Contract No.: ED-06-CO-0029
Lab Internal Ref No.: 02-01

Effect of Connected Mathematics 2 (CM2) on the Math Achievement of Middle School Students in Selected Schools in the Mid-Atlantic Region: A Cluster Randomized Controlled Trial

Supporting Statement A for Request for OMB Approval of

Invitation Letter, Interest Letter, Memorandum of Understanding, Teacher Consent Form, Teacher Demographics Survey, Monthly Online Teacher Survey, Parental Consent Form, Child Assent Form, Student Interest in Math Inventory, TerraNova 2nd Edition (CAT),

August 8, 2007

Taylor Martin
Kelli Millwood
Richard Brown
Ed Coughlin

Submitted to:
Institute of Education Sciences
555 New Jersey Ave., NW Room 504E
Washington, DC 20208-5500

Submitted by:
Regional Educational Laboratory-Mid-Atlantic
Penn State University
108 Rackley Building
University Park, PA 16802

Project Officer:
Dr. Ok-Choon Park

Project Director:
Kyle Peck

TABLE OF CONTENTS

LIST OF EXHIBITS

INTRODUCTION

This document presents the Supporting Statement requesting approval for a plan to collect data to assess the impact of a program of mathematics instruction. The project is sponsored by the Institute of Education Sciences within the U.S. Department of Education. Specifically, we are requesting approval to collect data including letters of interest and consent forms, pre- and posttests of mathematical achievement, a survey of student engagement, and teacher reports of classroom activities.

The advent of the *No Child Left Behind Act* (NCLB) of 2001 (P.L. No. 107-110) made clear the need to align standards, curriculum, instruction, and assessment goals with proven instructional practices in mathematics that use developmentally appropriate teaching methodologies that address the needs of all subgroups. In the critical area of improving achievement in middle school mathematics, a program called *Connected Mathematics* 2 (CM2) was designed to combine National Council of Teachers of Mathematics (NCTM) curricula with authentic, relevant problem solving practices to help students build math skills in a useful multidisciplinary context. CM2 has been researched previously, but few of the studies met the standards of the What Works Clearinghouse (WWC). The primary reason was that they did not employ experimental designs that randomly assign students or groups of students (such as classrooms or schools) to an intervention group participating in CM2 and a control group not participating in the curriculum program. This study aims to address this problem and provide a measurement of the impact of CM2 in keeping with NCLB's goals of making educational decisions based on rigorous methods. The Regional Educational Laboratory: Mid-Atlantic (REL-MA) is planning a cluster randomized trial of the *Connected Mathematics 2* Program on the mathematics achievement of sixth graders in the Mid-Atlantic region.

This document provides an overview of the planned data collection. The forms and data collection instruments to be used are described. These include parental consent forms, letters of interest and agreement, and assessment instruments to address the research questions. The submission describes the planned data collection in detail and includes an estimate of respondent burden associated with these efforts.

A. JUSTIFICATION

1. Circumstances That Make Data Collection Necessary

Statement of Need

Connected Mathematics and the revised *Connected Mathematics 2* are used widely across the country. In recent reviews by the Association for the Advancement of Science (AAS) and the US Department of Education's Mathematics and Science Expert Review Panel, *Connected Mathematics* received the highest rating for middle school mathematics curricula. However, there have been no studies of this program to date that meet the standards of evidence of the Institute for Educational Science's What Works Clearinghouse (WWC). Studies that meet these standards are clearly needed.

The current evidence base on *Connected Mathematics 1* shows that on balance the curriculum is associated with positive and moderately sized effects that may have failed to reach significance because of a lack of power due to relatively small sample sizes. The inability to make a generalized causal inference about the effects of *Connected Mathematics 1* based on the WWC review of *Connected Mathematics 1* stems, in part, from the lack of adequately powered and well implemented CRTs on the program. The proposed study will address this and other methodological limitations of previous studies to determine whether the positive and educationally substantive effects are both internally valid and statistically significant.

Rationale

Since the National Commission on Excellence in Education released the report *A Nation At Risk* (1983), there has been increased focus on improving mathematics education in the United States. The primary concern is U.S. students' level of performance in mathematics. The Third International Mathematics and Science Studies (TIMSS) and the National Assessment of Educational Progress (NAEP) results show that U.S. students in general do not perform as well as students in other industrialized nations in mathematics and science. NCLB legislation has focused attention on mathematics achievement even more, leading to reassessments of curricula and standards at the state level across the country.

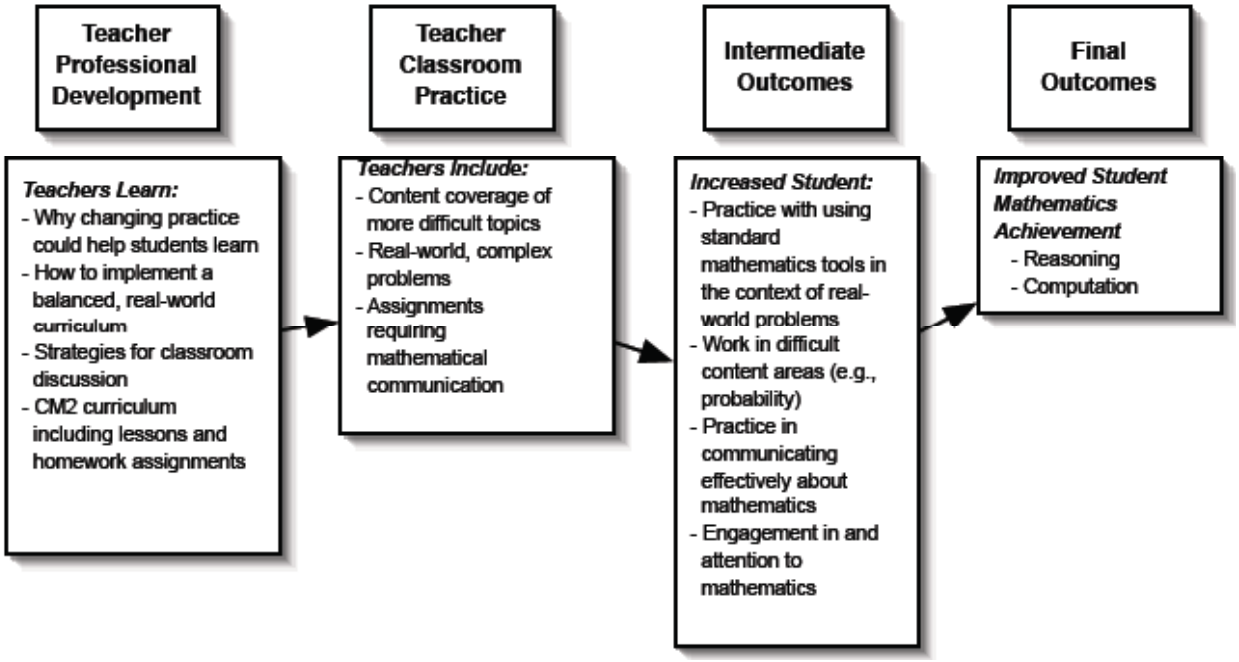
Many novel mathematics curricula have been developed over the past decade in order to provide guidance for improved mathematics instruction. Particularly, these curricula are more in keeping with the National Council of Teachers of Mathematics' (NCTM) *Principles and Standards for School Mathematics* than previously existing curricula (NCTM, 2000). The NCTM document identifies many characteristics of mathematics instruction that aim to increase student achievement in mathematics.

