Technical Report #6 – January 2004

Structural Equation Models Assessing Relationships Among Student Activities, Ethnicity, Poverty, Parents' Education, and Academic Achievement

Jeff Joireman

Martin Abbott

The Washington School Research Center (WSRC) is an independent research and data analysis center within Seattle Pacific University. The Center began in July 2000, funded through a gift from the Bill and Melinda Gates Foundation. Our mission is to conduct sound and objective research on student learning in the public schools, and to make the research findings available for educators, policy makers, and the general public for use in the improvement of schools. We believe that sound data and appropriate data analysis are vital components for the identification of school and classroom practices related to increased student academic achievement.

Washington School Research Center
3500 188th St. S.W., Suite 328
Lynnwood, WA 98037
Phone: 425.744.0992
Fax: 425.744.0821
Web: www.spu.edu/wsrc

Jeffrey T. Fouts, Ed.D.	**Martin L. Abbott, Ph.D.**	**Duane B. Baker, Ed.D.**
Executive Director	Senior Researcher	Director -
Professor of Education	Professor of Sociology	School Information Services

Structural Equation Models Assessing Relationships Among Student Activities, Ethnicity, Poverty, Parents' Education, and Academic Achievement

---

A Technical Report For
The Washington School Research Center

Foreword

Previous research and technical reports produced by the Washington School Research Center have explored the relationships among important school and student variables and academic achievement. Family level income, student ethnicity, the school environment, and specific classroom practices have all been shown to be important determinants of student success in school. At the school level, family income as reported in the percentage of students qualifying for free/reduced lunch is generally the strongest predictor of student achievement. However, the effects of poverty are being at least partially overcome in some schools by other factors. In this sixth technical report produced by the Center, Professors Joireman and Abbott continue the analysis of state-wide data bases to add to our understanding of the factors that help explain the variability we see in student achievement levels.

Their results based on 9th grade data show that, indeed, family income, parents' education levels, and student ethnicity are important in predicting student achievement. Using a sophisticated structural equation modeling approach they are able to sort out the unique contributions each of these factors makes to student achievement, while considering three *student* factors at the same time. The three "student activities" they include in the analyses are the amount of homework a student does, the number of school or community activities in which the student participates, and the amount of television the student watches. As it turns out, all three of these variables are significant predictors of student achievement *apart* from family income, parents' education levels, and student ethnicity.

While the statistical analyses are complex, the results can be stated relatively simply. Homework is positively related to achievement. Too much time spent watching TV is related to lower achievement and, unfortunately, the amount of TV watched tends to be greater among non-white students. Participation in school and community activities appears to encourage higher achievement beyond the effects of ethnicity or low-income status. While participation is less in non-white populations, it is greater among low income students. This suggests that encouraging participation may be an important way to affect achievement especially since schools cannot change ethnicity or income.

Interestingly, participation in school and community activities appears to be almost as strong a predictor of school success as does the amount of time a student spends doing homework. Perhaps this finding indicates just how important it is that we create schools where students feel “connected” and where they are important members of a community engaging in an important endeavor.

Jeffrey T. Fouts
Executive Director
Washington School Research Center

Table of Contents

Structural Equation Models Assessing Relationships Among Student Activities, Ethnicity, Poverty, Parents' Education, and Academic Achievement

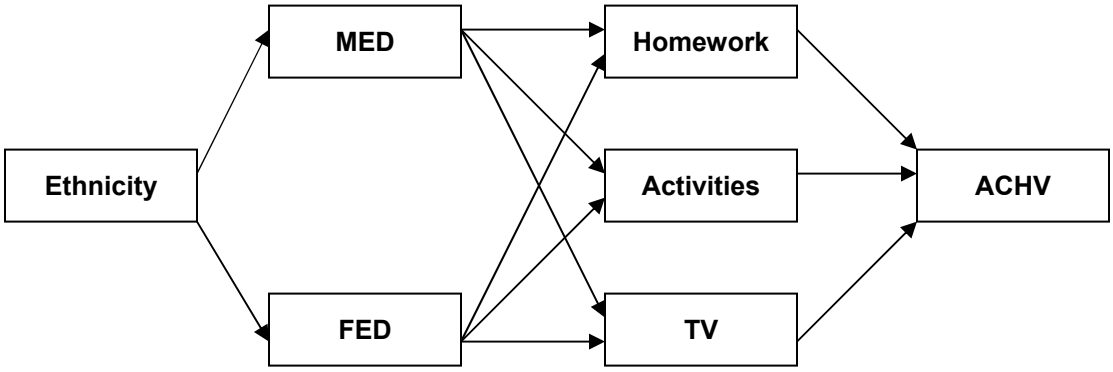
Introduction

In the face of dwindling resources for education, school administrators must make difficult decisions regarding funding of various programs. Within the current school reform climate, extracurricular activities represent one likely category of programs to be targeted for cuts. Although this strategy is understandable, critics argue that cutting such activities may have adverse consequences. In this report, we present a study examining the relationship between extracurricular activities and academic achievement within a large sample of high school students from Washington State.

Past research on the relationship between extracurricular activities and academic achievement has typically focused on one of two competing models. One of the earliest theories postulated that extracurricular activities and academic achievement were at odds, such that time devoted to extracurricular activities would interfere with academic achievement (e.g., Coleman, 1959, 1961). More optimistic models presume that, far from a zero-sum relationship, extracurricular activities and academic achievement are complementary, such that extracurricular activities can enhance students' development (Holland & Andre, 1987) and/or commitment and identification with school (Finn, 1989), which in turn may enhance students' academic achievement, adjustment, and probability of staying in school. Several recent studies provide support for this complementary model (Camp, 1990; Mahoney & Cairns, 1997; Marsh, 1992). The present study represents an attempt to replicate these findings, and locate extracurricular activities within a theoretical model involving ethnicity and parents' education. In addition, the present study examines the role of two additional "student activities" (homework and TV) within this model. We assume that ethnicity will be related to parents' education, which in turn will predict student activities, which in turn will predict reading and math achievement (see Figure 1).

Figure 1

Initial Theoretical Model – Individual Level



Method

Participants and Procedure

Data used for the analyses in this report are 2001 and 2002 9th grade ITBS/ITED data from Washington schools. In addition to achievement scores, individual student responses to several opinion questions included in the ITBS/ITED database were analyzed along with school-level ethnicity and low-income measures.[1] Structural Equation Modeling was performed using AMOS 4.0 software.

