Abstract Title Page

Title: Teacher-Led Math Inquiry in Belize: A Cluster Randomized Trial

Authors and Affiliations:

Krystal M. Hinerman, Darrell M. Hull, Qi Chen & Dana D. Booker - University of North Texas

Emma I. Naslund-Hadley - Inter-American Development Bank

Author Note:

This presentation was prepared with financial support from the Japan Special Fund of the Inter-American Development Bank. The views and opinions expressed in this presentation are those of the authors and do not necessarily reflect the official position of the Inter-American Development Bank or Japan International Cooperation Agency.

Abstract Body

Background / Context:

In Belize, the percentage of untrained teachers (i.e. teachers with the equivalent of a high school education or less) is one of the highest in the region of Latin America and the Caribbean. In response to growing student cohorts, the increased demand for more classroom teachers has led to significant decrease in the proportion of trained teachers at both the primary and secondary levels. Fewer than 45% of teachers at primary and secondary levels are trained, whereas some countries in the region report more than 80% are trained.

Perhaps as a result of the lack of training, a content test that was applied to a representative sample of teachers nationwide as part of our study design revealed important gaps in teachers' math content knowledge. When taking the national primary school examination (PSE) that Belizean six-graders take every year, the teachers scored on average a grade of "B". Although no systematic mapping has documented the pedagogical approach used in Belizean math classrooms, anecdotal evidence suggests that teachers rely heavily on the presentation and repetition of math procedures, sometimes combined with drills, practice, and memorization of concepts, procedures, and formulas. Students spend most of their time copying from the blackboard and plugging numbers into formulas. The common practice in the classroom is not one that actively engages students in activities that may help them develop analytical and critical-thinking skills. In addition, since the curriculum is not differentiated to meet the needs of distinct groups of students, there is often little engagement initiated by teachers in the lessons being taught.

Purpose / Objective / Research Question / Focus of Study:

We hypothesized that increased math content and pedagogical skill enhancement of teachers would lead to improved student outcomes on content knowledge in mathematics. The Teacher-Led Math Inquiry (TLMI) project posed three research questions: 1.When controlling for pretest differences, do students exposed to the TLMI treatment exhibit higher post-treatment math achievement than students in the control group at the end of one school year? (Intent to treat model); 2. To what extent does implementation fidelity as an added variable in the ITT model moderate post-treatment math achievement? (Test of treatment model); and 3. How did teachers view implementation of the TLMI approach? (Qualitative results)

Setting:

Within the nation of Belize, the Ministry of Education, Youth & Sports coordinates management of schools under six districts (Belize, Cayo, Corozal, Orange Walk, Stann Creek, and Toledo). The focus of the present study was on schools in the Belize District. In 2011, the Belize District contained 68 primary schools, ranging in enrollment from N = 12 to 1056 students per school (Mdn = 207) inclusive of eight grades that in Belize are referred to as Infant 1 and 2 (generally aged 5 and 6 years, respectively), and 1st through 6th Standard (comprising ages 7 – 12 years, respectively). Within the Belize district, approximately 60% of the primary schools are categorized as urban schools, as most of these are within Belize City. In the Belize District in 2011 there were N = 51 Government Aided Schools, N = 6 Government Schools, and N = 11 Private Schools.

Population / Participants / Subjects:

A sample of N = 24 schools were randomly drawn from the Belize District with 12 schools randomly assigned to implement the TLMI program, and 12 schools assigned to a control group. Students (N = 7564) from the 24 schools enrolled in standards Infant I to Standard 6 are described in Table 1.

Intervention / Program / Practice:

The TLMI approach is based on Colburn's (2000) classification of types of inquiry. The intervention provided students with materials and step-by-step instructions for individual investigation/exploration of concepts being taught. For teachers, TLMI was referred to as *Visible and Tangible Mathematics*. The Visible and Tangible Math model was introduced into schools as a school-wide approach that involved principals, school administrators, teachers and students. Following initial conversations about the new model, school goals were defined and teachers were assisted in the development of individual professional development plans that included steps and measurable indicators of progress toward implementation. To support teachers in the implementation of their individual plans, teachers attended face to face workshops, had contact with math teaching experts online, and received monthly in-class visits of math mentors. Table 2 includes key performance indicators for mentor observations. The intervention took place throughout the 2011-2012 school year with eight workshops offered from October to March. The workshops were followed by a capstone two-day conference in April.

