



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 B 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

B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS.....	4
1. Respondent Universe and Sampling Procedures.....	4
2. Statistical Methods for Sample Selection and Degree of Accuracy Needed.....	6
3. Methods to Maximize Response Rates and Deal with Nonresponse.....	8
4. Test of Procedures and Methods to be Undertaken.....	10
5. Individuals Involved.....	11

LIST OF EXHIBITS

Exhibit A. Invitation Letter

Exhibit B: Letter of Interest

Exhibit C: Memorandum of Understanding

Exhibit D: Informed Consent Form for Teachers

Exhibit E: Teacher Demographic Survey

Exhibit F: Monthly Online Teacher Survey

Exhibit G: Parental Consent Form

Exhibit H: Child Assent Form

Exhibit I: Student Math Interest Inventory (Eccles-Wigfield Survey)

B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS

1. Respondent Universe and Sampling Procedures

The sampling universe for this study is all middle schools with sixth grade students in the Mid-Atlantic region (NJ, DE, PA, D.C., MD). Following describes the recruitment plan that will be used to develop this sample universe. Schools will be recruited through a strategic recruitment plan that will be centrally managed by the REL-MA and implemented by REL-MA's Lab Extension Specialists (LESs).

A number of principles that are common sense or intuitive have been highlighted in the literature as important for a successful recruitment campaign. Synthesizing the work of Boruch (1997) in the social and behavioral sciences and Friedman, Furberg, and Demets (1998) in the biomedical sciences produced the following principles of a successful recruitment campaign:

1. Evaluate the likelihood of obtaining sufficient study participants within the allotted time;
2. Set up the infrastructure for recruitment with principal investigators and an experienced and organized recruiter in charge of recruitment and other staff required for operations;
3. Announcements of the trial should precede intention of recruitment;
4. Recruitment should begin no later than the first day of the recruitment period—commitment and willingness to spend a considerable amount of time recruiting is as important as a good plan;
5. Multiple strategies should be used;
6. Contingency plans should be available in case recruitment lags.

Conversations with Task 2.0 PIs in other Labs revealed that those that have been operating prior to the most recent awarding of REL: MA contracts are building on long standing relationships with State Education Agencies (SEAs) and Local Education Agencies (LEAs). REL: MA has begun the process of establishing such relationships in the Mid-Atlantic region. Another important lesson learned from other labs that will be implemented by REL: MA is the development of clear, non-technical descriptions of study goals, methods, roles and responsibilities of the various organizations involved, and the reasons to participate. Obviously, schools that clearly understand the benefits of participating and the goals of the study will be more likely to participate.

Public Awareness

The Task 1.1 leader, working with the state coordinators serving Delaware, New Jersey, Pennsylvania, Maryland and the District of Columbia, will launch an awareness campaign through the LESs to educate SEAs, LEAs and other education stakeholders about the Lab's work in the area of randomized control trials. LESs will serve as regional ambassadors introducing the Lab to constituents, explaining the general purpose of the Lab and the Lab's conduct of RCTs through a number of regional and local forums. These forums will bring together every superintendent in the region, although not necessarily at the same time. Educational media briefings will be held to build an additional base of knowledge and support for school participation in the Connected Math RCT.

Tracking

For early identification of schools interested in participating in the Connected Math RCT, throughout the awareness campaign the Lab will track level of interest by individual educators as well as schools and districts. Through the Customer Relationship Management Software, the Lab will track interest in participating in the RCT via the Lab's email address and calls to the 800 number. Information related to potential interest in or

barriers to participation in the RCT will be documented and used to supplement the school list, developed from the Common Core of Data.

Recruitment Training for the LESs

LESs will be trained in all aspects of the CM2 curriculum to empower them to describe the intervention to potential study participants, and to develop standardized protocols for use during the recruitment stage. LESs will also be trained in the methodological basics of a CRT; however, methodological aspects of the training will be limited to those aspects most relevant to explaining the costs and benefits of the CRT to interested participants at schools. For example, LES training on methodological issues would cover random assignment (there are ethical questions about random assignment that LES must be prepared to address) but not multi-level analysis.