Variables

Table 1 presents a description of the variables used in this study. At the individual level of analysis, all variables save reading and math achievement were based on student self-reports. Students reported their ethnicity, parents' education, number of hours per week they spent doing homework, number of (extracurricular activities) they had participated in during the past 12 months, and the number of hours per day they spent watching TV.[2] Reading and math achievement were based on students' standardized scores from the Iowa Test of Basic Skills (ITBS). To examine relationships at the school level, we aggregated students' responses within each school and used the percentage of students on free or reduced lunch as a proxy for low income. To enhance the stability of our findings, schools with fewer than 5 students were dropped from the (school level) analyses, as were schools labeled A (alternative), I (institutional), or S (special).

[1] Data were provided by the Office of the Superintendent of Public Instruction in Washington State.

[2] Original response options were presented using letters and clear descriptions. The scales reported here (in parentheses) reflect our recoding of the responses into a numerical form appropriate for analysis.

Table 1

Variables Used in the Current Study in Order of Appearance in Theoretical Models

Variable's Abbreviation	Variable's Operational Definition and Coding for Analysis
Ethnicity	Coded 1 (non-white) or 2 (Caucasian).
Low Income	Percentage of students in a school on free or reduced lunch. Used as a proxy for low income.
MED	Mother's Education. Coded 1 (didn't complete high school) thru 5 (completed an advanced degree).
FED	Father's Education. Coding same as MED.
Homework	Number of hours per week student engages in homework. Coded 1 (none) thru 8 (10 or more).
Activities	Number of activities run by student's school or community that student has engaged in over past 12 months. Coded 1 (0 activities) to 5 (4 or more).
TV	Number of hours of TV student watches per day. Coded 1 (none) to 7 (5 or more).
Reading	Standardized reading score from the Iowa Test of Basic Skills.
Math	Standardized reading score from the Iowa Test of Basic Skills.

Results

Student Level Analyses

Data were analyzed using structural equation modeling techniques designed to evaluate how well a causal model, like the one shown in Figure 1, represents the data (e.g., whether ethnicity influences student activities through its influence on parents' education). At the individual (i.e., student) level of analysis, we first tested the initial model shown in Figure 1. After evaluating the fit of this initial model, we conducted further analyses in an effort to improve the fit of the model according to several fit indexes. To identify the most significant and meaningful modifications, we examined modification indexes and freed paths that were most likely to improve the fit of the model and which made theoretical sense. To evaluate the fit of the models, we focused on a variety of different types of fit indexes including the Goodness of Fit Index (GFI), the Adjusted Goodness of Fit Index (AGFI), the Comparative Fit Index (CFI), and the Root Mean Squared Error of Approximation (RMSEA). Following convention (e.g., Byrne, 2001), models with GFI, AGFI, and CFI values greater than .90, and a RMSEA less than or equal to .10 were judged as providing a reasonable fit to the data.[3] Table 2 summarizes the results of the initial and final path models involving reading and math in the years 2001 and 2002. As can be seen in the left half of this table, the initial models for reading did not fit the data especially well (i.e., AGFIs and CFIs < .90, and RMSEAs > .10). Accordingly, we modified the model by adding three paths.

[3] For completeness, we also examined the Chi-Square test and its associated probability value. However, because the Chi-Square test is sensitive to large sample sizes, like those employed in the current study, our discussion of the models' fit focuses only on the four fit indexes.

Table 2

Summary of Initial and Final Path Models: Individual Level Analyses

	Reading				Math			
	2001		2002		2001		2002	
Fit Measure	Initial	Final	Initial	Final	Initial	Final	Initial	Final
GFI	0.95	0.98	0.95	0.98	0.96	0.98	0.96	0.98
AGFI	0.85	0.92	0.84	0.92	0.87	0.92	0.87	0.92
CFI	0.80	0.94	0.80	0.94	0.83	0.94	0.83	0.94
RMSEA	0.14	0.10	0.15	0.10	0.13	0.10	0.13	0.10

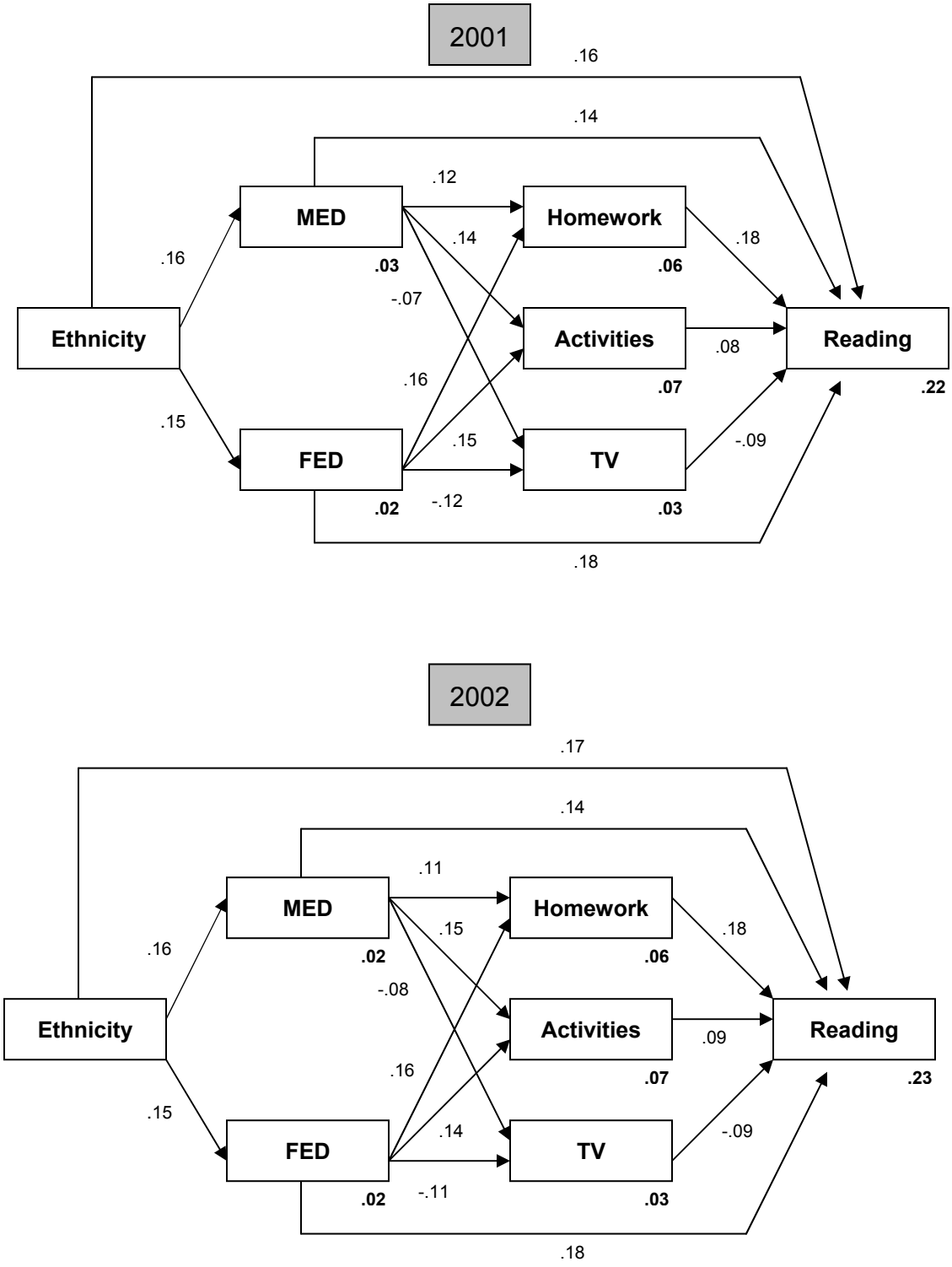
Note. DF = degrees of freedom. P = probability level. GFI = Goodness of Fit Index. AGFI = Adjusted Goodness of Fit Index. CFI = Comparative Fit Index. RMSEA = Root Mean Squared Error of Approximation.