The training addressed instruction of mathematical concepts in the way teachers would implement instruction in their own classrooms: through inquiry and hands on activities which varied considerably based on the concept taught and level of student. Classrooms were equipped with cost-effective manipulative materials required for the implementation of the model, including flash cards, geo-boards, geometric solids, blocks, and counters. Since most of the manipulative materials can be made at home, the mentors also assisted in the creation of teacher made materials such as ten frames and hundred charts. The aim of the TLMI treatment program was to teach teachers to understand mathematics in a tangible and visible way using the same approach they would use with their students. TLMI operates under a pressure/support model where 'pressure' comes in the form of a professional development Certificate in Primary Mathematics Teaching, and 'support' in the form of in-school mentors.

Research Design:

The design was a pretest posttest control group design with random assignment to treatment and control groups, otherwise known as a cluster-randomized trial. Schools (N = 24) were randomly drawn from the Belize District. Each school was asked to commit to participation in the study despite the 50% probability that their school may or may not be selected to implement the TLMI intervention. Upon commitment from all schools in the sample, every student from the 24 schools was assessed on a brief measure of general cognitive ability. Average school scores on this measure, combined with Urban/Rural status, and school size were used to form three strata or blocks containing 8 schools per stratum. Random assignment of schools to treatment (TLMI) or control (business as usual) groups was performed within strata, with 4 schools assigned to the

TLMI treatment from each stratum. Consequently, there were 12 schools eventually assigned to the TLMI treatment group, and 12 schools assigned to a control group, resulting in balance on three pretest covariates (general cognitive ability, urban/rural status, and school size). In order to retain participation of schools and teachers assigned to the control group, the control group teachers received a separate teacher training intervention program related to behavior management. The alternative intervention provided for the control group was focused on character development and positive discipline. None of the teachers in control group schools received intervention training in TLMI instruction or any form of math intervention, thus math instruction in the control group was business-as-usual.

Data Collection and Analysis:

Pretest and Posttest math content assessments, using the same form, were administered at the beginning and end of the 2011-2012 school year constructed from the Michigan Math Leadership Academy (MMLA) item pool. Item selection was based on alignment with objectives from the TLMI intervention and was reviewed by mentors and developers of the TLMI intervention prior to pretest administration. Eight different grade-appropriate assessment instruments were constructed from the MMLA item pool. The instrument for Infant 1 consisted of 10 items, with 25 items included in all other standard/grade instruments. Each level of the MMLA measures assessed student achievement with standard-appropriate math objectives in 5 domains: measurement, number operations, geometry, algebra, and data analysis. Specific domains assessed in each instrument are detailed in Table 3. In addition, each teacher participating in TLMI was observed on four occasions during the year (December, January, February and March) and assigned a global categorical rating of implementation by the trainers observing their classroom based upon observation and brief discussions with the teachers following observation. The categorical ratings consisted of: 4 = "got it," 3 = "almost there," 2 = "getting it," and 1 = "barely started." A single measure of implementation fidelity was created for each teacher based upon a sum of the four ratings, producing a scale from 4 to 16 (M = 10.81, SD = 2.97).

After examining the potential to model 2- and 3-level multilevel models, a design effect of 7.38 indicates a 2-level multi-level model is most suitable for examining the research questions (Maas & Hox, 2005; McCoach & Adelson, 2010). Post-treatment math z-scores were the outcome for all models. To better equate grade levels with different content assessments, z-scores were used. Post-test math z-scores were used as outcome measures for all models. All hierarchical models were conducted using HLM 7 (Raudenbush & Bryk, 2002). All models employed multiply imputed datasets where pooled results of five datasets are presented. In order to determine variance explained through adding covariates, sequential models adding additional predictors were examined as recommended (Raudenbush & Bryk).

To answer research question 1, Math Pretest was used as Level 1 predictor. The school group designation (treatment=1, control = 0) was included as a Level 2 predictor. To address research question 2 that examined how implementation fidelity affects post-treatment scores, an additional series of HLM models including the fidelity of implementation variable as a Level 2 predictor were evaluated. Table 4 provides models tested for Intent to Treat and Test of Treatment models. Focus group meetings were held with teachers from all schools implementing TLMI at the conclusion of the school year, and comments were aggregated across schools to respond to research question 3.

