Predicting Student Achievement: A Path Analysis Model on A Mathematics Coaching Program

Scott A. Zollinger[1] and Patti Brosnan

Department of Mathematics Education, The Ohio State University

(Received 25 September 2011; accepted 28 November 2011)

In response to calls for mathematics education reform, researchers at The Ohio State University, working with the Ohio Department of Education, developed the Mathematics Coaching Program (MCP). Coaches from 164 schools have participated in this classroom embedded professional development program designed to promote standards-based instructional practices. Preliminary results indicated MCP has a positive impact on student achievement. To provide supporting evidence of these results, researchers developed a path analysis model consisting of seven components to determine factors that were predictors of student achievement both before and after schools participated in MCP. Dependent variable components were pre- and post-MCP test scores and independent variable components represented economically disadvantaged students, non-white students, disability students, number of years in MCP, and coach mathematical content knowledge. Results indicated that the initial and modified theoretical models were not acceptable fits for our fourth grade sample data, but many parameter values were consistent with previous research. Disability, SES, and ethnicity were significant predictors of pre-MCP test scores and were negatively correlated. For post-MCP test scores, t-values of disability and ethnicity decreased to non-significant levels but t-values for SES nearly doubled. The number of years a school participated in the program was a significant predictor of post-MCP test scores but coach mathematical content knowledge was not. Overall, the path analysis model did not test as an acceptable fit for this data, but the final version represents a starting point for future testing.

I. INTRODUCTION

Equity is one of six principles addressed by the National Council of Teachers of Mathematics (NCTM) in the *Principles and Standards for School Mathematics* (2000). In their discussion of equity, NCTM (2000) emphasizes that schools are obligated to see that all students should have access to and participate in an equitable mathematics program that is responsive to prior knowledge, intellectual strengths, and personal interests. However, history indicates that equitable mathematics programs are lacking and economically disadvantaged students, students with disabilities, and non-white students, as well as other student subgroups, for many years, have been victims of low expectations (National Council of Teachers of Mathematics, 2000). While these ideas are not necessarily new to the field of mathematics education, they seem to have progressed to the forefront in the last few years. As a result, a push has been made for mathematical instruction to move from a traditional teacher-centered transmission of information to a more student-centered construction of knowledge. Instead of rote memorization of rules, formulas, and procedures, students are expected to learn mathematics with conceptual understanding that is enhanced through classroom interactions (National Council of Teachers of Mathematics, 2000).

While student-centered instruction has become more prevalent in mathematics education, the critical element in raising the expectations for groups of students that have traditionally been left behind remains the classroom teacher. They are the ones in the trenches with students every day and their instructional decisions continue to profoundly influence how knowledge is constructed. Teachers, though not always prepared to do so, are burdened with the responsibility of choosing worthwhile mathematical tasks that pique students' interests and encourage mathematical thinking. NCTM (2000) suggests that teachers would be bettered prepared if they took advantage of opportunities to reflect on and refine instructional practices through collaboration with colleagues. Instead, they often work in isolation, rarely communicating about instructional practices with anyone outside their classroom. While teacher collaboration can be addressed and improved in a variety of ways, this report focuses on results from classroom embedded professional development associated with a mathematics coaching program designed to improve instructional practices and student learning.

II. THE MATHEMATICS COACHING PROGRAM (MCP)

In an effort to address issues of equity, improved instructional practices, and student learning with respect to mathematics education, researchers at The Ohio State University (OSU) in conjunction with the Ohio Department of Education (ODE) developed the Mathematics Coaching Program (Brosnan & Erchick, 2008). This program is a classroom embedded professional development program that promotes standards-based instructional practices through improved content knowledge and improved pedagogical content knowledge. Since January of 2006, 164 schools throughout Ohio have participated in the program and preliminary results indicated MCP has had a positive impact on classroom instruction and student achievement (Brosnan & Erchick, 2008). These results were consistent with results from a study of 4th graders using 2000 National Assessment of Educational Progress (NAEP) data, where Lubienkski (2006) found that NCTM standards-based instruction and knowledge were positively related to student achievement.

According to the MCP design, teachers from each of the 164 schools were hired to become mathematics coaches and they were required to participate in four days of professional development each month throughout the school year. Two days were spent working with the other coaches and researchers from OSU to improve mathematical content knowledge and pedagogy while also increasing awareness of social justice issues in mathematics classrooms. Coaches were encouraged to actively participate in numerous discussions with their colleagues about how to improve mathematical instruction and thus student learning. They spent the other two days of professional development in smaller regional groups with an MCP facilitator in which they discussed issues that occurred during their coaching experiences. The MCP model was structured so that each coach worked collaboratively with three or four teachers for a period of approximately six weeks and then moved on to a new group of teachers. Expectations were that coaches would co-plan and co-teach lessons with the classroom teacher and that they would conduct pre- and post-conferences to discuss and reflect on what happened in the classroom and the direction of future instruction. Coaches were also expected to help analyze student achievement data and use it to guide instructional practices within the classrooms. Although coaches' roles and responsibilities varied from day-to-day, the previous description provides a general framework for the MCP model.