The final path diagrams are shown in Figure 2.[4] Each diagram displays the standardized regression coefficients for the relevant paths in the model which, unless otherwise noted, were all statistically significant at the .05 level (non-significant paths are denoted *ns*). Below each outcome (or endogenous) variable, we also present the percentage of variance in that variable that is explained by the relevant set of predictor variables. As an example, the top diagram shows that 6% of the variance in homework is explained by the set of two predictors pointing to it (i.e., mother's and father's education). Finally, it should be recalled that we categorized students as either non-white (coded 1) or Caucasian (coded 2). Hence, positive regression coefficients involving the variable "ethnicity" indicate that Caucasians scored higher on the relevant outcome variable, whereas negative coefficients indicate that minorities scored higher on the outcome variable.

[4] In addition to the three additional structural paths discussed in the paper, we also estimated the (positive) correlation between mother's and father's education in all models. However, in the interests of simplifying the presentation of the diagrams, we have omitted this correlation.

Figure 2

Final Path Models for Reading: Individual Level Analyses



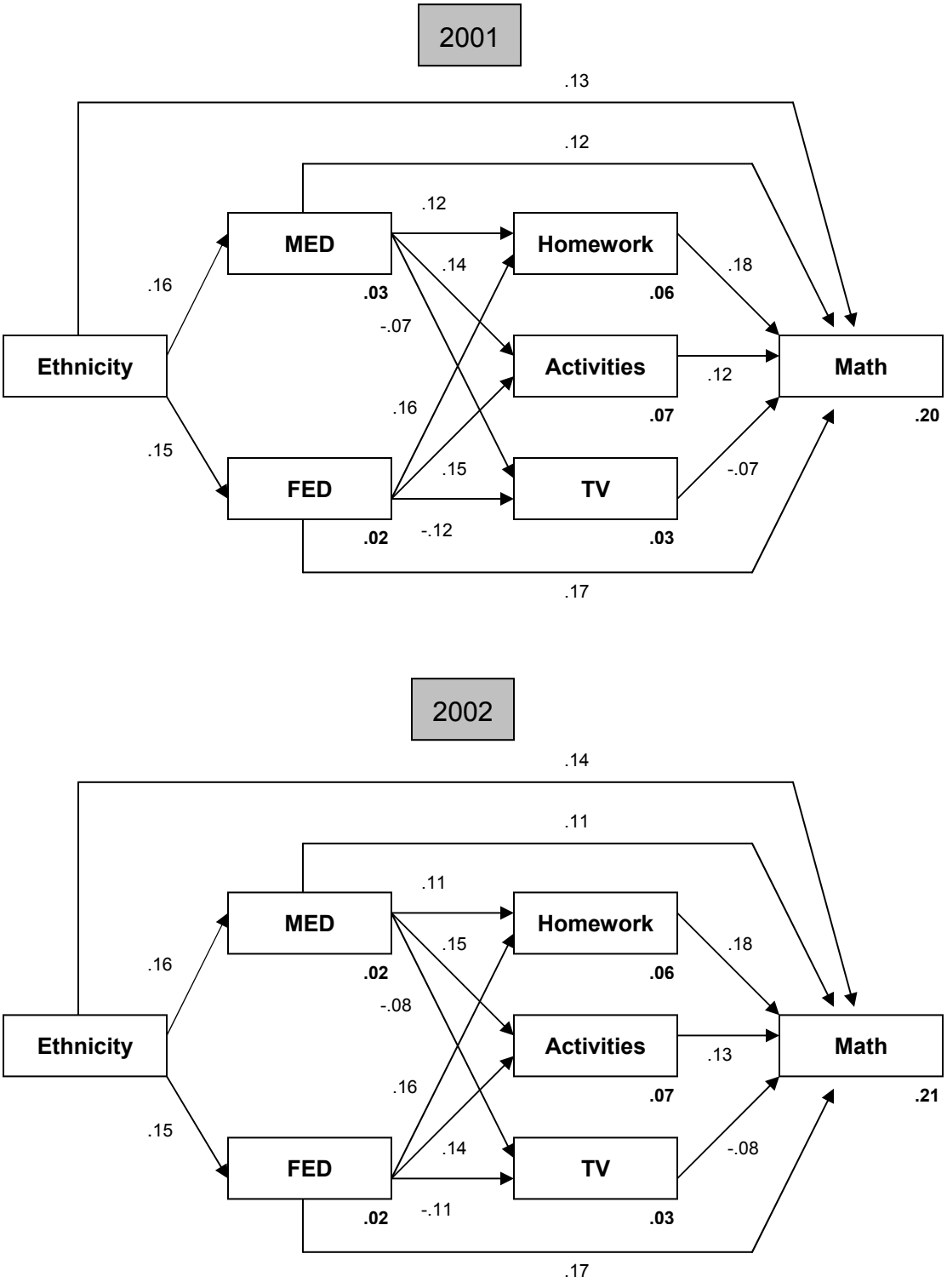
We begin by discussing the initially hypothesized paths (see Figure 1), followed by the three additional paths that were necessary to achieve an adequate fit to the data. As can be seen in the top of Figure 2, Caucasian students reported that their parents had higher levels of education, as indicated by the positive regression coefficients along those paths. Mothers' and Fathers' education was, in turn, positively related to the number of hours of homework and extracurricular activity participation, and negatively related to the number of hours of students spent watching television. Homework and Activities were, in turn, associated with higher scores in reading, whereas TV was associated with lower scores in reading. Thus, each of the hypothesized paths shown in Figure 1 was statistically significant and in the predicted direction. However, as noted, three additional paths were necessary in order to achieve an acceptable level of fit to the data. As can be seen, the three additional paths represented significant direct positive effects of ethnicity, mothers' education, and fathers' education on reading, and the entire set of six predictors explained 22% of the variance in reading. An examination of the 2002 results reveals a nearly identical pattern.