Findings / Results:

Intent to Treat (ITT) Models. The unconditional random intercepts model with no predictors indicated 28.59% (ICC = 0.286) of the variance in classroom post treatment math z-scores is explained at the classroom level. In the ITT model, after controlling for pre-treatment group differences by adding the pre-treatment math scores as a Level 1 predictor, including the treatment indicator as a Level 2 predictor explains an additional 1.95% of the variance in the intercept and 0.23% of the variance in the slope over the Random Intercept model with no Level 2 predictor. In summary, the results indicate that after controlling for pre-treatment group differences, the students in classrooms who received the math intervention program had higher post-treatment math scores than those in the control group.

Test of Treatment (TOT). The test of treatment models included 5738 students clustered into 242 classrooms. The final fidelity model after controlling for pre-treatment group differences, indicated a significant main effect for fidelity of implementation (γ_{01} = 0.001, SE = 0.009, t = -0.112, df = 108, p = 0.911) which explained an additional 2.77% of the Level 2 intercept variance over the model with no Level 2 predictor. Table 5 and 6 present the results of the ITT and TOT models along with the additional variance explained.

Qualitative Results. Tables 7 through 11 represent teacher responses to topics of interest in focus group discussions conducted at each TLMI school at the conclusion of the school year. All responses by teachers were open/free response, and percentages reported in these tables represent the proportion of teachers reporting statements that were consistent across each theme.

Conclusions:

The present study indicates TLMI has a substantive significant positive effect on student math performance in comparison to traditional instructional practices in Belize, specifically in the Belize District. Going back to the nineteenth century, teaching mathematics using manipulative materials has long been demonstrated to improve math performance. Meta-analytical evidence of such effects has been demonstrated to be quite strong in education since (Sowell, 1989), where students working directly with materials such as beansticks, geoboards, paper folding, or other manipulative materials under supervision achieved average achievement effects on specific objectives in primary school of .265 (Cohen's d metric).

Given that the TLMI treatment effect (ITT) for the present study of 1.95% is the R^2 equivalent to a Cohen's d of approximately .28 (Rosenthal, 1994), the findings from this impact evaluation are consistent with those in the literature, but in the context of a developing country where few investigations have been conducted. Moreover, the test of treatment (TOT) analysis that incorporated variations of implementation by teachers suggests that it may be possible for TLMI to generate R^2 effects of 2.77% (Cohen's d equivalent = .34).

Ultimately, if the program is utilized throughout the country or expanded within certain districts, improved student performance should be re-evaluated with other relevant outcome measures such as student enjoyment of mathematics and persistence in problem-solving, as these outcomes could be considered theoretically linked to TLMI; but also more distal measures such as performance on the Primary School Exam (PSE) in addition to proximal math performance measures such as those used in the present study that are benchmarked externally.

Appendices

Appendix A. Extended References

Arthur, M. W., Hawkins, J., & Pollard, J. A. (2002). Measuring risk and protective factors for substance use, delinquency, and other adolescent problem behaviors. *Evaluation Review*, *26*(6), 575-601. doi: 10.1177/019384102237850

Bishop, A. J. (1980). Spatial abilities and mathematics education—A review. *Educational Studies in Mathematics*, *11*(3), 257-269.

Bishop, A. J. (1983). *Space and geometry*. In R. Lesh & M. Landau (Eds.), Acquisition of mathematics concepts and processes (pp. 175–203). New York, NY: Academic Press.

Boggan, M., Harper, S., & Whitmire, A. (2010). Using manipulatives to teach elementary mathematics. *Journal of Instructional Pedagogies*, *3*(1), 1-10.

Bruner, J. S. (1964) 'The course of cognitive growth'. *American Psychologist 19*, 1–15.

Clarkson, P., & Presmeg, N. (2008). *Critical issues in mathematics education: Major contributions of Alan Bishop*. (Eds.). New York, NY: Springer.

Colburn, A. (2000). An inquiry primer. Science Scope. Special Issue. Retrieved May 25, 2013 from: http://www.experientiallearning.ucdavis.edu/module2/el2-60-primer.pdf

Galton, F. (1883). *Inquiries into the human faculty and its development*, London : Macmillan.