The ultimate goal of MCP was, and still is, to improve student learning and achievement. Initially, the program approached and worked with many of the lowest performing schools in Ohio based on Ohio Achievement Test (OAT) results. During the last six years, a wide variety of schools participated in MCP but the majority were low performing in terms of mathematics achievement. Many students from the participating schools represented at least one of the historically low-performing subgroups including economically disadvantaged, non-white, and disability. The MCP program philosophy was consistent with that of NCTM's equity principle in the belief that all students could learn and understand mathematics. The goal was to use classroom embedded coaching as an avenue for improved student achievement and development of consistently higher results across typically low-performing subgroups.

A. Purpose of the Study

The intent of this study was to provide supporting evidence of the impact of MCP on student achievement with an emphasis on low-performing student subgroups. A path analysis model was developed to determine factors that were predictors of student achievement both before and after schools participated in MCP. The dependent observed variables were OAT scores prior to participation in MCP and OAT scores from the end of the final year of participation in MCP. Five independent observed variables were investigated: 1) percentage of economically disadvantaged students; 2) percentage of non-white students; 3) percentage of disability students, 4) number of years a school participated in MCP; and 5) measured mathematical content knowledge of coaches. The path analysis model was designed based on theoretical expectations supported in the literature and then tested using the student version of LISREL 8.8 (Jöreskog & Sorbom, 1993). The relationships between the predictors and tests scores were evaluated and model modifications were made until the best possible fit was achieved for the given components.

B. Theoretical Model Components and Rationale

Seven observed variables were included in this study, three of which (pre- and post-MCP test scores and years in MCP, were not necessarily situated in the model based on literature but more based on variable type. The location of these three variables is briefly discussed and then a more detailed description of the remaining components follows. Pre- and post-MCP test scores were treated as dependent variables and the remaining components as independent variables placed according to their expected influence on these scores. The model design included a path from pre-MCP test scores to post-MCP test scores as we expected these two variables to be highly correlated. The variable for years of participation in the program was ordinal, ranged from one to three, and could only be situated to predict the post-MCP test scores. The remaining components of the model are discussed within the context of the literature.

C. Non-White and/or Economically Disadvantaged Students

Many studies discuss these two subgroups together, an approach used in this report as well. In a study using NAEP data from 1990, 1996, and 2000, findings indicated that White students, especially those of high socioeconomic status (SES), were experiencing more mathematics education focused on problem solving and critical thinking instead of memorization and practice of rules as compared to other groups of students (Sarah Lubienski, 2002). Additionally, the author found that the majority of black students continued to view mathematics as memorization. A separate study found that non-number curricular emphasis was more positively correlated with higher SES students than lower SES students (S Lubienski, 2006). Recent 4th grade NAEP data (Aud & Hannes, 2010) showed that from 2007 to 2009, there was no measureable change in average mathematics scores for ethnic groups (White, Black, Hispanic, Asian, American Indian) but that scores did remain higher than those of years prior to 2007. The mathematics achievement gap between White and Black and White and Hispanic students also did not change from 2007 to 2009. While there were some gains prior to 2007, the most recent data indicated that progress in this area had stalled.

Studies centered on standards-based curriculum reform offer evidence that students achieve a greater level of mathematical understanding when learning through this type of instructional practice (Boaler, 2002; Gutstein, 2003; R. Reys, B. Reys, Lapan, Holliday, & Wasman, 2003). In two of the studies (Boaler, 2002; Gutstein, 2003), non-White and low income students participated in open-ended projects, relevant to the real world and their own lives, that promoted multiple strategies, student interaction, and the assumption of different roles for students and teachers. Students' attitudes toward mathematics changed to seeing it as a valuable tool, they became better at explaining their mathematical reasoning, and they gained in confidence (Gutstein, 2003). Students of all levels and from all backgrounds were able to develop a conceptual understanding of mathematics using the reform-oriented curriculum (Boaler, 2002).

Based on this literature, researchers from MCP expected that schools participating in a program where coaches encouraged standards-based instructional practices with their teachers would show gains in student achievement relative to race and SES and that the results of these two subgroups would likely be correlated. Each subgroup was situated as an independent observed variable or predictor of both pre-and post- MCP test scores with the expectation that their level of significance would decrease after participation in the program.

D. Disability Students

Disability students encounter a wide variety of difficulties associated with learning mathematics. Their disabilities can range from a mild learning impediment to sever behavioral problems and can be both mental and physical. They often have diverse needs and require modifications in instruction and assessment practices. Research has shown that students diagnosed with learning disabilities perform lower on mathematics achievement tests than students without learning disabilities and that the gap widens with time (Fusaro & Shibley, 2008; Judge & Watson, 2011). Additionally, students with learning disabilities frequently come from lower SES backgrounds and belong to minority groups (Judge & Watson, 2011). Often, many of the modifications necessary for instruction require some aspect of standards-based reform. For example, students who struggle with auditory learning may require some type of visual representation or manipulative to better understand the concept being studied. One of the main areas of focus for standards-based instructional practices is the use and acceptance of multiple representations and/or solution strategies (National Council of Teachers of Mathematics, 2000). Certainly, some disability students need more specific modifications but instructional practices that allow them to construct their own knowledge based on previous knowledge and preferential learning methods would seem like a good place to start the modification process.