One such curriculum is *Connected Mathematics*. According to the developers and publisher, the program was designed to provide more balanced instruction that involves students in using mathematical reasoning in real-world contexts as defined by the NCTM Standards. In addition, a revision to the curriculum completed in 2005 was made to adapt the program so that it would better meet the needs of all students. This study will evaluate the impact of the revised *Connected Mathematics 2* (CM2) program. CM2 has the potential to improve mathematics achievement. Specifically, by bringing more balanced, real – world instructional practices into schools, it may improve students' engagement and performance in mathematics.

We developed Figure 1 to illustrate our theory of action by which CM2 could affect mathematics achievement. First, teacher professional development can provide educational resources and strategies in the form of the specific curriculum and training on practices of implementing it. In turn, this training and follow-up training should impact teachers' classroom practices as they implement CM2. This change should involve greater use of real-world problems, and balanced instruction that approaches all the NCTM goals, including

work on multiple topics and in different modes – both conceptual and procedural. These changes in classroom practice will lead to students engaging in different sorts of activities that should give them practice in a more representative set of mathematical topic areas and with the balance of mathematical reasoning tools, from basic memorized facts to extended reasoning strategies to solve novel problems. In addition, these practices should lead to higher student engagement in mathematics. Finally, these changes in practice and difference in student engagement should increase student achievement, particularly in contexts involving complex mathematical reasoning.

FIGURE 1. THEORY OF ACTION



Overview of the Study Design

REL-MA will recruit approximately 70 schools from the Mid-Atlantic Region (New Jersey, Delaware, Pennsylvania, Maryland and District of Columbia). Schools will be randomly assigned to the intervention or the control group. Connected Mathematics 2 will be implemented as the primary curriculum at the sixth grade level in the academic years 2008-09 and 2009-10 in the intervention schools (N = 35). The control schools (N = 35) will employ traditional curricula. The intervention and control groups will be compared on student achievement and on student engagement.

Overview of the Data Collection Plan

Table 1 provides an overview of the data collection plan, details of instruments used will be described later. During the recruitment year, principals of potential schools will complete letters of interest to allow the study team to assess the school's eligibility for participation. In Year 1 (2008) of the study, only teacher demographic and implementation fidelity data (site visits and online teacher surveys) will be collected as the teachers gain skill and confidence in implementing the curriculum. At the start of Year 2 (2009), the student engagement and achievement baseline measures will be administered (Eccles-Wigfield, 1995). During this year, the study team will continue site visits and the teachers will continue to complete online surveys. At the end of Year 2, students will complete the outcome achievement and engagement measures.

TABLE 1

DATA COLLECTION PLAN

Respondent	Mode	Timeline	Key Data
Principals	Letter of Interest	Fall-Spring 2007-2008	Eligibility for School's Participation
Teachers	Demographic Survey	Fall 2008	Teacher Characteristics
None	Retrieved from the Common Core of Data	Summer 2008	School Characteristics Student Characteristics (at school level)
Teachers	Site Visits	Fall-Spring 2008-2009	Implementation Fidelity
Teachers	Online Teacher Surveys	Fall-Spring 2008-2009	Implementation Fidelity
Students	Terra Nova Baseline	Fall 2009	Student Achievement Data
Students	Student Engagement	Fall 2009	Student Engagement Data
Teachers	Site Visits	Fall-Spring 2009-2010	Implementation Fidelity
Teachers	Online Teacher Surveys	Fall-Spring 2009-2010	Implementation Fidelity
Students	Terra Nova Outcome	Spring 2010	Student Achievement Data
Students	Student Engagement	Spring 2010	Student Engagement Data

2. How, by Whom, and for What Purpose the Information Is To Be Used?

The authors of this study propose using a multi-site cluster randomized trial (or multi-site CRT). The data will be collected to answer the following key research questions addressed by the study:

1. Does middle school students' use of CM2 as a comprehensive math curriculum cause higher student math achievement compared to students who use other curricula?

2. Does middle school students' use of CM2 cause higher levels of engagement in doing mathematics compared to students who use a traditional curriculum?

We expect the study to generate multiple kinds of information from consenting/assenting participants including: a Teacher Demographic Survey, Monthly Online Teacher Survey, Site Observations, Student Engagement Survey, and Terra Nova Student Achievement Test. Below we discuss how the different data will be used, by whom, and for what purposes.

Description of Data Collection

Data collection efforts and analyses will be led by Co-PIs with support from Analytica and a statistical analysis expert from University of Southern California.

The data collection plan for this research starts with schools agreement to participate and is followed by teacher consent, teacher demographics surveys, professional development, notifying parents of the waiver of consent, student assent, student pre-tests, one full year of data collection, and student post-tests. The respondents, mode of data collection, timeline, and key data to be collected at each stage are presented in Table 2. Table 2 is very similar to Table 1, but it focuses more on data collection. Details about each instrument follow the table.