The Recruitment Process

1. Analytica, a partner in REL: MA responsible for the independent random assignment of target units, will create a list of possible schools for recruiting states from a Mid-Atlantic region subset of the Common Core of data. The schools will be listed by state, locale (urban, suburban, rural), district, and number of sixth grade teachers and students.
2. The list will be augmented with data from the CM2 publisher such as schools that are using CM2 already because these schools will be excluded from the list of schools that receive an announcement about the trial (discussed below).
3. The schools in this final list will also meet the following criteria:
 - Not on the current list of *Connected Mathematics* (1 or 2) users;
 - Willing to commit to a two-year window for the study.
4. Using the final lists, a letter will be mailed to the districts, schools, and intermediate units (where applicable), inviting them to participate in the study (See Exhibit A).
5. Lab Extension Specialists (LESs), in consultation with the co-PIs, will call and the co-PIs will meet with superintendents, principals to discuss participation with them. Where applicable. LESs and the PI will contact directors and curriculum coordinators of intermediate units to recruit districts in their service area. These intermediate unit personnel are influential in curricular decisions in schools.
6. Co-PIs will follow up by meeting each superintendent/designee that has expressed interest and personnel (principals, sixth grade math teachers) at each school to explain the purpose of our project, CM2 curriculum, and their time investments for the CRT.
7. The superintendent or designee will sign a letter of interest that contains the name of the school, the number of sixth grade teachers and classrooms, demographic information verification and a description of the sixth grade mathematics curriculum and materials currently in use (See Exhibit B).
8. If appropriate, the superintendent will then place the CM2 CRT on the agenda of the School Board.
9. The School Board's final approval for participation will place the district into our pool of participants to be randomly assigned to the intervention or control group.
10. Schools will sign a memorandum of agreement (See Exhibit C). This contract includes an agreement to use CM2 as the regular school curriculum. The schools will also agree to have each participating teacher attend the CM2 training during the Summer and to make available two additional professional development windows of two hours each for additional facilitated training during each of the years of the study.
11. Customizing the human subjects approvals for each district will be done by the LES in consultation with school legal counsel and/or IRB.

Voluntary Sample of Schools

While a stratified random sampling process in some situations provides certain benefits, in this case allowing interested schools to volunteer and then sampling from this pool of self-selected schools is more appropriate for a number of reasons:

- First, due to the extensive commitment required to effectively implement the intervention, willingness and interest on the part of the participating school is critical for implementation fidelity.
- Second, the nature of the commitment will likely result in a large percentage of randomly selected schools denying participation, thereby attenuating the representativeness of the participating sample which is the primary advantage of random selection.
- Finally, whereas a random sample allows for inferences beyond the participating sample, our analytic design incorporates a fixed effects model, which restricts our inferences to the participating sample. Thus, a self-selection recruitment strategy is more consistent with the analytic assumptions and approaches proposed herein.

Issues Faced in the Recruiting Plan

The literature clearly points to the idea that successful recruitment depends not only on a well-developed plan but also on fidelity of implementation of that plan. Friedman, Furberg, and DeMets, (1998) point out that consistent monitoring of implementation of the recruitment plan is important for achieving recruitment goals.

Though the early awareness campaign, the extended period of time for recruitment (1 year), and the follow-up direct contacts from the Lab should be very effective tools for recruitment, we recognize the need to consider problems that may arise. Historically, recruitment of schools into randomized experiments has been extremely challenging because of the requirement that some schools not use the new, potentially more effective intervention (in this case CM2) if they are assigned to the control condition. REL: MA has also recognized that recruitment of enough schools to meet the power requirements discussed in the next section is critical to the success of the proposed CRT and has been the Achilles heel of previous randomized trials conducted in the social, behavioral, and education sciences (Boruch, 1997; Orr, 1998).