In this path analysis model, the disability subgroup was situated as an independent, observed variable or predictor for both pre- and post-MCP test scores and initially uncorrelated with any of the other predictors even though Judge and Watson (2011) suggested a possible correlation with non-White and SES subgroups. The same expectation applied for the disability subgroup that applied for the SES and non-White subgroups as disability was expected to be a significant predictor of pre-MCP test scores and then become less significant in predicting post-MCP test scores.

E. Coach Mathematical Content Knowledge

According to Ball, Hill, and Bass (2005), teachers' mathematical knowledge is critical to their ability to select appropriate instructional materials, to assess students' progress, and to make sound judgments with respect to presentation, emphasis, and sequencing. Their research indicated that mathematical knowledge of many teachers was inadequate and that those who scored higher on measures of mathematical knowledge for teaching produced better gains in student achievement. Analysis of 700 first and third grade teachers and approximately 3000 students found that teachers' performance on "knowledge for teaching" questions significantly predicted gain scores on the Terra Nova which was considered a reliable and valid standardized test (Ball, Hill, & Bass, 2005) . They noted that teachers of higher poverty and non-White students tended to have less mathematical knowledge than teachers of non-minority students and suggested that improving teacher knowledge may be one way to close the student achievement gap.

Professional development and college courses were suggested as potential avenues for inservice teachers to improve mathematical knowledge. Ball, Hill, and Bass (2005) studied the California K-6 mathematics professional development institutes and found teachers did learn content knowledge and their gains were related to the length of professional development and to curricula that focused on proof, analysis, exploration, communication, and representations – essentially standards-based curricula. A separate study found that teachers taking more college courses had significantly higher levels of teaching outcome expectancy – the belief that the educational system can work for all students regardless of outside influences (Swackhamer, Koellner, Basile, & Kimbrough, 2009). Ball, Hill, and Bass (2005) state that "…knowing mathematics for teaching demands a kind of depth and detail that goes well beyond what is needed to carry out the algorithm reliably."

A primary focus of MCP was to improve coaches' mathematical knowledge so that they could then pass the knowledge on to their classroom teachers and thus affect student achievement. Using the intent of MCP and the supporting literature, coach mathematical content knowledge was included in the model as an independent observed variable or predictor for post-MCP test scores. The expectation was that after participating in MCP, the coaches' improved mathematical content knowledge would result in student achievement gains and thus would be a significant predictor of post-MCP test scores. This predictor was also expected to be correlated with the number of years in the program.

II. METHOD

Between January of 2006 and May of 2010, MCP conducted monthly, two-day professional development sessions – from September through May – for coaches from 164 schools. Coaches then returned to their schools and worked with classroom teachers with the intent of improving classroom instruction and student learning. The schools considered for this study varied in terms of typology, student population, grade level, and previous academic performance level. Some schools, for a variety of reasons (often funding issues), chose to remove themselves from MCP without completing one full year in the program. As a result of a combination these variables, not all 164 schools had sufficient data to be included in this analysis.

Sample

The first limiting factor taken into account for this study was grade level. The path analysis could only be tested for one grade level at a time and only two, third and fourth grade, were present in enough schools to provide a minimum of 100 data points. The decision was made to use fourth grade data because it offered the most initial data points (124). However, data from seven schools, all from the same district, were eliminated from the study due to problems between MCP and their teachers union. According to their union, the coaches were allowed to participate in the monthly professional development sessions but they were not allowed to do any collaborative work with classroom teachers. Since the whole idea of embedded professional development is collaborative work with classroom teachers, these coaches were unable to affect classroom instruction or student learning. Therefore, data from those seven schools were discarded for this study thus reducing the sample size to 117. Additionally, three other schools were part of an initial attempt to expand MCP by forming a regional satellite group. However, they were not completing the program according to MCP guidelines and were eventually dismissed. Therefore, data from those three schools also were not included thus reducing the total sample size for this analysis to 114 schools.

Data Description

The model design consisted of seven components; two observed dependent variables and five observed independent variables. The two dependent variables were pre- and post-MCP test scores (Pretest and Posttest) and the five independent variables were; 1) percentage of economically disadvantaged students (SES); 2) percentage of non-white students (Ethnicty); 3) percentage of disability students (Disabil); 4) number of years a school participated in MCP (Yrsprog); and 5) measured mathematical content knowledge of coaches (LMTscore). With the exception of years of participation in MCP and mathematical content knowledge of coaches, data were obtained from yearly ODE reports posted on their website. Values for percentages of economically disadvantaged, non-White, and disability students were obtained from ODE reports corresponding to the school's final year in MCP. Paired-sample t-tests indicated no statistical difference between percentages comparing a school's first year in MCP with their final year (See Table 1). The following is a brief description of the data used for each component of the analysis.