TABLE 2

DATA COLLECTION PLAN

Respondent	Mode	Timeline	Key Data
Schools	Agreement to Participate in Study	September 2007 – April 2008	Letter stating that the school will implement Connected Mathematics 2 for all 6th grade math classrooms during the 2008-2009 and 2009-2010 school years.
Teachers	Teacher informed consent	June 2008	Teacher agreement to participate in the study.
Teachers	Teacher Demographics Survey	June 2008	Demographic information and math teaching experience.
Teachers	Fidelity Observations of Intervention Teachers	September 2008 – May 2009 (3 times through the year)	Descriptive information to support the statistical analysis of data in the study. Description of intervention classroom's instructional objectives, strategies, and use of Connected Mathematics 2.
Parents	Waiver of Consent Form (IF they choose NOT to participate)	September 2009	Agreement to participate.
Students	Informed Assent Form	September 2009	Agreement to participate in the study.
Students	Pre-test	September 2009	Pre-test math assessment and engagement survey
Teachers	Monthly Online Intervention Teacher Survey	September 2008 – May 2010	Questionnaire to monitor fidelity of implementation via teacher feedback on monthly progress
Teachers	Fidelity Observations of ***all*** Teachers	September 2009 – May 2010 (3 times through the year)	Descriptive information to support the statistical analysis of data in the study. Description of classroom's instructional objectives, strategies, and in intervention classrooms use of Connected Mathematics 2.
Students	Post-test	April – May 2008	Post-test math and engagement assessment to assess change through the year.

a. Schools Agreement to Participate

Analytica, a partner in REL:MA will create a list of possible schools for recruiting from the Mid-Atlantic region. Letters will be sent to schools on the list inviting them to participate in the study (See Exhibit A). The superintendent or designee will sign a letter of interest and provide basic information about their schools (See Exhibit B). Schools selected for the sample will sign memorandum of agreement, where they agree to the terms and conditions of the study (See Exhibit C).

b. Teacher Informed Consent

Teachers participating in the study will sign an informed consent form that provides a written overview of the project detailing procedures if in the intervention or control condition (See Exhibit D).

c. Teacher Demographic Survey

After teachers sign the consent form they will complete a short teacher demographic survey (See Exhibit E). This survey will be used to collect baseline teacher characteristics that will be instrumental in determining balanced intervention and control groups.

d. Monthly On-line Intervention Teacher Survey

At the end of each month teachers in the intervention group will complete a brief online survey about their experiences with CM2 (See Exhibit F). The results from this survey will be used as a measure of implementation fidelity.

e. Parent Waiver of Consent Form

The parents of sixth grade students of participating schools will be sent a letter with a waiver of consent form to be returned if they do not consent to their child's participation in the study (See Exhibit G). This method will be used pending approval from each School District (local IRB or school legal counsel) and The Pennsylvania State University IRB. Generally speaking, waiver of consent can be used with USDOE studies that use academic interventions and outcomes.

f. Child Assent Form

Teachers will give students, who have parental/guardian permission to participate, a description of the study and explain it to the students. At that time the teachers will ask the students if they are willing to participate in the study, if students' would like to participate they will sign at the end of the assent form (See Exhibit H). Teachers will also explicitly advise students that they can withdraw their participation at any time during the project; this is also stated in the assent form. Child assent is required for the Institutional Review Board at The Pennsylvania State University.

g. Site Observations

Implementation fidelity will be assessed through protocols developed according to the CM2 materials and professional development to monitor implementation through three site visits in each intervention school for each of the two study years. These protocols are being developed in collaboration with the Principal Investigators of the CM2 study and the curriculum developer. Once finalized, they will be submitted to OMB as addenda. CM2 schools and all sixth grade classrooms contained within them will be observed three times during the "formative" year of teacher preparation (Year 1 of the study) and three times during the following "impact" year (Year 2) of teacher curriculum implementation in which impact data will be collected.

During Year 2 of the study, control classrooms will also be visited three times as well, simply to document any differences in practice as well as to confirm the curriculum and materials in use. In addition to teacher practice, the protocol will enable classroom observers to document any behavioral or psychological impacts that might be present in the control classrooms as a result of the study and document these impacts to inform interpretation of the study results. Issues of crossover (students in the control classrooms participating in instruction delivered in intervention classrooms or vice versa), spillover (teacher instructional practices used in the intervention classrooms adopted by teachers in control classrooms), and contamination effects (students in the control classrooms adopting the habits of mind imbued by instructional activities in the intervention classrooms), should not be an issue since intervention and control classrooms are not in the same school, but with the wide distribution of CM2 throughout the region other pathways for the materials or strategies to reach the control classrooms are conceivable.

h. Student Engagement Survey

A measure of student engagement, Student Math Interest Inventory, will be used at the beginning and end of Year 2 (see Exhibit I). This measure, developed by Eccles and Wigfield (1995) in an earlier study, of student motivation and attitude towards mathematics has been validated in several previous studies; evidence of construct validity for the instrument via the Confirmatory Factor Analysis procedure is provided. The authors report three scales that measure engagement in mathematics including: ability/expectancy perceptions, task difficulty perceptions, and task value perceptions. Reliability for the instrument is strong for all of the scales; alpha coefficients range from .62 to .95, and the goodness of fit measures all exceed .97.

This measure will be administered to students, in both conditions, only in Year 2 of the study– in Fall 2009 and Spring 2010. This is a brief survey that takes approximately 15 minutes to administer. The Student Interest Inventory will be the only measure of student engagement and will be used as an outcome measure to estimate the impact of CM2 on student engagement.

i. Terra Nova Student Achievement Test

Terra Nova 2nd Ed. (CAT) for the sixth grade level will be used at the beginning and end of Year 2 as a measure of student achievement. The Terra Nova 2nd Ed. (CAT) is an appropriate measure for this study for multiple reasons. First, the content objectives reflect the National Council of Teachers of Mathematics (NCTM) Standards, state and local curriculum documents, and the conceptual framework of NAEP. The NCTM's vision of having students reason mathematically and solve real life problems is a major focus. Second, it employs both selected and constructed response items to gain a deep understanding of what students know and can do. Third, the Terra Nova, 2nd Ed. (CAT) has desirable psychometric properties. It uses IRT to combine the selected and constructed response items on a single scale, improving interpretation across these different types of items. In addition, the national norming process for the TerraNova 2nd Ed. (CAT) was based on nationally representative student samples (more than 275,000 students in grades K-12) to obtain high reliability and validity.