Ji, P., Segawa, E., Burns, J., Campbell, R. T., Allred, C. G., & Flay, B. R. (2005). A measurement model of student character as described by the Positive Action program. *Journal of Research in Character Education, 3*(2), 109-120.

Martin, T. A. & Shipley, W. C. (2009). Shipley-2 Block Patterns Task. Los Angeles, CA: Western Psychological Services.

Maas, C. J., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *1*(3), 86-92.

McCoach, D. B., & Adelson, J. L. (2010). Dealing with dependence (part 1): Understanding the effects of clustered data. *Gifted Child Quarterly, 54*, 152-155.

Michigan Mathematics Leadership Association (2005). Test Generator. http://mathassessments.mscenters.org/index.php

Muthén, L. K., & Muthén, B. O. (1998-2010). *Mplus* (Version 6.1). Los Angeles, CA: Muthén & Muthén.

Piaget, J., & Inhelder, B. (1971). *Mental imagery and the child.* London: Routledge & Kegan Paul.

Pollard, J. A., Hawkins, J. D., Arthur, M. W., Pollard, J. A., Hawkins, J. D., & Arthur, M. W. (1999). Risk and Protective Factors Survey. *Social Work Research*, *23*, 145-158.

Presmeg, N. C. (1986). Visualisation in high school mathematics. *For the Learning of Mathematics, 6*(3), 42-46.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R. T., & du Toit, M. (2011). *HLM 7*. Lincolnwood, IL: Scientific Software International.

Raudenbush, S. W., Spybrook, J., Congdon, R., Liu, X., Martinez, A., Bloom, H., & Hill, C. (2012). Optimal Design Software for Multi-level and Longitudinal Research (Version 3.01). *Available from www.wtgrantfoundation.org or from sitemaker.umich.edu/group-based*.

Renshaw, P & Brown, R. (2008). Four models of the process of integrating everyday and scientific discourse: Replacement, interweaving, contextual privileging, and pastiche. *Language and Education*.

Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), *The Handbook of Research Synthesis*. New York, NY: Sage. pp. 239.

Shipley, W. C. & Gruber, C. P. (2009). Shipley-2 Vocabulary. Los Angeles, CA: Western Psychological Services.

Shipley, W. C., Gruber, C. P., Martin, T. A., & Klein, A. M. (2009). *Shipley-2 Manual*. Los Angeles, CA: Western Psychological Services.

Smith, S. S. (2009). *Early Childhood Mathematics* (4th ed.) Boston: Pearson Education Using manipulatives. Retrieved June 5, 2013, from: http://www.teachervision.fen.com/pro-dev/teaching-methods/48934.html

Sowell, E. J. (1989). Effects of manipulative materials in mathematics instruction. *Journal for Research in Mathematics Education*, *20*(5), 498-505.

Spearman, C. E. (1927). *The Abilities of Man: Their Nature and Measurement*. Macmillan: London.

Vygotsky, L (1987) *Thinking and Speech*. In R.W. Rieber & A S Carton (Eds.), The collected works of L.S. Vygotsky: Problems of general psychology. New York, NY: Plenum Press.

Werner, H. (1964). *Comparative psychology of mental development*. New York, NY: International Universities Press.

Appendix B. Tables and Figures

Table 1. Participant demographics and missing data percentages.

		Original Data		Final Data**	
	Item	Count	Percent	Count	Percent
	Total participants	7678	100%	7564	100%
	MATH treatment group	3651	48%	3017	47%
	CREST treatment (control) group	4027	52%	4547	53%
	No teacher ID*	270	4%	0	0%
Standard	Infant 1	895	12%	842	13%
	Infant 2	827	11%	818	12%
	Standard 1	1074	14%	862	13%
	Standard 2	1012	13%	810	12%
	Standard 3	1443	19%	1004	15%
	Standard 4	901	12%	824	13%
	Standard 5	773	10%	721	11%
	Standard 6	753	10%	682	10%
Demographics	No demographic information	1869	24%	863	11%
	Male	2959	39%	2886	44%
	Female	2850	37%	2814	43%
	Creole	3275	43%	3214	49%
	Garifuna	447	6%	437	7%
	Maya	160	2%	158	2%
	Metizo	1632	21%	1602	24%
	Other	268	3%	262	4%
Missing Data	Missing all pre and post scores*	954	12%	0	0%
	Missing pre and post math*	1072	14%	0	0%
	Missing pre and post resiliency	1529	20%	551	8%
	Missing pre math	1749	23%	667	9%
	Missing post math	1821	24%	706	9%
	Missing pre resiliency	2554	33%	1548	20%
	Missing post resiliency	2411	31%	1330	17%
	Have pre and post math	5180	67%	5180	79%
	Have pre and post resiliency	4242	55%	4236	65%

*Deleted from final models. ** Final data after deleting cases based on missing essential data.