We now turn to the results involving math. As can be seen the right half of Table 2, the initial models predicting math did not fit the data especially well. An examination of the modification indexes suggested freeing the same three additional paths that were freed in the models involving reading. After freeing these three paths, the final models fit the data well. As can be seen in Figure 3, the final models for math were identical to the final models just discussed for reading, with ethnicity predicting parents' education, parents' education predicting the three "students' activities" (homework, extracurricular activities, and tv) and the "student activities" predicting math achievement. As a set, the six predictors explained between 20% and 21% of the variance in math achievement.[5]

[5] We did not include paths among the three "student activities" (homework, extracurricular activities, and tv) as we had no a prior prediction concerning the causal direction of these relationships. However, it is relevant to note that homework and extracurricular activities were positively correlated (r = .20 in each year), while tv was negatively related to both homework (rs = -.14 and -.15 in 2001 and 2002, respectively). These results provide further support for the claim that engagement in extracurricular activities is not necessarily at odds with academic pursuits (in this case, homework).

Figure 3

Final Path Models for Math: Individual Level Analyses



School Level Analyses without Poverty

The individual level analyses just reported provided reasonable support for the proposed model. In an effort to examine the generalizability of those final models to a more macro level, we next conducted school level analyses by combining students' scores within each school to arrive at a score on each variable for each school. Ethnicity at the school level represents the percentage of Caucasian students in a given school, whereas the remaining school level variables represent the average level of a given variable within a given school. After aggregating scores in this manner, we subsequently dropped schools with an A, D, I or S classification (see above) and those schools with less than five students. We then assessed the fit of the final individual level models (i.e., those shown in Figures 2 and 3) at the school level. Consistent with our earlier approach, in the event of poor model fit, we subsequently modified the model by freeing paths that were meaningful and would produce a large improvement in model fit, as indicated by the modification indexes. Fit statistics for the initial and final models are summarized in Table 2, and the final path models for reading and math are presented in Figures 4 and 5, respectively.

As can be seen in Table 3, the initial models for both reading and math did not fit the data well (i.e., both the AGFI and RMSEA failed to reach their respective criteria of .90 and .10 for a reasonable fit). An examination of the modification indexes in each analysis suggested that model fit would be improved by freeing a path from Ethnicity to TV. As summarized in Table 2, freeing this one path improved model fit to a reasonable level. A close inspection of the final models shown in Figures 4 (reading) and 5 (math) reveals that many of the relationships were stronger at the school (vs. the individual) level of analysis. This pattern was especially evident in the case of ethnicity's impact on parents' education and the outcomes of reading and math achievement. As noted, at the school level, ethnicity (% Caucasian) also predicted (fewer) hours of TV watched per day. Several other comparisons between the individual and school level findings deserve brief mention. To begin, parents' education explained a much larger percentage of the variance in homework at the school as opposed to the individual level (38.5% vs. 6%), whereas parents' education explained similar amounts of the variance in extracurricular activities at the two levels of analysis. In addition, at the school level, more of the variance in TV was explained (33% vs. 5%), but this was primarily due to the fact that, at the school level, ethnicity was included as an additional predictor of TV (in addition to

parents' education). Also noteworthy is the finding that much more of the variance in reading and math was explained at the school level (74.5% and 71.5%) than at the individual level (22.5% and 20.5%). Finally, relationships between the three types of activities (homework, activities, and tv) and the two forms of achievement (reading and math) were fairly similar at the individual and group levels of analysis.

Table 3

Summary of Initial and Final Path Models: School Level Analyses, No Poverty

	Reading				Math			
	2001		2002		2001		2002	
Fit Measure	Initial	Final	Initial	Final	Initial	Final	Initial	Final
GFI	0.94	0.98	0.94	0.98	0.94	0.98	0.94	0.98
AGFI	0.70	0.91	0.74	0.88	0.70	0.91	0.74	0.88
CFI	0.94	0.99	0.96	0.98	0.94	0.99	0.96	0.98
RMSEA	0.20	0.10	0.18	0.12	0.20	0.10	0.18	0.12

Note. DF = degrees of freedom. P = probability level. GFI = Goodness of Fit Index. AGFI = Adjusted Goodness of Fit Index. CFI = Comparative Fit Index. RMSEA = Root Mean Squared Error of Approximation.