Pre- and Post MCP Test Scores. Student scores on the OAT were grouped into five categories; limited, basic, proficient, accelerated, and advanced. A score was considered passing if at the proficient, accelerated, or advanced levels. Yearly reports offered by ODE presented, by grade level, the percentage of students at each level and also one percentage for the combined levels of proficient, accelerated, and advanced – called proficient and above. This single combined percentage from the fourth grade mathematics portion of the OAT was used for the pre- and post-MCP test scores for each school. Test results from the spring prior to a school's first year of participation in the program were used as the pre-MCP test scores and OAT results from the spring of their final year of participation were used as post-MCP test scores.

Economically Disadvantaged Students. According to ODE, students were considered economically disadvantaged if they were eligible to receive the free or reduced-price lunch or if they were known to be recipients of or whose guardians were known to be recipients of public assistance. Based on these criteria, schools provided data to ODE who then reported the percentage of economically disadvantaged and non-disadvantaged students for each building. For the purposes of this model, we used the percentage of economically disadvantaged provided by ODE. Since this data was not offered by grade level, the assumption was made the fourth grade percentage was reflective of the school percentage.

Percentage of Non-White Students. Calculated from data provided by schools, the yearly report presented by ODE provided percentages of students in each school by ethnic subgroups. The subgroups included were White, Black, Hispanic, American Indian or Alaskan Native, Asian or Pacific Islander, and Multiracial. The three largest subgroups for MCP schools were White, Black, and Hispanic with White and Black the two predominant subgroups. The remaining subgroups were rarely large enough to report a percentage. Therefore, to arrive at the ethnicity component and include all non-White subgroups for the model, the percentage of White students was subtracted from 100 to obtain the percentage of non-White students and those results were entered as data for the ethnicity component.

Percentage of Disability Students. Similar to ethnic subgroup data, the annual ODE report provided a percentage of students classified as disabled for each individual school. For this model, data at the fourth grade level was again assumed to be reflective of the entire school and thus the percentages determined by ODE were used for the disability component.

Table 1. Subgroup Paired-Samples Test

Samples	Mean	SD	t	df	p-value ($\alpha = .05$)
Disabilpre - Disabilpost	-1.175	4.712	-1.411	31	.168
SESpre – SESpost	-2.678	5.919	-1.354	31	.186
Ethnictypre -Ethnictypost	-1.094	11.189	-1.045	31	.304

Years Participating in MCP. When schools elected to participate in MCP, their coaches were promised the opportunity to attend professional development sessions for three years. Beginning in September and ending in May, a two-day session was presented each month of the school year with different sessions offered to coaches based on their number of years of participation. At any given time, MCP included first-, second-, and third-year coaches. Data for this component was classified as ordinal and was limited to the values of one, two, or three.

Coach Mathematical Content Knowledge. As part of the Learning Mathematics for Teaching (LMT) Project , a group of researchers at the University of Michigan (UM) developed and tested a pool of questions to measure mathematical content knowledge for teachers (Heather Hill, Schilling, & Ball, 2004). The questions were analyzed for both reliability and validity. With approval from the researchers at UM, researchers from MCP used a subset of the questions to create an instrument for measuring coach mathematical content knowledge. This instrument included 26 number and operation questions, 28 patterns, functions and algebra questions, and 19 geometry questions each valued at one point. To achieve a reliability of between .7 and .8, researchers at UM recommended 12-25 questions for each scale and a sample size of at least 60 coaches (Ball et al., n.d.). As previously discussed, the coach sample size was 114 so minimum values for both conditions were met.

All coaches participating in MCP were expected to complete the LMT assessment during the fall and spring of their first year and then each spring for the remaining two years. The assessments were scored by graduate research assistants with a number correct for each category of questions and then a combined total number correct. The combined total (maximum = 73) for each coach was the data value used for this study. For consistency, the score used for each coach was the most recent LMT score available. For example, if their final year with the program was 2008, then the score used for this model was from the LMT assessment they completed in the spring of 2008. If for some reason that that score was not available, then the fall of 2007 score was used and this process continued until available data was found. If no data was available for a particular coach, then the missing value was estimated using the Similar Response Pattern Imputation (SRPI) in PRELIS, the data manipulation and basic statistical analyses software associated with LISREL 8.8. Data from all remaining variables were selected for the matching process and conditions warranted that eight missing values were estimated using SRPI. In schools with more than one coach, an average of the coaches' scores was used and when a coach was in two buildings, the same score was used for both buildings.

III. RESULTS

Preliminary Analysis

School/Coach Descriptive Data. Since there are large differences in middle and high school physics courses between China and USA, and in both countries there are no explicit direct training of reasoning ability in school curricula, the naturally formed two education systems provide a unique controlled setting for studying the effects from content learning on the development of reasoning ability. The result suggests that training on content knowledge in physics (or science and mathematics in general) in current traditional school education settings, which is carried out at a substantial level of complexity in China, doesn't affect the development of general scientific reasoning abilities.

All values of skewness and kurtosis were in the slight-to-moderate range thus the recommendation was that the maximum likelihood (ML) estimators of LISREL were acceptable for the model (Lei & Lomax, 2005; Schumacker & Lomax, 2010). Boxplots (See Figure 1) of variable data indicated four outliers, two in disability data and one each in post-test scores and LMT scores. All values were approximately three standard deviations from their respective means so the input data were reviewed and verified a second time. They were found to be true outliers and therefore considered for deletion before the model run. However, since the sample size was only 114, the decision was made to include the outliers in the data.