Table 2. Key Performance Indicators for Mentor Observations.

KPI	Teacher should know how to:	Observable Evidence	
1. Student progress needs to be accurately assessed, recorded and used to inform instruction	Assess student performance and keep accurate records of progress	None	Teacher uses workbook tests only, does not keep continuous records
		Some	Authentic testing is occasionally used, records are kept
		Good	Assessment records are linked to the new teaching methods
		Strong	Teachers are using accurate assessment records to inform instruction
2. There is a need to address individual learning differences	Give opportunities for small group work, mixed groups and individual attention	None	Traditional rows and seating, students facing the teacher, instruction always delivered to the whole class
		Some	Students working on different pages and activities, some small groups
		Good	Most students engaged, small group instruction based on student needs, groups encouraged to work together, student show ability to work in this environment, teacher notices “lost” students
		Strong	All students engaged, individual student needs are addressed, groups are fluid, students teach each other, teacher effectively respondes to “lost” students
3. There is a need to teach students to understand underlying math concepts	Create effective plans for instruction of math concepts using hands on materials	None	No materials evident, no use of materials, paper and pencil tasks only
		Some	Materials are present and students are using them
		Good	Materials use is daily in student focused activities
		Strong	Students are fully engaged in hands on activities
4. Students need to increase their level of ease with basic facts in order to facilitate solving more advanced equations	Teach strategies that will help development of mental math	None	No table charts or any assists for basic facts and families are posted in classrooms
		Some	Basic facts, timed quizzes are used, charts and tables are posted for student use
		Good	Teacher asks if answers are “reasonable,” students self-check and use assists without prompting
		Strong	Strong math dialogue is used, teacher asks for reasons “why” or “how do you know?” more frequently

Table 3. MMLA Mathematics Assessments Per Standard by Strand.

Strand	Domain	Infant 1	Infant 2	Standard 1	Standard 2	Standard 3	Standard 4	Standard 5	Standard 6
Numbers & Operations	Meaning, notation, place value, and comparisons		3, 4, 5, 6, 7	14	1, 2, 3, 4, 5, 6, 7, 18	1, 4, 9, 11, 12	6	1, 2	
	Number relationships and meaning of operations		8, 9, 10, 11, 12, 13	1, 2, 3, 4, 5, 9, 10, 11, 12, 13	12, 14, 15, 16, 17	2, 8, 10, 13	8		1, 5, 6, 7, 13
	Fluency with operations and estimation		14, 15, 16	6, 7, 8	8, 9, 10, 11, 13	3, 5, 6, 7, 14, 15, 16	1, 2, 3, 4, 5, 7	3, 4, 5	2, 4, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19
Algebra	Patterns, relations, functions, and change							6, 7, 8, 14	20, 22
	Representation							9, 10, 11, 15, 16, 17, 18	21
	Formulas, expressions, equations, and inequalities							12, 13	
Measurement	Units and systems of measurement	1	17, 18	16, 17	19, 20, 21, 22	17	9, 10, 11, 12	19, 20, 21	
	Techniques and formulas for measurement	2, 3, 4	1, 2	15		19, 20, 21	13, 14, 15		
	Problem solving involving measurement			18	23	18		22	
Geometry	Geometric shape, properties, & mathematical arguments	5, 6, 7, 8, 9, 10	19, 20, 21	19, 20, 21, 22		22	16, 17, 18, 19, 20, 21	23, 24, 25	
	Location and spatial relationships		22, 23	23, 24					
	Spatial reasoning and geometric modeling		24		24				
	Transformation and symmetry					23			23, 24, 25
Data & Probability	Data representation		25	25	25	24, 25	22		
	Data interpretation and analysis						23, 24, 25		3
	Probability								

Note: Numbers in boxes represent item numbers.

Table 4. Intent to Treat and Test of Treatment Models*.