This measure will be administered to students, in both conditions, only in Year 2 of the study– in Fall 2009 and Spring 2010. The full assessment will require 90 minutes to administer. The burden of completing this measure is not included in the collection since this is an assessment of the student's abilities. There is no copy of the assessment included in this package, as this is a copyrighted measure. However, a copy can be furnished to OMB upon request (with permission of the copyrighter).

All scoring of the assessment will be contracted with the publisher. Terra Nova 2nd Ed. (CAT) will be the only measure of student achievement and will be used as an outcome measure to estimate the impact of CM2 on student achievement

3. Use of Information Technology

The data collection plan was designed to lead to efficient and accurate collection of data while minimizing respondent burden. School characteristics (such as % rural) and student characteristics (such as % migrant students, measured at the school level) will be gleaned from the Common Core of Data. In addition, intervention teachers will complete online surveys. Other than these data, no data will be gathered electronically.

4. Efforts to Identify and Avoid Duplication

In order to achieve REL-MA's research goals most effectively, it will be necessary to collect primary student achievement and engagement data for several reasons. REL-MA will take the following steps to identify and avoid duplication:

1. *Students spread across states*. The data will be collected from schools spread across the Mid-Atlantic region. Each of these states administers a separate state achievement test. While there are many similarities for any given grade level on these assessments, they are different enough that basing the impact estimate of *Connected Mathematics 2* (CM2) on student performance on these tests would lead to difficulty in interpreting the results. The achievement tests that REL-MA administers will therefore not duplicate tests students currently take.
2. *Validity of data*. The tests will be administered by proctors not associated with the school. This has two positive advantages over teacher-administered achievement tests. One, there is less chance that data will be lost. The tests will be administered at each school in one administration and the booklets collected and sent back to the vendor immediately. Two, the proctors will be blind to the condition of the participants taking the test eliminating any bias that could occur due to any action taken with knowledge of condition. The difference in administrators makes our data gathering effort unique.
3. *Timing of tests*. It is important to administer the pretest at the start of the school year rather than relying on student test scores from the prior year for two reasons. One, some students may have attended summer school or engaged in other activities that could affect their performance. This could lead to an incorrect assessment of initial level of mathematics achievement. If so, it would affect the estimate of the impact of CM2. Two, it is important to compare student performance at both the start and end of the year on the same test. While some end-of-year achievement tests are calibrated to be comparable, others are not. A valid assessment of the impact of CM2 requires an adequate measure of progress on the items related to the sixth grade content goals. Since there is not an existing pretest, this assessment will not duplicate any other effort.
4. *Types of items assessed*. In addition to the practical reasons mentioned above, the Terra Nova Multiple Assessment Test provides open-response as well as closed-response items. These items are a key demonstration of students' ability to reason mathematically, an important measure of interest for the study. These items are not included on many states' achievement tests. The open-response items are not part of assessments students typically complete, therefore these measures do not duplicate existing measures.
5. *The engagement measure*. This measure is not a pre-existing source. However, understanding student engagement in the curriculum is key to the theory of action presented earlier and therefore to understanding the changes in student performance that may occur during the study. As this measure is not given normally, it does not duplicate other efforts.

5. Impacts on Small Businesses and Other Small Entities

There is no impact on small business from this study.

6. Consequences to Federal Programs or Policies if Data Collection is Not Conducted

If the data collection efforts described were not conducted, it would significantly hamper the Department of Education's (DOE) assessment of the impact of CM2 on student engagement and achievement. As described earlier, the No Child Left Behind Act requires that education decision makers base policies and programs on scientifically based research. The DOE funds several research efforts that address the efficacy of mathematics

curricula. Information on the impact of CM2 can inform the decisions on funding these efforts by contributing to a better understanding of the characteristics of effective mathematics curricula.

More generally, the recent growth in the number of schools and districts employing at least some aspects of NCTM Standards-based curricula is a major part of efforts to improve mathematics education flowing from NCLB legislation. Understanding the effects of curricula like CM2 will provide more evidence for DOE in developing evidence-based approaches to mathematics instruction and informing parents and schools about those approaches.

7. Special Circumstances

None of the issues listed as inconsistent with usual policy for this section are relevant in the current study.

8. Solicitation of Public Comments and Consultation with People Outside the Lab

a. Federal Register Announcement

A 60-day notice to solicit public comments was published in the Federal Register on February 1, 2007 (vol. 72, p. 4695). We are still within the 60 day comment period. We will address all comments received.

b. Consultations Outside the Agency

Study staff have consulted with a number of people with expertise in cluster randomized trials, multi-level analyses and mathematics instruction. These include members of our Technical Working Group (experts gathered to guide rigorous designs, as stipulated in all IES lab contracts) and Dr. Richard Brown, study CRT Advisor (see Table 3).

TABLE 3

CONSULTANTS

Expert	***Affiliation***
Dr. Robert Boruch	University of Pennsylvania
Dr. Rebecca Maynard	University of Pennsylvania
Dr. Richard Brown	University of Southern California*
Dr. Herbert Turner	Analytica
Dr. Mike Lopez	National Center for Latino Child and Family Research
Willa Spicer	NJPSA
Scott Joftus	Cross & Joftus, LLC

*Also a study team member

9. Justification for Respondent Payments

REL-MA will not provide monetary incentives to schools, teacher, parents, or students. All participating teachers will receive a $25 non-monetary incentive per year for participating in the two years of the study. Schools selected for the intervention condition in this study will have the opportunity to participate in the high-quality professional development that comes with CM2 and will have access to the CM2 curriculum and materials. As they would for the school's standard professional development activities, teachers will receive the base pay rate, for outside of contract time, per day (approximately $18-30 per hour, depending on location and

experience) as a stipend for participating in this professional development. Schools randomly assigned to the control condition will receive $1000 worth of equipment for their school (i.e., laptop, LCD projector, etc.).

10. Confidentiality Assurances

The Education Sciences Reform Act of 2002, Title I, Part E, Section 183 requires "All collection, maintenance, use, and wide dissemination of data by the Institute" to "conform with the requirements of section 552 of title 5, United States Code, the confidentiality standards of subsection (c) of this section, and sections 444 and 445 of the General Education Provision Act (20 U.S.C. 1232g, 1232h)." These citations refer to the Privacy Act, the Family Educational Rights and Privacy Act, and the Protection of Pupil Rights Amendment.