Correlation Matrix. The pre-model analysis also included using SPSS to determine the Pearson Correlation coefficients for each pair of variables to be studied in the model. Analysis indicated several significant correlations ($p < .05$) and the resulting correlation matrix shown in Table 3 was used as input for the path analysis model.

Model Testing

Path analysis modeling was used to assess the relationship between a mathematics coaching program and student achievement. The software used for this analysis was the

Table 2. School/Coach Descriptive Statistics

Variable	N	Mean	SD	Skewness	Kurtosis	Outlier
Pretest	114	60.18	17.56	-.05	-.72	NA
Posttest	114	65.15	17.07	-.72	-.01	19
Disabil	114	15.69	4.73	.57	.54	30.2, 28.2
SES	114	63.17	23.01	-.15	-.80	NA
Ethnicty	114	39.30	29.78	.50	-1.00	NA
Yrsprog	114	1.75	.69	.38	-.85	NA
LMTscore	114	54.92	10.35	-.54	-.17	24

Table 3. Correlation Matrix

	Pretest	Posttest	Disabil	SES	Ethnicty	Yrsprog	LMTscore
Pretest	1						
Posttest	.580**	1					
Disabil	-.280**	-.238*	1				
SES	-.408**	-.558**	-.338**	1			
Ethnicty	-.313**	-.330**	-.009	.491**	1		
Yrsprog	-.421**	-.053	.014	.101	-.007	1	
LMTscore	.019	.001	.090	.048	-.048	.118	1

**Correlation is significant at the .01 level

*Correlation is significant at the .05 level

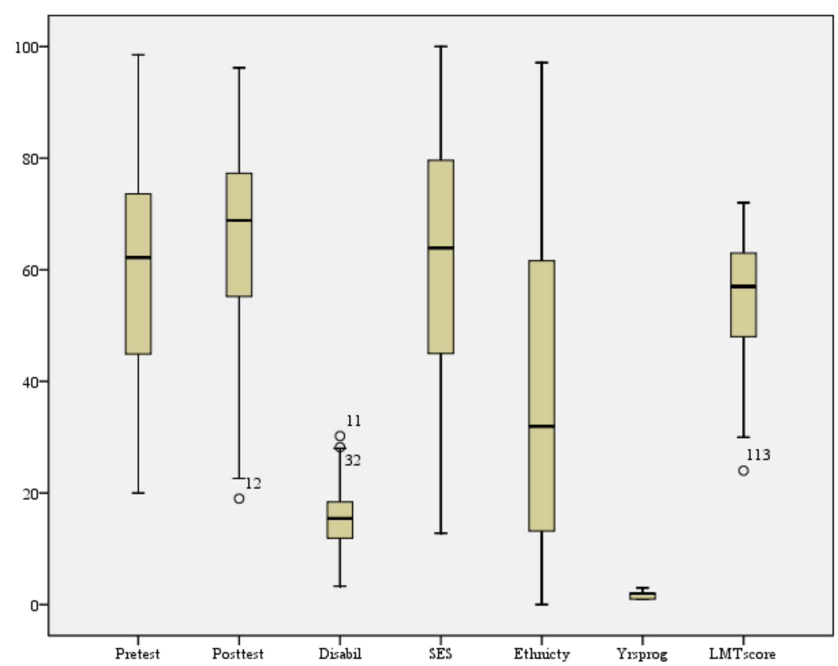

Figure 1. Boxplots with outliers for observed variables. The numbers represent the position of the data point within the input table, not values of the outliers.

student version of LISREL 8.8 (Jöreskog & Sorbom, 1993) and the estimation method used was maximum likelihood. The initial theoretical model is shown in Figure 2. Although all variables in the model were observed variables, "dummy" latent variables were used in the LISREL syntax so that specific independent variable correlations could be designated. The software automatically correlated all independent variables and in some instances, this made no sense. For example, the percentage of disability students at a school should not have been correlated with the number of years a school participated in MCP. Therefore, error variances of all observed variables and covariances of non-correlated "dummy" latent variables were initially set equal to zero.

Model 1 Results. Results from the initial path analysis model indicated that disability (t = -2.388), SES (t = -2.568), and ethnicity (t = -2.033) were all significant predictors ($|t| > 1.96$) of pre-MCP test scores but they explained only 19% (R = .190) of the variance. The negative values indicated that an increase in percentage of students in each category resulted in lower test scores, consistent with the previously mentioned research. For the post-MCP test scores, only pre-MCP test scores (t = 7.028), SES (t = -4.940) and years in the program (t = 3.143) were significant predictors with the combined six explaining 54% (R = .536) of the variance. Additionally, the covariance between SES and ethnicity (t = 4.685) was significant. This model was over-identified since the number of distinct values in the sample variance-covariance matrix (S) was greater than the total number of free parameters (See Table 4). The resulting degrees of freedom was 10.