	Intent to Treat
Level-1 Model	$\text{PostMath}_{ij} = \beta_{0j} + \beta_{1j}\text{*PreMath}_{ij} + r_{ij}$
Level-2 Model:	$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{*Treatment}_j + u_{0j}$ $\beta 1j = \gamma_{10} + \gamma_{11}\text{*Treatment}_j + u_{1j}$
Mixed Model:	$\text{PostMath}_{ij} = \gamma_{00} + \gamma_{01}\text{*Treatment}_j + \gamma_{10}\text{*PreMath}_{ij} +$ $\gamma_{11}\text{*Treatment}_j\text{*Pre-Math}_{ij} + u_{0j} + u_{1j}\text{*Pre-Math}_{ij} + r_{ij}$
	Test of Treatment
Level-1 Model:	$(\text{PostMath})_{ij} = \beta_{0j} + \beta_{1j}\text{*(Pre-Math)}_{ij} + r_{ij}$
Level-2 Model:	$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{*(Fidelity}_j) + u_{0j}$ $\beta_{1j} = \gamma_{10} + \gamma_{11}\text{*(Fidelity}_j) + u_{1j}$
Mixed Model:	$\text{Post-Math}_{ij} = \gamma_{00} + \gamma_{01}\text{*(Fidelity}_j) + \gamma_{10}\text{*(Pre-Math)}_{ij} + \gamma_{11}\text{*(Fidelity}_j)$ $\text{*(Pre-Math)}_{ij} + u_{0j} + u_{1j}\text{*(Pre-Math)}_{ij} + r_{ij}$

*With *i* referring to students, *j* referring to teachers, β as student level coefficients, γ as teacher level coefficients, and all random error terms being estimated.

Table 5. Hierarchical Linear Models: Regression Coefficients, Statistical Tests and Variance Components

		M1: Unconditional				M2: Means as Outcomes				M3: Random Intercepts				M4: Random Slopes and Intercepts			
A.	**INTENT TO TREAT MODELS**																
	Element	Coef	SE	*t*	df	Coef	SE	*t*	df	Coef	SE	*t*	df	Coef	SE	*t*	df
γ_{00}	Intercept	0.052	0.037	1.411	281	-0.037	0.053	-0.701	280	0.042	0.031	1.375	281	-0.024	0.044	-0.544	280
γ_{01}	Treatment indicator					0.170*	0.073	2.328	280					0.127*	0.061	2.076	280
γ_{10}	Pre-test Math Zscore									0.439**	0.017	25.485	281	0.441**	0.025	17.919	280
γ_{11}	Treatment indicator X Pre-test													-0.004	0.034	-0.118	280
σ^2	Level 1 Residual Variance	0.854				0.854				0.696				0.696			
$\tau \pi$	Level 2 Intercept Residual Variance	0.341				0.333				0.224				0.220			
τ_1	Level 2 Slope Residual Variance									0.035				0.035			
B.	**TEST OF TREATMENT**	M1: Unconditional				M2: Means as Outcomes				M3: Random Intercepts				M4: Random Slopes and Intercepts			
	Element	Coef	SE	*t*	df	Coef	SE	*t*	df	Coef	SE	*t*	df	Coef	SE	*t*	df
γ_{00}	Intercept	0.062	0.039	1.580	241	0.062	0.039	1.601	240	0.056	0.033	1.676	241	0.056	0.033	1.710	240
γ_{01}	Fidelity of Implementation					0.007*	0.002	2.859	240					0.005*	0.002	2.414	240
γ_{10}	Pre-test Math Zscore									0.428**	0.019	22.919	241	0.427**	0.019	22.884	240
γ_{11}	Fidelity X Pre-Test													-0.001	0.001	-0.582	240
σ^2	Level 1 Residual Variance	0.860				0.860				0.705				0.705			
$\tau \pi$	Level 2 Intercept Residual Variance	0.332				0.319				0.225				0.219			
τ_1	Level 2 Slope Residual Variance									0.035				0.035			

*p <.05 **p <.001

Table 6. Hierarchical Linear Models: Additional Variance Explained

	Included Predictors		Intent to Treat (ITT)			Test of Treatment (TOT)		
	Level 1	Level 2	Level 1	Level 2		Level 1	Level 2	
Model vs Comparison	Student	Teacher	σ^2	$\tau \pi$	τ_1	σ^2	$\tau \pi$	τ_1
Model 2 vs Model 1	NA	**Treatment***	-0.01%	2.31%		0.00%	3.82%	
Model 3 vs Model 1	**Pretest***		18.50%	34.14%		18.01%	32.11%	
Model 4 vs Model 3	**Pretest**	**Treatment***	-0.01%	1.95%	0.23%	0.00%	2.77%	0.34%

*Indicates added variable compared to nested model. Bolded variables indicate statistically significant associated coefficients.