The Director shall ensure that all individually identifiable information about students, their academic achievements, their families, and information with respect to individual schools, shall remain confidential in accordance with section 552a of title 5, United States Code, the confidentiality standards of subsection (c) of this section, and sections 444 and 445 of the General Education Provision Act.

Subsection (c) of section 183 referenced above requires the Director of IES to "develop and enforce standards designed to protect the confidentiality of persons in the collection, reporting, and publication of data".

Subsection (d) of section 183 prohibits disclosure of individually identifiable information as well as making any of the publishing or communicating of individually identifiable information by employees or staff a felony.

REL-MA 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. REL-MA 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. All members of the study team having access to the institution-level data have been certified by The Pennsylvania State University Institutional Review Board (IRB) as having received training in the importance of confidentiality and data security. All institution-level identifiable information will be kept in secured locations and identifiers will be destroyed as soon as they are no longer required

Data collection activities will also be conducted in compliance with The Privacy Act of 1974, P. L. 93-579, 5 USC 552 a; the "Buckley Amendment, " Family Educational and Privacy Act of 1974, 20 USC 1232 g; The Freedom of Information Act, 5 USC 522; and related regulations, including but not limited to: 41 CFR Part 1-1 and 45 CFR Part 5b and, as appropriate, the Federal common rule or ED's final regulations on the protection of human research participants.

REL-MA has worked with the Institutional Review Board at The Pennsylvania State University to seek and receive approval on the study and all its controls. All investigators on the *Connected Mathematics* Study have a long history of protecting confidentiality and privacy of records, and consider such practice a critical aspect of the scientific and legal integrity of any data collection effort. We plan to use ongoing, long-standing techniques that have proven effective in the past. Every member of the study team must be certified in conducting research with human subjects by the appropriate Institutional Review Board. The forms shown in Exhibits D, G, and H show the teacher informed consent, parent waiver of consent, and child assent. It will be very important that parents or legal guardians of sample members understand that information is being collected regarding their children, and that this information is being held confidential. We will use the following process to inform

parents/guardians of data collection and confidentiality procedures and to obtain their consent. As part of participating in the study, schools will ask parents to provide their waiver of consent for study activities at the start of the school year. Parents will be informed at this time that the study is voluntary. We will then use passive consent procedures and inform all parents whose children enter the sample of the study and give them the chance to withdraw, as well as informing them that they can withdraw their child at any time. The parental consent form is shown in Exhibit G, child assent form is show in Exhibit H, and the teacher consent form is shown in Exhibit D.

In addition, the CM2 Study Partners will employ the following safeguards to carry out confidentiality assurances (based on the PSU IRB application):

1. All students will use usernames and passwords and they will be identified by an assigned ID number. Only the co-PIs, the study manager, and Dr. Peck (Director of the Lab) will have access to the information linking the students with their ID number. Pre- and post-test data will be stripped of names once ID numbers have been affixed, and at no time will the results for individuals, teachers, classes, schools, or districts be reported.
2. The only data that will be reported will be aggregated data describing the treatment and control conditions, and the pre-test performance of students classified as "high," "average," and "low." Requests for any other information will be denied.
3. All data and forms collected from the students will be stored in a secured file cabinet at the Penn State Lab Offices.
4. All identifying information will be replaced with the ID numbers when scores and demographic data is entered into the statistical analys programs and HLM files for data analysis.
5. 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 individuals, schools, or districts to anyone outside the study team, except as required by law.
6. All copies of the informed consent forms will be maintained in the Penn State Lab Offices in a locked cabinet with a signed copy returned to the participating teachers.
7. Similarly, any waiver of consent form returned will be carefully noted to remove the student's information from the study and data analysis. This will be conducted by the PI and the forms will be maintained in the locked cabinet.
8. Access to sample selection data is limited to those who have direct responsibility for selecting the sample. At the conclusion of the research, these data will be destroyed.
9. Identifying information on schools, students, and parents is maintained on separate forms, which are linked to the interviews only by a sample identification number. These forms are separated from the interviews as soon as possible.
10. Access to the hard copy documents collected from respondents is strictly limited. Documents are stored in locked files and cabinets. Discarded material is shredded.
11. Computer data files are protected with passwords and access is limited to specific users. With especially sensitive data, the data are maintained on removable storage devices that are kept physically secure when not in use.

11. Additional Justification for Sensitive Questions

The data collection instruments largely cover non-controversial subject matter in standard ways. For example, they do not cover topics often considered sensitive such as risk-taking behavior, sexual or political topics, criminal activities, etc.

12. Estimate of Total Annual Hour Burden

Participation in all data collection activities is completely voluntary, with no sanctions or penalties being applied for respondents who choose not to provide information or who do not answer specific questions. Table 4 presents the estimated respondent burden for the various data sources. As summarized in Table 2, the estimated respondent burden for data collection is 33,025 respondents and 19, 223 hours (Note that these estimates include administration of all pre and post assessments of students). The annual average burden hours will be 6, 408 (The Student TerraNova 6th grade Achievement test is an assessment. Therefore the burden hours are exempt from PRA.)

TABLE 4

ESTIMATE OF TOTAL ANNUAL HOUR BURDEN[a]

Instrument or Data Source	Average Number of Respondents	Number of Responses per Respondent (total number of responses in parentheses)	Average Time Per Response	Total Burden (Hours)
Recruitment Year				
School/District letter of interest	300	1 (300)	20 minutes	100
School/District memorandum of understanding	70	1 (70)	10 minutes	12
Year 1				
Teacher Consent Form	210	1 (210)	5 minutes	18
Teacher Training (CM2, 5 days in summer)	210	1 (210)	40 hours	8400
Teacher Demographic Survey	210	1 (210)	15 minutes	53
Site Visits	210	3 (630)	1 hour	630
Year 2				
Parental consent form	10,500	1 (10,500)	10 minutes	1750
Child Assent Form	10, 500	1 (10,500)	10 minutes	1750
Student Attitude Measure (Eccles and Wigfield)	10, 500	2 (21, 000)	15 minutes per administration (at pre- and posttest)	5250
Site Visits for implementation fidelity for 2nd year of using CM2	210	3 (630)	1 hour	630
CM2 Teacher Monthly Progress Reports	105	18 (1, 890)	20 minutes	630
TOTAL	**33,025**	**33 (46, 150)**		**19,223**
ANNUAL AVERAGE	**11,008**	**11 (15, 383)**		**6,407**