To evaluate the fit of the hypothesized model with the sample data, four global fit indices were analyzed according to criteria summarized by Shumacker and Lomax (2010, p. 76). The Chi-Square ($\chi 2$) test of statistical significance was used to determine whether the sample covariance matrix (S) was significantly different from the model-implied covariance matrix. Ideally, $\chi 2$ would not be significant or would at least be approaching the number of degrees of freedom (df). Other fit indices were also examined because the $\chi 2$ statistic is influenced by sample size. The root-mean-square error of approximation (RMSEA) was another test of model fit with acceptable values of .05 to .08. The standardized root-mean-square residual was analyzed for an acceptable value of less than .05. The fourth value utilized, the goodness-of-fit index (GFI), indicates an adequate fit with a value of at least .90 but is preferred to be .95 or greater. Based on these criteria, the hypothesized model was not a good fit for the sample data as indicated in Table 5. While the GFI (.899) was at the low end of the adequate range, the remaining indices were not within the range of acceptable values.

Analysis of the standardized residuals (SR) and modification indices (MI) led to the consideration of several model modifications. The largest positive or negative values of the standardized residuals and the largest modification indices suggested the most beneficial modifications based entirely on fit but in most cases, those modifications made no substantive sense. A path from post-MCP test scores to pre-MCP test scores (SR = -2.13, MI = 23.2) was considered and rejected based on the irrational idea of using post-test scores to predict pre-test scores. Another path, this one from years in the program to pre-MCP test scores (SR = -4.55, MI = 22.5), again was illogical because the number of years with MCP would have had no impact on a school's test scores from the year prior to their first participation in MCP. The first modification that seemed justifiable was to add a covariance between the disability and economically disadvantaged variables (SR = 3.59, MI = 17.5). As was discussed earlier, research indicated that students with learning disabilities more often come from lower SES backgrounds and belong to minority groups (Judge & Watson, 2011). Thus, the only other suggested modification that made sense was to add a covariance between disability and ethnicity (SR = -.10, MI = 4.6). Since the MI for the first acceptable modification was much larger than the latter and for analysis purposes, only one modification at a time was made, the covariance between disability and economically disadvantaged was added and the new model was tested. The second theoretical model is shown in Figure 3.

Model 2 Results. Results from the modified path analysis model indicated that disability (t = -2.195), SES (t = -2.361), and ethnicity (t = -1.986) all remained significant predictors of pre-MCP test scores and the variance explained increased to approximately 22% (R = .218). The negative values were consistent with both research and the first model. For the post-MCP test scores, again pre-MCP test scores (t = 7.028), SES (t = -4.561) and years in the program (t = 3.143) were significant predictors with a slight increase in explained variance to 55% (R = .546). The covariance between SES and ethnicity (t = 4.941) remained significant and the added covariance between disability and SES (t = 3.889) was also significant. See Table 6 for a complete list of t-values for both models. With the added covariance, one degree of freedom was lost but the model remained over-identified since the number of distinct values in S remained greater than the total number of free parameters and the resulting degrees of freedom was 9.

Based on the χ^2 difference test ($\chi^2_{\text{diff}} = 17.711$, $df_{\text{diff}} = 1$), the model modification resulted in a significant change in χ^2. However, χ^2 was still significant ($\chi^2 = 26.649$, $p = .002$) and not close to the number of degrees of freedom ($df = 9$). While the remaining global fit indices improved toward the ranges previously established, only the GFI (.937) was close to acceptable. The second theoretical model tested was a better fit than the first model but was still at a less-than-acceptable level. See Table 7 for a comparison of parameter values and fit indices between the two models.

Since the modified theoretical model did not test as a good fit, modification indices and standardized residuals were again analyzed for possible additional modifications. However, the only suggested modifications were similar to those rejected for the first model – adding a path from post-test scores to pre-test scores and adding an error covariance between years in the program and pre-test scores. Therefore, both options were again rejected for the same reasons as previously mentioned and model testing for the given

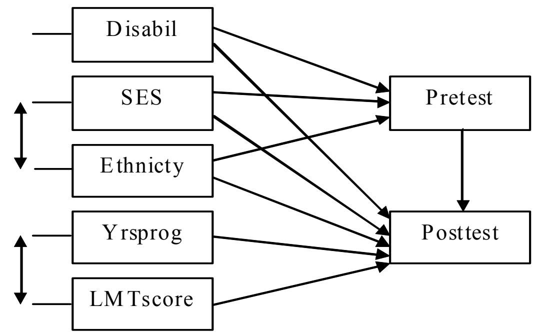

Figure 2. Initial theoretical path analysis model for predicting student achievement based on OAT scores.

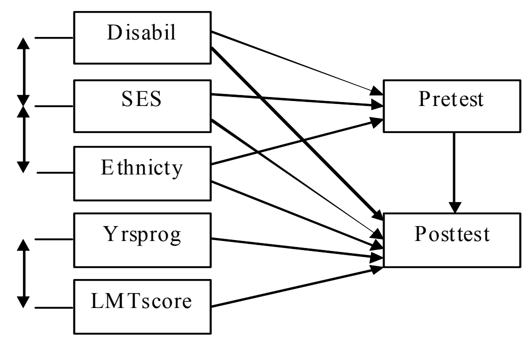

Figure 3. Modified theoretical path analysis model (added covariance between disability and SES) for predicting student achievement based on OAT scores.