Table 7. Teacher report of impact of TLMI on school, classroom and/or students.

Teacher Impact		Selected Comments
a. improve pedagogy/method learning	37%	…given me ideas and techniques to teach math concepts more hands-on, interactive and in more meaningful ways.
b. improve attitude/motivation/enjoyment	17%	...a lot more confident in the teaching of mathematics; …motivates teachers to do a much better job…
c. improve understanding of student learning	9%	It open(ed) my eyes to see why students are not getting the (math).
Student Impact		
a. improve academic involvement	43%	The manipulatives were a wonderful addition because it allows the students to be more interactive in the lessons.
b. improve conceptual understanding	37%	…allowed students to better relate to the concepts…; more room for growth and real-life experience
c. improve attitude/motivation/enjoyment	35%	Children in my class enjoy math time.; …remove the idea that math is not for everyone;
d. improve social skills and/or behavior	7%	Students have learned to cooperate and participate more in group and class activities.

Table 8. Teacher report on the methods and materials provided by TLMI

Teacher Impact		Selected Comments:
a. improve teaching enjoyment and attitude	21%	Math is much more fun and the time seems shorter; I, myself, enjoy math more; …makes teaching a lot easier;
Student Impact:		
a. improved attitude/motivation/enjoyment	35%	I have seen how it makes the classroom [livelier] and lets children believe that math can be fun.
b. improve conceptual understanding	49%	…are able to understand how they get it; …my students readily and easily grasp concepts;
c. improve social skills and/or behavior	9%	…helped me to get my kids to learn to share and work with each other.

Table 9. Best experiences with TLMI as reported by teachers

Teacher Impact		Selected Comments:
a. improved attitude toward math teaching	12%	The realization that math is not as difficult and complicated as I thought it was.
b. learning teaching strategies and methods	47%	The best thing…was learning new, fun ways to help my students learn math better other than the chalk and talk.
c. interaction with program instructors	12%	…instructors that were very helpful and encouraged me to give out my very best in the classroom…made me believe…
d. workshops	10%	…learn about using the materials at the workshops and with this knowledge felt more competent in the classroom.
e. collaboration with colleagues	12%	…many teachers shared experience about way[s] to teach or how they taught a given topic.
f. materials, manipulative and/or moodle	20%	The materials are fun to work with; …able to go on Moodle and use other resources
Student Impact:		
a. improve learning environment	16%	...seeing children's faces light up as they learn…; students looked forward to learning without the teacher involved.
b. manipulatives/materials	27%	…use of materials in a way that made math easy, interesting and full of fun

Table 10. Most challenging experiences with TLMI as reported by teachers

Teacher Impact		Selected Comments:
a. put ideas into action/articulating new ideas	17%	…to be able to articulate with confidence then explain the different idea or steps in using the tools…
b. unlearning old ways of teaching	12%	My challenge was to take what I learned and use it in my class, especially when I [am] used to one way.
c. completing/sending assignments/journal	20%	Time kept running on me. I really had to make an effort and give time to complete my assignments.
d. teaching students how to use manipulatives	11%	Getting the students to appreciate manipulatives and take care of them while using them correctly and not as a toy.
e. lack of manipulatives	14%	To create the materials needed because we didn't get ours; …materials were sometimes hard to come by.
f. using and/or access to moodle/computer/internet	20%	…I have no access to internet where I live; getting the information off the internet
g. prep-time for activities/work load/scheduling	15%	…additional work we had to add into our already packed schedule.

Table 11. Teacher reports on ways in which the TLMI program could be improved

a. more demo lessons/supervision	27%	Maybe if the mentors would model some of their techniques by taking over a classroom full of real students!
b. longer program/time/more often/summer seminars	20%	…more time in teaching the methods; Program should have been longer!
c. more manipulatives	26%	…more manipulatives that cannot be handmade; …we can share but not at the same time