Figure 4

Final Path Models for Reading: School Level Analyses Without Poverty

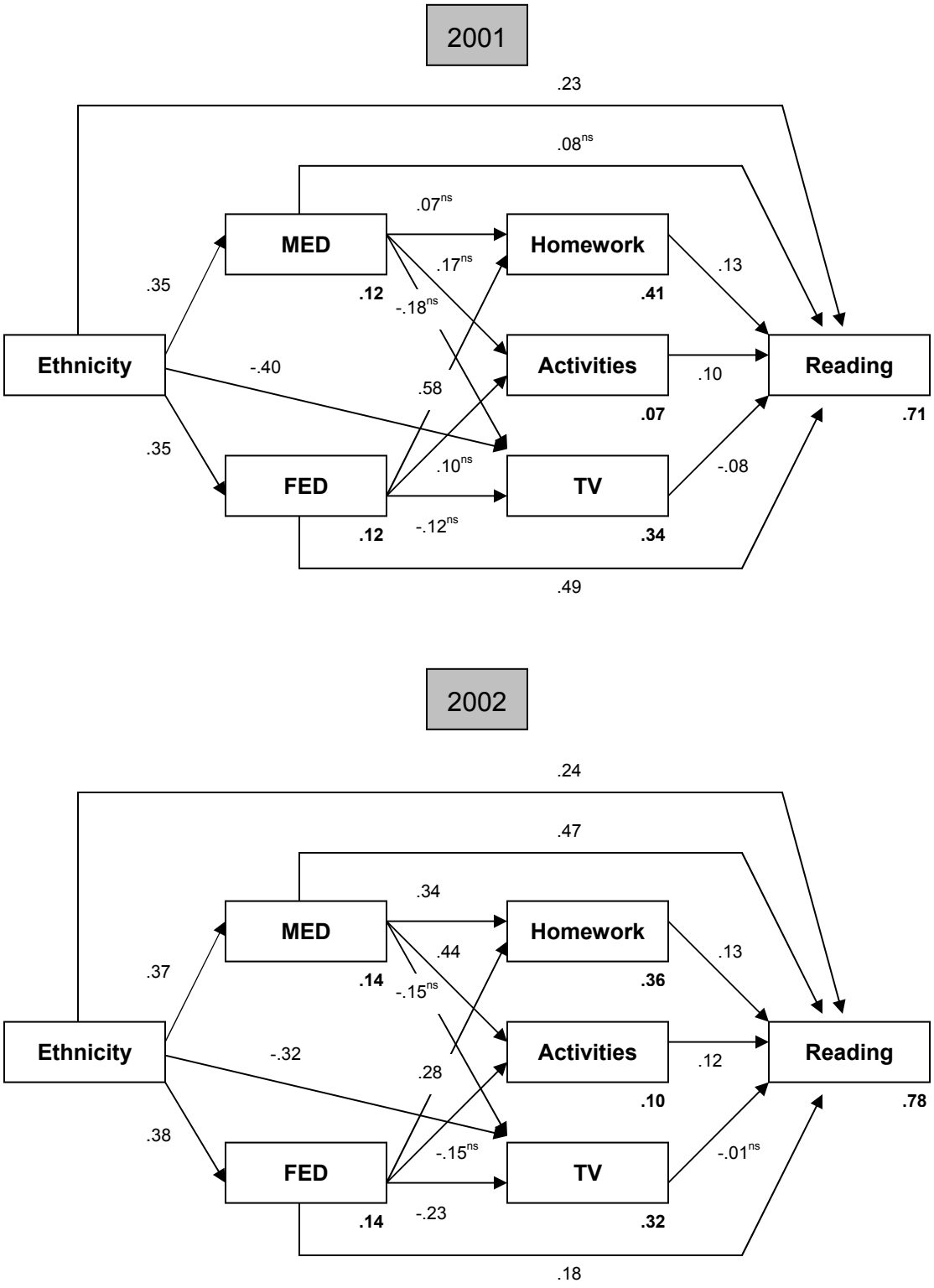
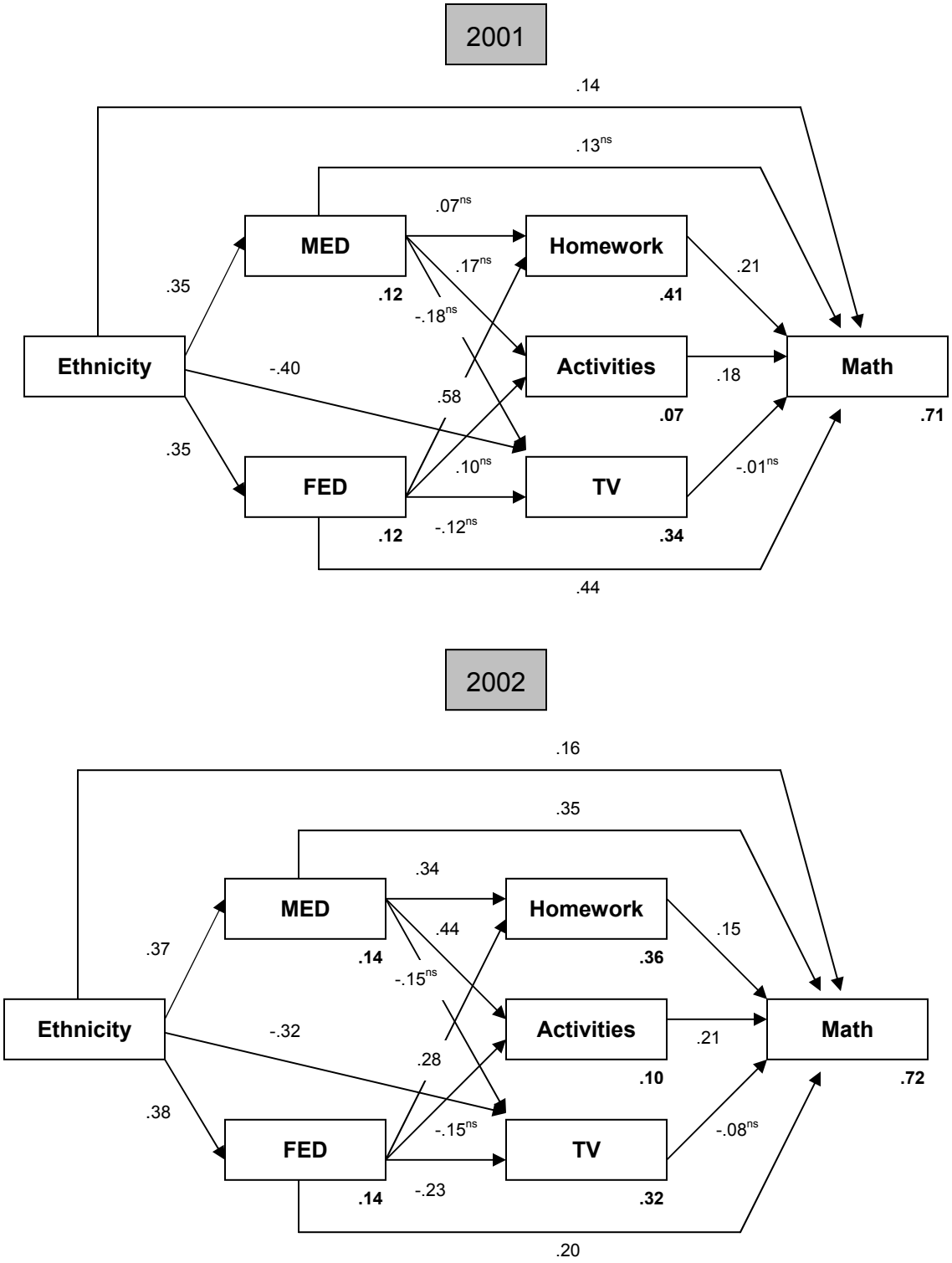