REL: MA will take the following steps to prevent and remedy recruiting problems:

- 1) Establish short- and long-term goals and track progress using charts and tables generated from the tracking system.
- 2) The PENN Center for Education Leadership will critically review all plans for the recruitment campaign and related materials (e.g., letters of invitation, scripts for explaining the study, and so forth).
- 3) REL: MA's Task Force on Recruiting will evaluate progress in December of the recruiting year. If one-third or less of the intended sample of schools have committed to the study at that point, the Task Force will take the following steps:
 - a. Review the incentives provided and consider changes to encourage participation.
 - b. Increase the level of recruiting effort on the part of REL-MA staff (e.g., more contact with principals and district level staff)
 - c. Consider contracting with the PENN Center for Education Leadership to conduct recruiting.

2. Statistical Methods for Sample Selection and Degree of Accuracy Needed

A serious flaw in previous studies on *Connected Mathematics*, as reported in the WWC review of middle school math curricula, was the misalignment between the unit of assignment and the unit of analysis which prohibited accurate computation of standard errors and assessment of whether positive effects of *Connected Mathematics* were real, or due to chance (i.e., not statistically significant). As stated earlier, in this study the school is the unit of assignment, therefore, it is a cluster randomized trial with random assignment at the school level. The school level is defined as all sixth grade classrooms in schools placed in the random assignment pool.

Based on this definition, we assume clustering at the school level when calculating statistical power. This assumption is consistent with Schochet (2005) who notes (p.21):

“For school-based experimental evaluations, one design option is not to sample classrooms within the intervention and control schools. For this option, either all relevant classrooms in the selected schools are included in the research sample or students are assigned directly to the research sample without regard to the classrooms they are in.”

Clearly, a multi-level modeling approach is appropriate for these nested data, with students nested within schools (that is, there is no sampling of classrooms or students within schools). Based on the strategy to recruit a self-selected sample of schools, albeit a diverse one, schools that eventually make it into the assignment pool are not a random sample from a population of schools. In this study, the school population is defined as schools on the REL Mid-Atlantic school list that were mailed an invitation to take the first steps towards participating in the Connected Math CRT. We assume, then, that the effect size estimates for schools are fixed. This “fixed effects approach” is recommended by Schochet (2005), who upon review of the sampling methodology of various trials in education concludes that, “the fixed effects case is usually more realistic in evaluations of education interventions.” Using the fixed effects model constrains inferences about effect sizes to the schools in the study. To assess potential generalizability of the study findings beyond these schools, as Schochet (2005) suggests, we will examine the pattern of effect size estimates across the characteristics of schools in the study. This will enable us to ascertain the extent to which CM2 could be effective with schools beyond those included in the study on extrastatistical grounds, if appropriate (see Keppel and Wickens, 2004, p. 530).

Our assumptions regarding the magnitude of intra-class correlations are based on the work of Schochet (2005). We assumed, then, a Level 2 ICC of 0.15 for the schools in our study, a value that Schochet uses in presenting power estimates. Although there is bound to be a small amount of clustering of classrooms within schools, it is not large enough to be modeled as between classroom variance in our power assumptions.

We will use the Terra Nova pretest of standardized student achievement. The test will be administered by REL Mid-Atlantic proctors and monitored by REL Mid-Atlantic researchers at the beginning of the school year. Analytically, the student-level test results will be aggregated to the school-level and entered as the school-level covariate in the power analysis. It was assumed that the covariate has a strong linear association with the outcome, and that this association is similar within the intervention condition. Based on the work of Bloom et al. (2005), it was also assumed that the school-level pretest is as effective a covariate for classroom outcomes as a student-level pretest would be for student outcomes. Based on the work of Schochet (2005) we assumed that for a multi-level model that included student-level baseline test scores as explanatory variables, R^2 , was at least .50. Work by Bloom et al. (2005) has shown that the pretest R^2 can be 0.56 or higher. However, we conservatively assumed R^2 to be 0.50 for power calculations. This assumption is consistent with Schochet’s assumption.