Table 4. Level of Identification

# Distinct Values in S	# Free Parameters
$p(p + 1) / 2$	9 path coefficients
	2 predictor covariances
	5 predictor variances
	2 equation error variances
7(8) / 2 = 28	18 total free parameters

p = # of observed variables

Table 5. *Goodness of Fit Statistics Model 1*

Statistic	Value	Fit Indication
χ^2 (p value)	44.360 (0.000)	$\chi^2 > df = 10$, (not acceptable)
RMSEA	.174	Not .05 to .08 (not acceptable)
SRMR	.124	Not less than .05 (not acceptable)
GFI	.899	Not greater than .90 (not acceptable)

Table 6. T-Values for Models 1 and 2

Paths	**Model 1**	**Model 2**
Disability → pre-test	-2.39*	-2.20*
SES → pre-test	-2.57*	-2.36*
Ethnicity → pre-test	-2.03*	-1.99*
Disability → post-test	0.60	0.56
SES → post-test	-4.94*	-4.56*
Ethnicity → post-test	0.37	0.36
Years in program → post-test	3.14*	3.14*
LMT score → post-test	-0.26	-0.26
Pre-test → post-test	7.03*	7.03*
Correlation of Independent Variables		
SES, ethnicity	4.69*	4.94*
Years, LMT	1.25	1.25
Disability, SES	--	3.89*

*Significant predictor or correlation for $|t| > 1.96$

Table 7. Standardized Estimates and Selected Fit Indices for Models 1 and 2

Paths	Model 1	Model 2
Disability → pre-test	-.20	-.20
SES → pre-test	-.25	-.25
Ethnicity → pre-test	-.20	-.19
Disability → post-test	.04	.04
SES → post-test	-.37	-.37
Ethnicity → post-test	.03	.03
Years in program → post-test	.20	.20
LMT score → post-test	-.02	-.02
Pre-test → post-test	.50	.50
Equation Error Variances		
Pre-test	.81	.78
Post-test	.46	.45
Correlation of Independent Variables		
SES, ethnicity	.49	.49
Years, LMT	.12	.12
Disability, SES	--	.34
Selected Fit Indices		
χ^2	44.360	26.649
df	10	9
p value	.000	.002
RMSEA	.174	.132
SRMR	.124	.100
GFI	.899	.937

components seemed to have progressed as far as possible. The only additional option for this model would have been to remove a non-significant path but research indicated the given paths were justified so the decision was made to reject this option for the present study.

III. DISCUSSION AND CONCLUSIONS

The proposed theoretical model did not fit the fourth grade sample data at an acceptable level but many of the parameter values were consistent with previous research. Disability, SES, and ethnicity were all significant predictors of pre-MCP test scores and were negatively correlated thus indicating that as student populations increased in these demographics, test scores decreased. Previous research indicated similar findings (Aud & Hannes, 2010; Fusaro & Shibley, 2008; Judge & Watson, 2011). Results for the post-test scores were somewhat conflicting. With both models, the t-values of disability (t_1 = 0.60, t_2 = 0.56) and ethnicity (t_1 = 0.37, t_2 = 0.36) decreased to non-significant levels and were positively correlated, but the t-values for SES (t_1 = -4.94, t_2 = -4.56) nearly doubled and remained negatively correlated . Given the social justice component of the MCP professional development sessions, the expectation would be to make all three variables less of a factor in predicting test scores. Results indicated possible success in this regard with respect to disability and ethnicity but the opposite effect with SES. One possible explanation for this finding was that the mean percentage of SES students per school (*M* = 63.17, *SD* = 2.16) was substantially larger than the percentages for disability (*M* = 15.69, *SD* = .44) and ethnicity (*M* = 39.30, *SD* = 2.79). One coach likely would have had less ability to affect all students in the larger percentage group than in the smaller groups. This result also suggests that the structure of the social justice component of MCP might have unintentionally resulted in greater influence with respect to ethnicity and disability than socioeconomic status and could necessitate a review of this component of MCP.

The two independent variables representing specific aspects of MCP also offered conflicting results. The number of years a coach/school participated in the program was a significant predictor of post-MCP test scores and positively correlated in both models (t_1 = 3.14, t_2 = 3.14) but the LMT score representing coaches' mathematical content knowledge was not a significant predictor (t_1 = -0.26, t_2 = -0.26) of post-MCP scores. The first result would be expected of any coaching program with goals of improving student achievement but the second result was, at least initially, somewhat surprising and inconsistent with previous research suggesting that teacher content knowledge impacts student achievement (Ball et al., 2005; Swackhamer et al., 2009). The fact that the number of years in the program was positively correlated with LMT scores in both models (t_1 = 1.25, t_2 = 1.25) was consistent with research but the lack of significant correlations was not and might suggest one possible reason for this result. If the mathematical content knowledge of coaches did not consistently increase over time, then the content knowledge passed on to teachers from coaches – second-hand content knowledge – would likely have been even more inconsistent. This, combined with the fact that the amount of time classroom teachers were in direct contact with students far exceeded that of coaches, might help explain why LMT results were not good predictors of student test scores. Another important factor to remember with respect to coaches' LMT scores was that when data was not available from their final year of participation in MCP, scores from prior years were used. When scores from prior years were used, coaches obviously had not yet received all PD sessions aimed at improving mathematical content knowledge.