Figure 5

Final Path Models for Math: School Level Analyses Without Poverty

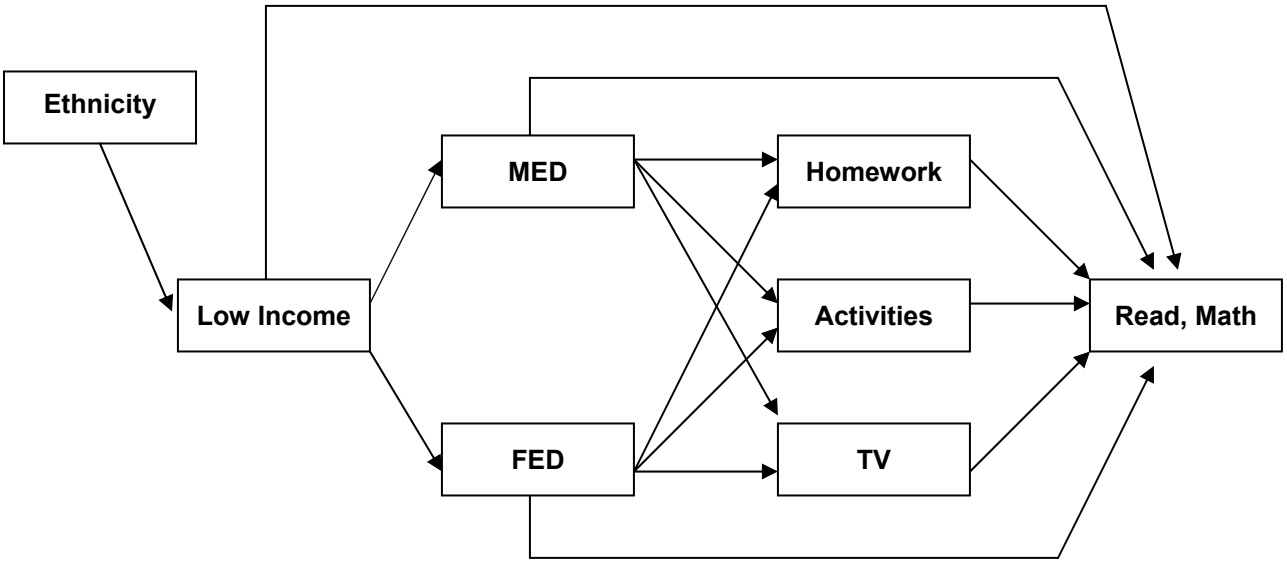


School Level Analyses with Poverty

In a final set of school level analyses, we examined whether low income might serve to carry (i.e., mediate) the influence of ethnicity on the relevant outcomes of interest, as past research has shown that the influence of ethnicity on reading and math is largely explained by low income (Abbott & Joireman, 2001). To accomplish this, we fit an additional set of models treating low income (% of students on free or reduced lunch) as a mediator of the ethnicity effects (see Figure 6). As before, in the event of poor model fit, the initial model was modified by freeing paths that were meaningful and which would substantially improve model fit. Table 4 summarizes the fit indexes for the initial and final models for reading and math.

Figure 6

Initial Model: School Level Analysis With Poverty



As can be seen in Table 4, the initial models did not fit the data especially well. As a result, we modified the initial models in an effort to achieve satisfactory model fit. The final models for reading and math are presented in Figures 7 and 8, respectively. The key finding of interest is whether the direct relationship from ethnicity to reading and math would be necessary to achieve satisfactory model fit. If not, this would provide some support for the idea that low income mediates the ethnicity-achievement relationship. If, by contrast, a direct path from ethnicity to achievement is needed to

achieve satisfactory fit, then the effect of ethnicity is less likely to be completely due to low income. As shown in Figure 7, ethnicity (% Caucasian) was negatively related to low income (i.e., Caucasians reported higher incomes), and low income predicted lower levels of reading achievement, two findings consistent with a mediation model. However, inconsistent with a mediation model, ethnicity continued to have a direct effect on reading, even when low income was included in the model. Thus, low income did not mediate the ethnicity-reading relationship.

Table 4

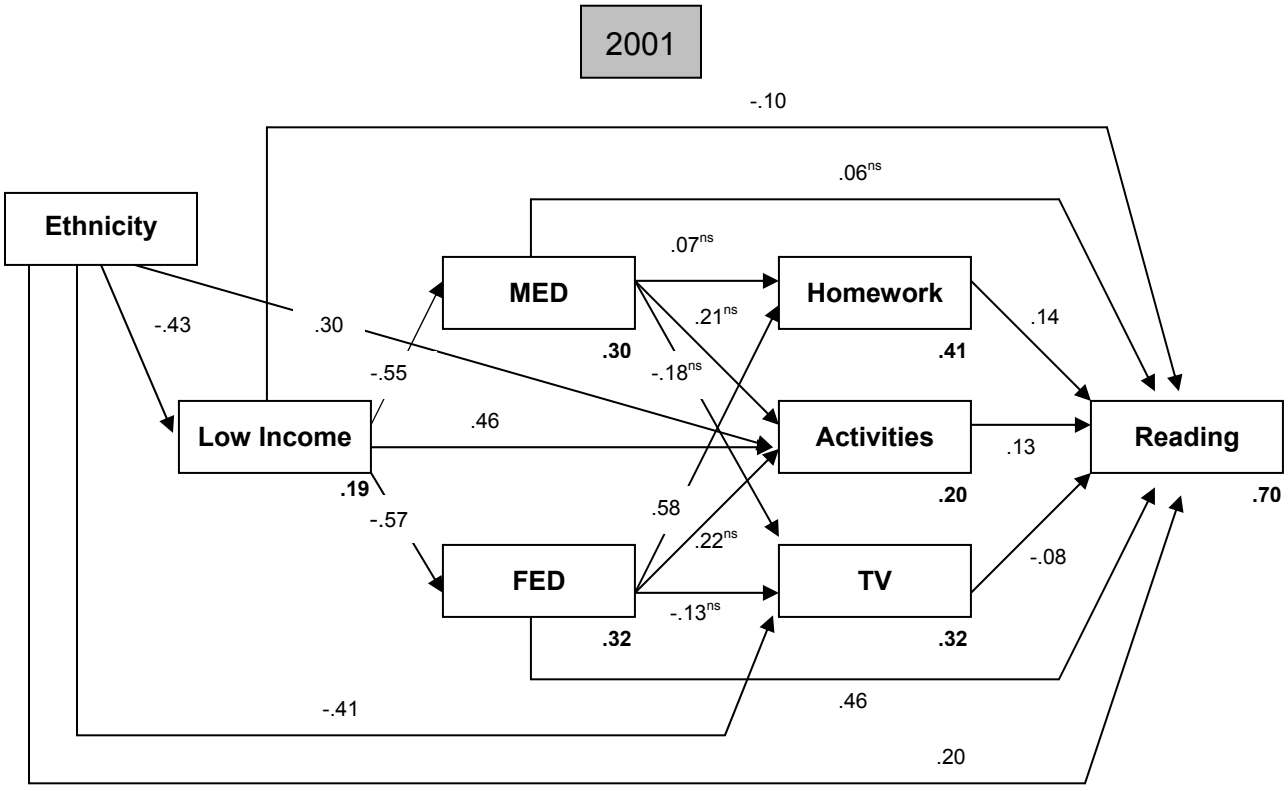
Summary of Initial and Final Path Models: School Level Analyses, With Poverty