[a] *Estimates based on an assumption of 70 schools with 6 classrooms per school and 25 students per class*

10, 500 students and 210 teachers

Table 5 presents our estimate of respondent cost burden. The data collection for the study includes schools' letters of agreement to participate, teacher informed consent form and parent waiver of consent form. Where needed, we will also use school information request forms if we cannot get adequate descriptive data on-line. Students will complete a pre and post-test of the Terra Nova 2nd ed. (CAT), Basic Multiple Assessment, at the sixth grade level test. During the formative and impact year for intervention teachers, and for the study year only for control teachers, REL-MA will conduct three classroom observations at the beginning, middle, and end of the school year to monitor implementation fidelity and to provide descriptive information for the data analysis process. The classroom observation protocol, under development for Pearson at Chapman University, will be released Fall 2007. This observation protocol will not require teachers to submit any written response related to each visit. Any questions that the reviewer might have outside the observational items in the protocol would be asked of the teachers in brief asides during the observation. The protocol will be copyrighted by Pearson and cannot be made publicly available. In addition, we will collect baseline year administrative records for all students in the sample using data that is readily available the State Departments of Education in each state where the study is planned. If any additional information is necessary the school or district staff will assist the team.

Researchers in the project will not be gathering any written responses from teacher before, during, or at the conclusion of the teacher training. Connected Mathematics does ask for teacher feedback in the course of the training to guide their delivery, but none of these data will be used within the study. Two questions will be included in the demographic survey regarding teachers' perceived readiness to teach the mathematics curriculum in the coming year, but this survey has already been accounted for in the calculated burden.

Our assumptions for the study participants include 70 schools, 210 teachers, and 10,500 students (we assume, on average, each teacher will teach 2 classes.) Additional details about respondent burden are provided in Exhibit H. In all, total respondent hours are 60.390. These hours include 1750 hours estimated for the parents' waiver of consent forms. However, it must be noted that the forms need to be completed and returned ONLY in the instance where the parent does NOT allow their child to participate in the study. Therefore this estimate is actually much higher than our expectations of approximately 1% of participants actually completing the forms.

TABLE 5

RESPONDENT BURDEN ESTIMATES (BASED ON ACTIVITIES IN Table 4)

Informant	Number of Responses	Number of Response Activities	Average Time per Respondent (Hours)	Total Respondent Time (Hours)	Estimated Hourly Wage (Dollars)	Estimated Cost Burden to Respondents (Dollars)
Parents	10500	1	10 min.	1750	$14.95[1]	$26,152.60
Teachers (Intervention)	105	5	13 hours (Training time is reimbursed at district rates)	1365	14.95[2]	$20,406.75
Teachers (Control)	105	3	3 hours 20 minutes	350	14.95	$5,232
School District Staff	70	1	20 minutes/ school	23.3	$10.02[e]	$233.78

13. Estimate of Total Cost Burden to Respondents

There are no start-up costs to respondents. Burden hour costs to respondents are discussed above in Section A. 12. of this document.

14. Estimate of Total Costs to the Government

For the data collection activities for which OMB approval is currently being requested, the overall cost to the government is $4,232,367. This includes:

- $320,462 for study design, OMB clearance, and planning
- $345,767 for recruiting and random assignment of schools to the study
- $1,657,238 for the purchase of materials, training of teachers, and monitoring of first-year implementation
- $1,063,420 for pre-testing, additional training, and monitoring during the course of the study year
- $845,480 for final data collection, data analysis, report preparation and dissemination

Thus, the overall costs to the government of the full range of the cluster randomized control trial over the entire study period will be $4,232,367 during the study period (2007-11). This estimate is based on the evaluation contractor's previous experience managing other research and data collection activities of this type. The average annual cost over 3 years will be $1,410,789.

15. Program Changes or Adjustments

This is a program change of 6, 408 because this is a new collection.

[1] 2003 Statistical Abstract of the U.S. Table No. 636: Average Hourly Earnings by Private Industry Group: 1980-2002 (estimate in table is for 2002).

[2] 2003 Statistical Abstract of the U.S. Table No. 251: Average Salary and Wages Paid in Public School Systems: 1980-2002 (estimate in table is for 2002).

16. Tabulation, Analysis and Publication of Results

In this section, we describe the ways we will use the data collected. The data will primarily be used to describe the study sites and sample, describe the implementation of the curricula used, and estimate the effects of CM2 on student achievement and engagement outcomes.

Describing the Study Sites

The data collected will be used to characterize the schools in which the study was conducted. This information will assist policy makers and other stakeholders in interpreting the results of the study. In addition, it will be included in the data analysis as described in Section B. The data will be collected from the Common Core of Data.

Describing the Study Sample

The demographic information from the Common Core of Data, and the student achievement and engagement pretests will be used to characterize the participants in the intervention in the study. These data will be used in analysis and will be helpful for interpretation of the results by educational decision makers and others. In addition, it will inform the study team on the similarity between the intervention and control participants in the study. The teacher survey will provide background information on teachers' preparation and education. Again, this information is necessary for interpreting the results of the study and describing the level of similarity between the intervention and control teachers who were randomly assigned to condition.

Describing the Implementation of the Intervention and Control Curricula

Data from the site visits and teacher reports will be used to assess implementation fidelity. For the CM2 group, it is important to ensure the teachers are using the materials as the vendor recommends and the professional development explains. For the control group, it is important to ensure that methods similar to CM2 are *not* being used.

Estimating the Effects of the New Curricula on Academic Outcomes

Consistent with standards of reporting on randomized experiments as articulated in Boruch (1997) for the social sciences, in the Consolidated Standards on Reporting Trials (CONSORT) in biomedicine, and in Flay et al. (2005) in prevention research, our analysis of the data will begin with an assessment of the flow of participants and clusters through the trial. To this end, we will report the number of school units, classrooms, and students for each group (intervention and control) through the following stages of the trial:

- Assessment for eligibility
- Allocation
- Follow-up
- Analysis

The core curricula used by intervention and control classrooms is an important feature of this trial (or CRT). In the intervention classrooms, the CM2 curricula is the core curricula. In the analysis of the impacts of CM2, the impact is based on a contrast of the CM2 classrooms and "curriculum as usual" classrooms, and knowing what "curriculum as usual" means would be most useful. Therefore, we will present a profile of the curricula used in the CM2 and control classrooms.