Although results for the LMT scores were unexpected, they might prove useful for developing a better fitting model for the fourth grade data. Since the options for improving this model were limited but nine degrees of freedom remain available, possible alternatives would be to replace LMT scores with a better fitting component or just add one or more new components to the current model. One possible component that could be added is the *Learning About Mathematics and Pedagogy* (LAMP) instrument scores that measure both content and pedagogy. Developed by MCP researchers, coaches and classroom teachers complete this instrument at the beginning and end of the academic year. Teachers are given 10 selections of student work which they analyze for student thinking and then offer suggestions for teacher instructional decisions. Since this directly assesses classroom teachers and includes the pedagogy component, both of which the LMT does not assess, the LAMP might prove to be a better predictor of student achievement and thus a better fit for the model. Concerns about using this instrument include subjective scoring and that some type of average score would have to be used to represent the teachers at each grade level for every school. Teachers must also sign a consent which states their agreement to or denial of use of their data and those choosing not to allow MCP to use their results would reduce available data. Additionally, the magnitude of LAMP data far surpasses that of the LMT and would require substantial additional time for analysis.

Overall, the path analysis model did not test as a good fit for this sample data but the final version seems like a good place to start for future testing. Decisions must be made about whether or not to include LMT scores in future models and what additional components could be added. Coach and teacher LAMP scores offer one possible direction for future model testing with MCP data. As more data becomes available, other options are likely to be considered as well.

ENDNOTES AND REFERENCES:

[1] Correspondence concerning this article should be addressed to Scott A. Zollinger, Department of Mathematics Education, 287 Arps Hall, 1945 North High Street, Columbus, OH 43210. E-mail: zollinger.4@buckeyemail.osu.edu

Aud, S., & Hannes, G. (2010). The condition of education 2010 in brief (No. NCES 2010-029). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2010029

Ball, D., Bass, H., Blunk, M., Cole, Y., Delaney, S., Hill, H, Phelps, G., et al. (n.d.). Learning Mathematics for Teaching (LMT) Project. Retrieved August 19, 2011, from http://sitemaker.umich.edu/lmt/appropriateness

Ball, D., Hill, H, & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach 3rd grade and how can we decide ? American Educator, 29(3), 14-17.

Boaler, J. (2002). Learning from teaching: Exploring the relationship between reform curriculum and equity. Journal for Research in Mathematics Education, 33(4), 239.

Brosnan, P., & Erchick, D. (2008). Mathematics coaching: Findings from a program showing students' and teachers' mathematics content improvement. Presented at the National Council of Teachers of Mathematics Regional Conference. Retrieved from http://mcp-coaching.osu.edu/Documents/NCTMRegional08.pdf

Fusaro, J., & Shibley, I. (2008). Learning disabilities and reading mathematics achievement of eighth graders in the USA. Journal of the International Association of Special Education, 9(1), 51-56.

Gutstein, E. (2003). Teaching and learning mathematics for social justice in an urban, Latino school. Journal for Research in Mathematics Education, 34(1), 37.

Hill, Heather, Schilling, S., & Ball, D. (2004). Developing measures of teachers' mathematics knowledge for teaching. The Elementary School Journal, 105(1).

Jöreskog, K., & Sorbom, D. (1993). LISREL 8 : Structural equation modeling with the SIMPLIS command language. Chicago Ill.: Scientific Software International.

Judge, S., & Watson, S. (2011). Longitudinal outcomes for mathematics achievement for students with learning disabilities. Journal of Educational Research, 104(3), 147-157.

Lei, M., & Lomax, R. (2005). The effect of varying degrees of nonnormality in Structural Equation Modeling. Structural Equation Modeling, 12(1), 1-27.

Lubienski, S. (2006). Examining instruction, achievement, and equity with NAEP mathematics data. Educational Policy Analysis Archives, 14(14), 1-33.

Lubienski, Sarah. (2002). A closer look at black-white mathematics gaps: Intersections of race and SES in NAEP achievement and instructional practices data. Journal of Negro Education, 71(4), 269.

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

Reys, R., Reys, B., Lapan, R., Holliday, G., & Wasman, D. (2003). Assessing the impact of standards-based middle grades mathematics curriculum materials on student achievement. Journal for Research in Mathematics Education, 34(1), 74.

Schumacker, R., & Lomax, R. (2010). A beginner's guide to structural equation modeling (3rd ed.). New York: Routledge.

Swackhamer, L. E., Koellner, K., Basile, C., & Kimbrough, D. (2009). Increasing the self-efficacy of inservice teachers through content knowledge. Teacher Education Quarterly, 36(2), 63-78.