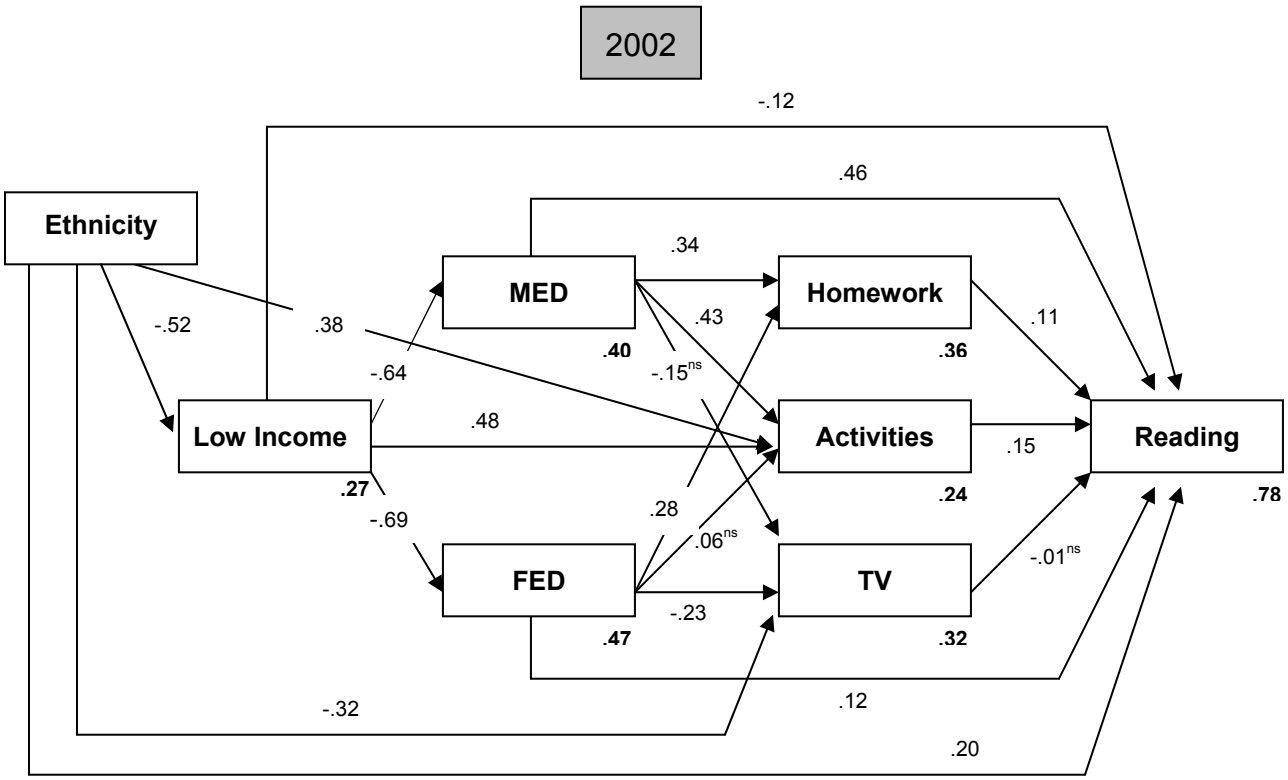
	Reading				Math			
	2001		2002		2001		2002	
Fit Measure	Initial	Final	Initial	Final	Initial	Final	Initial	Final
GFI	0.90	0.99	0.90	0.99	0.90	0.98	0.91	0.98
AGFI	0.69	0.93	0.70	0.93	0.71	0.92	0.73	0.92
CFI	0.90	0.99	0.92	0.99	0.91	0.99	0.93	0.99
RMSEA	0.20	0.07	0.19	0.07	0.19	0.08	0.17	0.08

Note. DF = degrees of freedom. P = probability level. GFI = Goodness of Fit Index. AGFI = Adjusted Goodness of Fit Index. CFI = Comparative Fit Index. RMSEA = Root Mean Squared Error of Approximation.

