical report also will be prepared that discusses the study rationale, presents the research questions, summarizes the findings, and highlights selected conclusions. There will be no risk of deductive disclosure in the reports as all findings will be presented in aggregate form.

Public- or Restricted-Use Data Files. The researchers will begin preparing a data file and associated codebook prior to the first data collection. They will seek advice from the TWG regarding the most appropriate format for these materials. The data file and codebook will be updated throughout the study. Before release the data file will be reviewed for the possibility of deductive disclosure and edited to reduce this risk. Upon completion of data collection, the data file and documentation in the codebook will be finalized as public- or restricted-use data files.

17. OMB EXPIRATION DATE

Not applicable. We are not seeking this and plan to display the expiration of OMB approval on data collection forms.

18. EXCEPTIONS TO CERTIFICATION STATEMENT

No exceptions to the certification statement are requested or required.

REFERENCES

- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.
- Colorado Department of Education. (2004). Fall 2004 Average Salaries by Setting. Retrieved September 7, 2006, from <http://www.cde.state.co.us/cdereval/rv2004AveSalbySetting.htm>
- Colorado Department of Education. (2005). Detail Job Classification Descriptions. Retrieved September 12, 2006, from <https://ade.cde.state.co.us/jobdefs.htm>
- Colorado Department of Education. (no date). Colorado Student Assessment Program: A Guide for Parents. Retrieved March 14, 2007, from http://www.cde.state.co.us/cdeassess/documents/parents/CSAP_Eng.pdf
- Dillman, D. A. (2000). *Mail and Internet Surveys: The Tailored Design Method* (Second ed.). New York: John Wiley & Sons, Inc.
- DuFour, R. (2005). What is a Professional Learning Community? In R. DuFour, R. Eaker & R. DuFour (Eds.), *On Common Ground: The Power of Professional Learning Communities*. Bloomington, IN: National Educational Service.
- Groves, R. M., Cialdini, R. B., & Couper, M. P. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly*, 56, 475-495.
- IRRE. (1998). Research Assessment Package for Schools (RAPS) Manual. Retrieved March 21, 2006, from http://www.irre.org/publications/pdfs/RAPS_manual_entire_1998.pdf
- James, J. M., & Bolstein, R. (1990). The effect of monetary incentives and follow-up mailings on the response rate and response quality in mail surveys. *Public Opinion Quarterly*, 54, 346-361.
- James, J. M., & Bolstein, R. (1992). Large monetary incentives and their effect on mail survey response rates. *Public Opinion Quarterly*, 56, 442-453.
- King, K. A., & Vaughan, J. L. (2004). Influence of paper color and a monetary incentive on response rate. *Psychological Reports*, 95, 432-434.
- Midgley, C., Maehr, M. L., Huda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., et al. (2000). *Manual for the Patterns of Adaptive Learning Scales (PALS)*. Ann Arbor, MI: University of Michigan.
- Murray, D. M., Hannan, P. J., & Baker, W. L. (1996). A Monte Carlo study of alternative responses to intraclass correlation in community trials: Is it ever possible to avoid Cornfield's penalties? *Evaluation Review*, 20(3), 313-337.
- Office of Management and Budget. (2006). *Standards and guidelines for statistical surveys*. Washington, D.C.: Author.
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, 66(4), 543-578.

- Raudenbush, S., Spybrook, J., Liu, X., & Congdon, R. (2006). Optimal design for longitudinal and multilevel research: Documentation for the "Optimal Design" software. Retrieved June 1, 2006, from <http://sitemaker.umich.edu/group-based/files/odmanual-20060517-v156.pdf>
- Ryan, R. M., & Deci, E. L. (2000). Self-Determination Theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68-78.
- Schmoker, M. (1999). *Results: The key to continuous school improvement* (2nd ed.). Alexandria, VA: Association of Supervision and Curriculum Development.
- Schmoker, M. (2005). No turning back: The ironclad case for professional learning communities. In R. DuFour, R. Eaker & R. DuFour (Eds.), *On Common Ground: The Power of Professional Learning Communities*. Bloomington, IN: National Educational Service.
- Schmoker, M. (2006). *Results Now: How we can achieve unprecedented improvement in teaching and learning*. Alexandria, VA: Association of Supervision and Curriculum Development.
- Shettle, C., & Mooney, G. (1999). Monetary incentives in U.S. Government surveys. *Journal of Official Statistics*, 15, 231-250.
- Stiggins, R. J., Arter, J. A., Chappuis, J., & Chappuis, S. (2004). *Classroom assessment for student learning: Doing it right - using it well*. Portland, OR: Assessment Training Institute.
- Teisl, M. F., Roe, B., & Vayda, M. (2006). Incentive effects on response rates, data quality, and survey administration costs. *International Journal of Public Opinion Research*, 18, 364-373.
- What Works Clearinghouse. (2006). WWC Study Review Standards. Retrieved November 15, 2006, from http://whatworks.ed.gov/reviewprocess/study_standards_final.pdf
- Whiteman, M. K., Langenberg, P., Jerulff, K., McCarter, R., & Flaws, J. (2003). A randomized trial of incentives to improve response rates to a mailed women's health questionnaire. *Journal of Women's Health*, 12, 821-828.
- Wiseman, F., Schafer, A., & Schafer, R. (1983). An experimental test of the effects of a monetary incentive on cooperation rates and data collection costs in central-location interviewing. *Journal of Marketing Research*, 20, 439-442.
- York-Barr, J., Sommers, W. A., Ghere, G. S., & Montie, J. (2001). *Reflective practice to improve schools: An action guide for educators*. Thousand Oaks, CA: Corwin Press, Inc.