To estimate the number of schools needed to achieve power of 0.80, we used Design VII (School randomization with fixed classroom effects, which means the only source of clustering is at the school level) and corresponding results from Schochet’s Table 3, page 35, which is accompanied by the following benchmark estimates:

- Two-tailed test
- Equal number of schools randomly assigned to the intervention and control conditions
- No sampling of classrooms within schools
- Between school ICC = .15
- Three classrooms per school per grade
- 23 students per classroom
- 80 percent of students in the sample will complete both pre- and posttest
- Proportion of variance explained by school-level covariate with an $R^2 = .50$
- MDE = 0.20

REL-MA will follow several procedures in order to maximize the likelihood of achieving the 80 percent response rate from students on the pre-and posttests. First, we will administer these tests during the regular school day during school hours. Second we will administer a make up day and follow up with as many students as possible to attain high response rates from students enrolled at both pre- and posttest. We will also employ several strategies described below to handle attrition, including following up to attain posttests from students who have left the school.

As Table shows, REL-MA must recruit sixty seven (67) schools to achieve statistical power as specified in this study. To guard against the improbable but potential loss of schools, REL Mid-Atlantic will recruit an addition 3 schools to increase the total number of schools to be recruited and randomly assigned to conditions to 70. Of course far fewer schools, 44, are needed to detect a larger MDE of 0.25.

TABLE 8

REQUIRED SCHOOL SAMPLE SIZES TO DETECT TARGET EFFECT SIZES FOR A SCHOOL RANDOMIZED DESIGN WITH SCHOOL-LEVEL CLUSTERING ONLY

Standard Deviation:	Schools Needed to Detect Impact (Schools within Districts: School Level Clustering; $R^2=0.5$)
.10	259
.20	67
.25	44
.33	26

In estimating power, we relied on Schochet’s work, rather than Raudenbush’s (through use of Optimal Design), because the latter imposes inapplicable assumptions for power estimates. Specifically, Optimal Design assumes all schools are sampled from the same district. This assumption is unrealistic for this proposed study given that schools will come from different districts.

3. Methods to Maximize Response Rates and Deal with Nonresponse

a. Recruitment

The first stage in recruitment is identifying appropriate schools and developing agreements to participate with those schools. We will follow several principles for successful recruitment gleaned from the literature, as described in detail in Section B.1. Respondent Universe and Sampling Procedures, these emphasize timely action, building relationships with schools and districts, and implementing multiple strategies with contingency plans should initial efforts flag.

The second stage is to obtain teacher, parent, and student consent and buy-in to the study. Teachers are very committed to improving mathematics achievement, and therefore the option to try CM2 and receive the materials and professional development without cost is likely to be attractive to many schools and teachers.

For parents and students, again the chance to engage in a curriculum that has shown positive benefits for mathematics achievement is likely to lead to a high participation rate. Students and their families will not, however, receive any monetary or other direct incentive to participate. The benefits to each respondent are outlined in Table 9.

TABLE 9
RESPONDENT BENEFITS

Respondent	Benefit
Schools (Administrators)	<i>Assistance from CM2 publisher:</i> Implementing the curriculum Site visits
Teachers	Professional development High quality educational program \$25 non-cash incentive for each year of participation
Students (Parents by extension)	<i>Curriculum that should increase:</i> Engagement Mathematics Achievement

b. Post-assignment Attrition

Post-assignment attrition of schools is problematic because it reduces power and can bias results. Randomization equates intervention and control groups at baseline (on expectations and in the long run), and this equating is expected to hold true for post-tests as well. Post-assignment attrition, however, may distort the pre-test equivalence, because attrition rarely is totally random. Attrition will be a serious problem in the study if a school’s likelihood of dropping out of the study can be linked either to the intervention or outcome variables. To guard against this, we have increased the number of schools to be recruited from the 67 called for in our power analysis to 70.

In this cluster randomized trial, non-severe overall attrition (20% or less) or non-severe differential attrition (7% or lower) at the student level (which is not the unit of assignment), has less of an impact on statistical power (Bloom et al., 2005). To guard against severe overall attrition at the student level, we assumed 20% attrition in our power analysis. To be able to conduct an intent-to-treat analysis at the student level, as stated in the previous section, every effort will be made to collect posttest data on students that for whatever reason drop out of the study. To guard against differential attrition of greater than 7%, we will track the attrition at the student level. Where there is missing data on individual student responses at the time of post-testing, we will employ a multiple imputation strategy to estimate the missing post-test values.