Baseline Comparison of Groups – Participant Characteristics

For reasons explained earlier (random assignment equates on long-run expectations), reporting the baseline characteristics for schools and teachers, verifying equating groups through random assignment, and identifying any baseline characteristics on which the groups are not in balance, is most important. Characteristics for which groups are unbalanced, defined as a statistically significant difference between the intervention and control groups as determined by t-tests for baseline characteristics measured on a continuous scale and by chi-square goodness of fit tests for variables measured on a categorical scaled, will be used to statistically equate the groups by adding these variables as covariates in the multi-level model presented subsequently.

After random assignment of schools, we will internally review (and eventually report) the following baseline characteristics for CM2 and control schools. A t-test and chi-square goodness of fit test will be used to test for statistically significant differences between groups on baseline characteristics. Data collection sources for characteristics are in parenthesis:

School Characteristics: CM2 vs. Control (Source: Common Core of Data):
a. % Urban
b. % Suburban
c. % Rural
d. % Small city

Sixth Grade Teacher Characteristics (at the school level) within CM2 and Control Schools (Source: teacher demographic survey, Exhibit E):
a. Gender (% female)
b. Ethnicity (% African American, % Native American, etc.)
c. Mean age in years
d. Mean years in current school
e. Mean Years in current district
f. Mean years teaching experience
g. % with bachelor's degree in mathematics
h. Highest degree (% bachelors, % masters, % PhD)
i. Mean # of course hours in mathematics
j. Mean hours of professional development in mathematics in the past three years

Student Characteristics (at the school level) within CM2 and Control Schools (Source: Common Core of Data):
a. Mean score on the Terra Nova, selected and constructed response items
b. Gender (% female)
c. Ethnicity (% African American, % Native American, etc.)
d. % Limited English proficiency
e. % Eligible for free or reduced lunch
f. % Migrant students
g. % Special education

To conclude, if there are any statistically significant effect sizes on baseline characteristics, especially the pre-intervention measures of achievement (which will also serve as the post-intervention measure of achievement), then these measures will be included in the multi-level model used to assess the impact of the intervention on the outcome as covariates. This and other methods are discussed next. Given the anticipated size of the sample at the school and student level, we would expect the number of baseline characteristics for which there are imbalances to be very small.

Analytic Models

In this cluster randomized trial in which schools are the unit of random assignment, any analysis of the data must take into account the clustering (or lack of independence) of student outcomes within schools. The lack of independence of student outcomes within a school means that any one student's score in the school can be used to predict another student's score in that school. This dependence is quantified by the Intra-Cluster Correlation (ICC) that can be defined, technically, as the proportion of variance that is between schools, or, intuitively, as the correlation between the outcome values of any two individuals in the same school (Bauer and Curran, 2006). If there is dependence among students within a school, then the ICC will be greater than zero (ICC > 0) and a multi-level model must be used to correct the standard errors for the dependence.

Using a multi-level model can be justified given the structure of the data for this study in which students are nested within schools. Thus, we will use a multi-level model with students at Level 1 and schools at Level 2. In other words, we will estimate the impact of CM2 on student engagement and student achievement using a two-level model with what Schochet (2005) refers to as "clustering at the school level." We use a two-level model, rather than a three-level one (where classroom level clustering may also be taken into account), because all classrooms within schools randomly assigned to the CM2 condition will be required to implement CM2 during the two-year study (see Schochet, p. 21).

There is growing interest within the field of education about the magnitude of ICCs. We will estimate the ICCs for each of our study outcomes using the following model that will be implemented using PROC MIXED in SAS:

$$\text{ICC} = \frac{\tau_\pi}{\tau_\pi + \sigma^2}$$

where τ_π is the between school (or Level 2) variance of the outcome of interest, σ^2 is the within school variance of the outcome of interest, and $\tau_\pi + \sigma^2$ is the total variance of the outcome of interest.

Estimating Intervention Effects

Recall that the research questions for this study are:

1. Does middle school students' use of CM2 as a comprehensive math curriculum cause higher student math achievement compared to students who use other curricula?

2. Does middle school students' use of CM2 cause higher levels of engagement in doing mathematics compared to students who use other curricula?

To address both questions, we use a multi-level model (with students at Level 1 and schools at Level 2) to estimate the impact of CM2 on student engagement and student achievement. In the Level-1 model, student outcomes are modeled as a function of the students' TerraNova pretests scores aggregated to the school level and any pre-intervention characteristics for which there are statistically significant imbalances. The inclusion of the TerraNova pretest in the model is to improve statistical precision of parameter estimates whereas inclusion of the pre-intervention characteristics for which there are statistically significant imbalances is to reduce bias in parameter estimates. (Bloom, Hayes, & Black, 2005; Raudenbush, Martinez, & Spybrook, 2005). The Level-1 model is specified as follows:

$$Y_{ij} = \pi 0_j + e_{ij} \sim N(0,\sigma^2)$$

where,

Y_{ij} is the student outcome for student i in school j;
π_{0j} is the average outcome of students in school j;
e_{ij} is a random error associated with student i in school k.