Figure 7

Final Path Models for Reading: School Level Analyses With Poverty



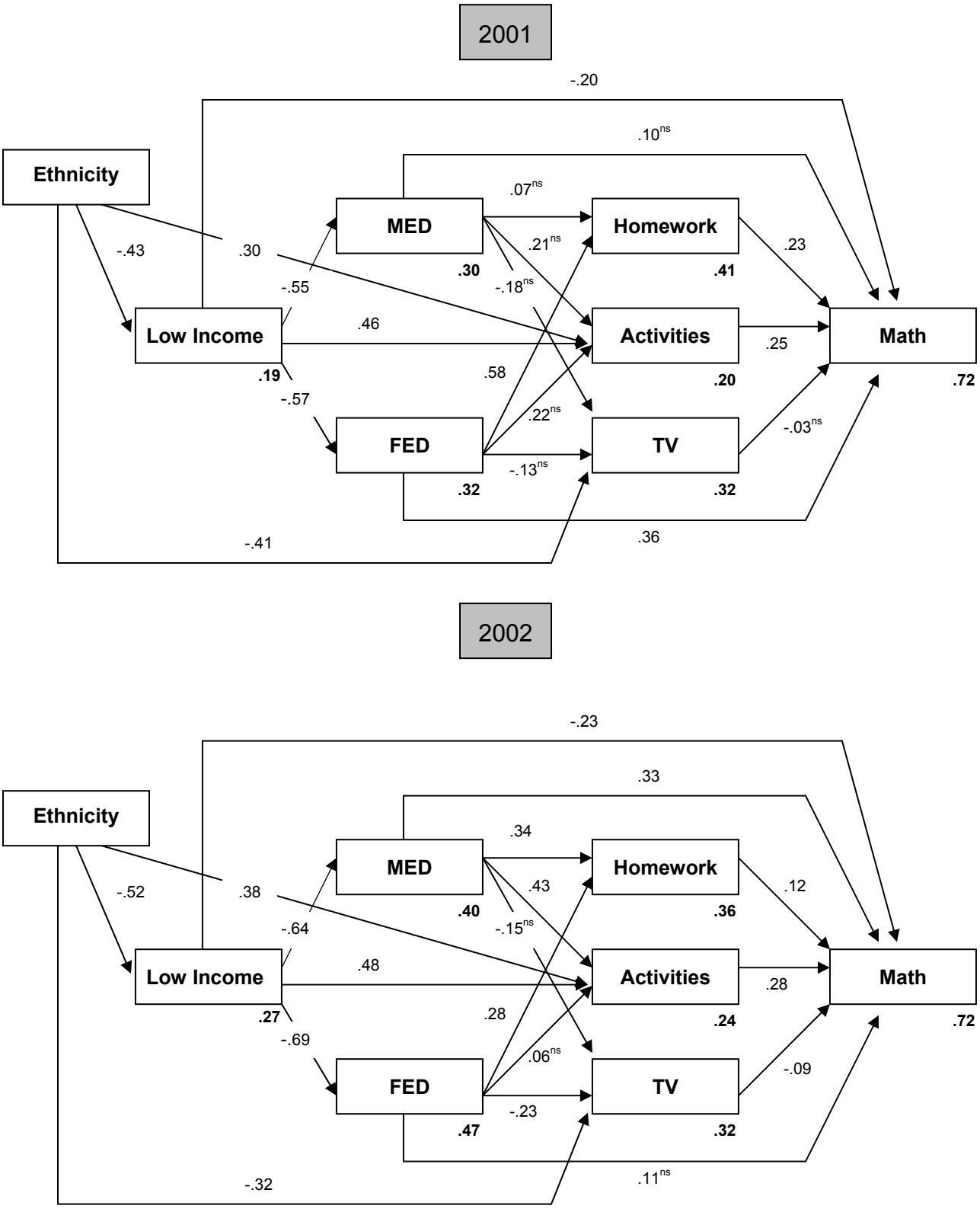


Results provided more support for the mediation hypothesis with regard to math achievement (see Figure 8). As can be seen in Figure 8, a direct path from ethnicity to math was not necessary to achieve satisfactory model fit. This suggests that low income mediated the relationship between ethnicity and math achievement.

One additional finding deserves mention. Specifically, the relationship between activities and math achievement was notably stronger when low income was included in the model (average path coefficient = .265) than when low income was not included in the model (average path coefficient = .195). This indicates that when income is controlled, extracurricular activities become a more significant predictor of math achievement. This pattern was not, however, observed for reading.

Figure 8

Final Path Models for Math: School Level Analyses With Poverty



Discussion

The goal of this project was to examine the relationships among ethnicity, parents' education, student activities (homework, extracurricular, and TV), and reading and math achievement. Our primary question was whether engagement in extracurricular activities conflicted with or complemented reading and math achievement. Results supported the contention that extracurricular activities are positively related to reading and math achievement, suggesting that extracurricular activities and academic achievement are not necessarily at odds. These findings support past research that has demonstrated links between engagement in extracurricular activities and a host of positive outcomes such as better grades, a stronger academic self-concept, better adjustment, and better career outcomes.

In addition to estimating the relationship between extracurricular activities and achievement, we evaluated a more comprehensive causal model in which ethnicity influences parents' education, which in turn influence the three "student activities," which in turn influence reading and math achievement. In general, results provided reasonable support for the proposed model (Figure 1), after several additional paths had been added to the model (direct paths from parents' education and ethnicity to reading and math). School level analyses revealed similar patterns, with the model variables explaining a much larger percentage of the variance in reading and math at this level. Additional school level analyses including low income as a predictor revealed some support for the claim that low income mediates the relationship between ethnicity and math achievement. However, ethnicity continued to exert a significant effect on reading, over and above low income. Thus, the current findings partly confirmed Abbott and Joireman's (2001) earlier findings that low income mediates the relationship between ethnicity and reading and math achievement.

While interesting and informative, the current findings should be interpreted within the context of several inherent limitations. First, the correlational nature of these data prevents firm conclusions regarding the direction of causality, even with the use of structural equation modeling techniques. Thus, longitudinal data are necessary to more clearly identify the causal direction of the relationships reported here. Another limitation of the current study concerns the fact that many of the model variables were based on

single item responses. It is especially worth noting that the current single item measure of extracurricular activities did not allow for a broad sampling of activities, and/or delineation between different types of activities. As such, future attempts to replicate these results should utilize more comprehensive sampling of extracurricular activities. In a similar vein, future studies should examine in greater detail how parents' education plays a role in students' activities, and how extracurricular activities encourage positive academic outcomes (e.g., via a more positive academic self concept; cf. Marsh, 1992).

References

Abbott, M. L., & Joireman, J. (2001). *Relationships among achievement, low income, and ethnicity across six groups of Washington State students. Technical Report 1*. Lynwood, WA: Washington School Research Center, Seattle Pacific University.

Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates.

Camp, W. G. (1990). Participation in student activities and achievement: A covariance structural model. *Journal of Educational Research, 83*, 272-278.

Coleman, J. S. (1959). Academic achievement and the structure of competition. *Harvard Educational Review, 29*, 330-351.

Coleman, J. S. (1961). *The adolescent society*. New York: Free Press of Glencoe.

Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research, 59*, 117-142.

Holland, A., & Andre, T. (1987). Participation in extracurricular activities in secondary school: What is known, what needs to be known? *Review of Educational Research, 57*, 437-466.

Mahoney, J. L. & Cairns, R. B. (1997). Do extracurricular activities protect against early school dropout? *Developmental Psychology, 33*, 241-253.

Marsh, H. W. (1992). Extracurricular activities: Beneficial extension of the traditional curriculum or subversion of academic goals? *Journal of Educational Psychology, 84*, 553-562.