Teacher attrition in the intervention schools will be managed within the normal processes of the school district. It is expected that if a sixth grade teacher leaves the school and is replaced by another sixth grade teacher, that teacher will also teach CM2 since the curriculum is used at that grade level. The replacement teacher will receive the same level of professional development (new teachers the second year of the study would be provided the same week-long training as other teachers received the first year) and collegial support. Replacement teachers will be noted as such and depending on the number of such teachers, a sensitivity analysis (including and excluding students from the analysis) will be conducted during the analysis phase.

Overall, our plan to manage attrition includes four stages: prevention, reporting attrition, classifying attrition, and bias reduction.

c. Prevention

The best solution to attrition is prevention. We will take the following steps to reduce attrition in our study:

- Clear explanation of study requirements to ensure that participating principals, teachers, and students fully understand the burden created by study participation.
- Provide \$1000 worth of equipment, such as a laptop computer, or LCD projector for the control school.

- Award all participating teachers a \$25 non-cash incentive for each year of participation.

We will also emphasize to principals, teachers, and students how they are making an important contribution to the field of math education by participating in the study.

d. Reporting

We will take the following steps to record attrition in the study:

- Quarterly phone calls to CM2 principals to inquire as to how the curriculum implementation is progressing and whether any teachers have transferred or left the school;
- Record the number of students leaving/entering schools in intervention and control schools to detect differential attrition.
- Record the number of teachers leaving/entering schools in intervention and control schools to detect differential attrition.
- Record the number of students leaving schools.

Once attrition is properly recorded, we will conduct descriptive analyses to determine whether attrition in general, or certain patterns of attrition over time, can be linked to any school characteristics. Descriptive analyses will be conducted for the whole sample and by intervention group to detect differential attrition. Assignment at the school level should greatly reduce study attrition. The authors do not expect any significant attrition at the school level, though the sample will be constructed so that a three school reduction can be tolerated. Attempts will be made to include any schools departing the study during the data collection as intent-to-treat cases. Individual teachers who leave during the course of the study will be replaced through normal district means and the CM2 curriculum will continue to be the assigned curriculum for that grade. The professional development plan has strategies for dealing with teacher turnover between years (a summer institute in the summer of 2009 will be held for new hires or teachers changing grade level) and collaborative professional development processes that are designed to take place during the course of the year will be a realistic solution to address intra-year changes, should they occur. Any replacement teachers will be noted as such in the analysis and a sensitivity analysis will be conducted to determine how sensitive the results are to inclusion of these teachers in the analysis.

4. Test of Procedures and Methods to be Undertaken

In this study, the data collection instruments are either nationally normed and tested standardized assessments (e.g., Terra Nova), or well-researched instruments designed and used by other researchers for similar purposes (e.g., the Eccles-Wigfield Attitude Survey). These instruments were selected for appropriateness for answering the research questions and for their psychometric properties (as described in Section A.2). Using instruments that have these properties helps to ensure that the estimates of the effects of CM2 derived from these measurements are accurate. We did complete a small sample pretest (9 students) of the Eccles-Wigfield Attitude Survey to determine if our estimates of completion time were correct. The test was administered to 9 students in a Southern California elementary school. The entire administration process, including teacher instructions, was approximately 10 minutes. The time to completion for a group of mixed ability students, including two main-streamed learning disabled students, ranged from 2 minutes to 4 minutes and 40 seconds. The mean time to completion was 3 minutes and 32 seconds. The Terra Nova is a timed test so we did not conduct this type of pretest.

5. Individuals Involved

Name	Title	Telephone
Taylor Martin	Assistant Professor, <i>U Texas - Austin</i>	(512) 232-9686
Kelli Millwood	Research Associate, <i>Metiri Group</i>	(310) 945-5157
Richard Brown	Assistant Professor, <i>University of Southern California</i>	(213) 740-9567
Ed Coughlin	Senior Vice President, <i>Metiri Group</i>	(310) 945-5153
Herb Turner	President, <i>Analytica</i>	(215) 808-8880

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).