The school average outcome estimated from Level-1 intercept π_{0j} in equation 0.2a will be modeled (in equation 0.2b below) as varying randomly across schools and as a function, at school-level, of CM2, student pretest (aggregated to the school level), and any school characteristics at baseline for which there are statistically significant imbalances. Schools are treated as fixed effects at Level 2, as shown in the following Level-2 specification:

$$\pi_{0j} = \beta_{00} + \beta_{01}*(CM2)_j + \beta_{02}*(Pretest)j + \sum_{r=1}^{n} \beta_{0r} * (BIC_r)_j + u_{0j}, \quad U_{0j} \sim N(0,\tau_\pi).$$

where

β_{00} is the average student outcome across all schools;
β_{01} is the difference in the average student outcome between the CM2 schools and the control schools, or the intervention effect;
CM2is an indicator variable for the CM2 schools: 0.5 = CM2, and -0.5 = control;
β_{02} is the average effect of student pretest on average student outcome;
Pretest is student pretest aggregated to the school level and grand-mean centered;
β_{0r} is the average rth effect of the rth baseline covariate for which there is a statistically significant imbalance;
BIC_r is the rth baseline covariate for which there is a statistically significant imbalance;
u_{0j} is a random error associated with school j on school average student outcome;

To address the first research question, we estimate the model with student achievement as the outcome (Y_{ij}). To address the second research question, we estimate the model with student engagement as the outcome (Y_{ij}). In addition to the statistical significance of the CM2 effect, we will also gauge the magnitude of the effect with the effect size index. Specifically, we will compute the effect size as a standardized mean difference (Hedges's g) by dividing the adjusted group mean difference (β_{01}) by the unadjusted pooled within-group

standard deviation of the outcome measure. It is possible that the intervention will affect not only the mean of CM2 schools but also the standard deviation. In this case, Glass' Delta (adjusted group mean difference divided by the control group standard deviation) would be used as the appropriate metric (Keppel and Wickens, 2004; Lipsey & Wilson, 2001).

Presentation of Results

Results estimated from the model will be presented in two ways. First, we will report the mean difference in outcomes between the CM2 and control classrooms which is β_{01} and its 95% confidence interval (1.96 * SEβ_{01}). Second we will report the mean difference in outcomes between the CM2 and control classrooms, β_{ij}, as an effect size in standard deviation units. The effect size will be estimated using the following formula (See Table 6):

$$\frac{\beta_{01}}{\sqrt{s^*}}$$

Where $\sqrt{s^*}$ is the pooled posttest standard deviation from the Level-1 model. If the assumption of homogeneity of group posttest variances is violated because the posttest variance of CM2 schools has been altered by CM2, then the control group standard deviation will be used in the denominator of equation 0.2 instead. In addition to the average difference between intervention and control groups, given that the model used here is a random intercept and random slopes model, we can report the estimated intercepts and slopes for each to show the variation in effect across schools (Bloom et al., 2005).

TABLE 6 6

CONNECTED MATHEMATICS 2 IMPACTS ON MATH ACHIEVEMENT FOR TWO HLM MODELS

Fixed Effect	***Parameter Estimate***	***Standard Error***	***t Ratio***	***p Value***	***Effect Size***
1. Average student outcome (math achievement/student engagement) across all schools: β_{00}					
2. Average difference between CM2 and Control classrooms on student outcomes (math achievement/student engagement), across all schools: β_{01}					
3. The average effect of student pretest scores on average student outcomes (math achievement/student engagement): β_{02}					
4. The average rth effect of the rth baseline covariate for which there is a statistically significant imbalance: β_{0r}					
Random Effect	***Variance Component***	***df***	***X^2***	***p Value***	
1. Random error associated with school j on average student outcome (math achievement/student engagement): u_{0j}					

Publication Plans and Time Schedule

Table shows the planned schedule for publication activities.

TABLE

SCHEDULE OF ACTIVITIES

Activity	Schedule
Annual Reports	June 2008, 2009, 2010
IES Draft of Final Report	March 2011
Conference Presentations	After approval of IES report
Peer reviewed journals	After approval of IES report

Interim Findings

REL: MA will report interim findings to IES in an annual report filed each project year.

IES Report

REL: MA will submit a final technical report for the study to IES for review. Our Technical Working Group will review this report prior to IES submission. We will then coordinate with IES to determine appropriate channels for sharing study findings as described in the report. Appropriate products will be generated for various target audiences (e.g., researchers, policy makers, and teachers).

Peer-reviewed Journals

Members of the study team may, following relevant IES guidelines, submit articles to peer-reviewed journals, though specific results of the current study will first be disseminated to IES.

Conference Presentations

Members of the study team may submit proposals for conference presentations subject to appropriate IES guidelines and contract restrictions. Presentations of preliminary data may be made at national scholarly conferences pending approval by IES.

17. Approval Not to Display the Expiration Date for OMB Approval

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

18. Exception to the Certification Statement

We are not seeking exceptions to the certification statement.

References

Bauer, D.J., and Curran, P.J. (2006). *Multi-level modeling of hierarchical and longitudinal data using SAS*. Cary, NC: SAS Institute.

Bloom, H.S., Richburg-Hayes, L., & Black, A.R. (2005). *Using covariates to improve precision: Empirical guidance for studies that randomize schools to measure the impacts of educational interventions* . New York: MDRC.

Boruch, R.F. (1997). *Randomized experiments for planning and evaluation: A practical guide*. Thousand Oaks, CA: Sage.

Eccles, J. and Wigfield, A. (1995). In the Mind of the Actor: The Structure of Adolescents; Achievement Task Values and Expectancy-Related Beliefs. *Personality and Social Psychology Bulletin*, *21*(3), 215-225.

Flay, et al. (2005). Standards of evidence: Criteria for efficacy, effectiveness, and dissemination. *Prevention Science*, 1-23.

Friedman, L.M., Furberg, C.D., and DeMets, D.L. (1998). *Fundamentals of Clinical Trials* (3rd Edition). New York, NY: Springer-Verlag.

Keppel, G., and Wickens, T.D. (2004). *Design and Analysis: A Researcher's Handbook* (4th Edition). Upper Saddle River, NJ: Prentice Hall.

Lipsey, M.W., & Wilson, D.B. (2001). *Practical Meta-Analysis*. Thousand Oaks, CA: Sage.

National Commission on Excellence in Education (April, 1983). *A nation at risk*. Washington, DC: U.S. Department of Education.

NCTM. (2000). *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Orr, L. (1998). *Social experiments: Evaluating public programs with experimental methods*. Thousand Oaks, CA: Sage.

Raudenbush, S.W., Spybrook, J., Liu, X., & Congdon, R. (2005, October). *Optimal design for longitudinal and multi-level research* (version 1.555). (Available at the WT Grant Foundation Website).

Schochet, P.Z. (June, 2005). *Statistical power for random assignment evaluations of education programs*. (Available from Mathematica Policy Research, Inc., P.O. Box 2393, Princeton, NJ, 08543-2393).