DOCUMENT RESUME

ED 434 929 TM 030 181

AUTHOR	Stone, Clement A.; Lane, Suzanne
TITLE	MSPAP Performance Gains from 1993-98 and Their Relationship to "MSPAP Impact" and School Characteristic Variables.
SPONS AGENCY	Department of Education, Washington, DC.
PUB DATE	1999-04-00
NOTE	19p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Montreal, Quebec, Canada, April 19-23, 1999).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Academic Achievement; *Achievement Gains; Educational Assessment; Educational Change; *Educational Objectives; Educational Practices; Elementary Education; *Institutional Characteristics; Middle Schools; *Performance Based Assessment; Science Education; State Programs; Structural Equation Models; Student Surveys; Teacher Surveys; Testing Programs
IDENTIFIERS	Impact Evaluation; *Maryland School Performance Assessment Program; Reform Efforts

ABSTRACT

A prevailing assumption underlying statewide performance-based assessments is that they not only serve as motivators in improving student achievement and learning, they also encourage instructional strategies and techniques in the classroom that are more consistent with reform-oriented educational outcomes. Given these high expectations, more comprehensive and direct evidence for the consequences of assessments (both negative and positive) need to be addressed. The purpose of this paper is to explore the relationship between changes in the scores from the Maryland School performance Assessment Program (MSPAP) science performance assessment from 1993 to 1998 and classroom instructional and assessment practices, student learning and motivation, students' and teachers' beliefs about and attitude towards the assessment, and finally, student characteristics. Using growth models estimated within a structural equation modeling (SEM) framework, several factors for each of these dimensions were observed to explain a significant amount of the variability in school performance. The paper discusses these factors as well as the design of evaluations that hope to study the impact of assessment programs on students, teachers, and schools. (Contains 3 figures, 3 tables, and 14 references.) (Author/SLD)

MSPAP Performance Gains From 1993-98 and Their Relationship to "MSPAP Impact" and School Characteristic Variables

Clement A. Stone

Suzanne Lane

University of Pittsburgh

Paper presented at the annual meeting of the National Council on Measurement in Education, Montreal, April 1999. Preparation of this paper was supported by a grant from the U.S. Department of Education, Assessment Development and Evaluation Grants Program (CDFA 84.279-A), for the Maryland Assessment Project. Our deepest appreciation is extended to the teachers, principals and students in Maryland for their invaluable time and effort spent on this project.

Abstract

A number of states are implementing statewide performance assessment programs that are being used for high-stakes purposes such as holding schools accountable to state standards. A prevailing assumption underlying statewide performance-based assessments is that they not only serve as motivators in improving student achievement and learning, they encourage instructional strategies and techniques in the classroom that are more consistent with reform-oriented educational outcomes (e.g., instruction focusing on reasoning and communication skills). Given these high expectations, more comprehensive and direct evidence for the consequences of the assessments (both negative and positive) need to be addressed. The purpose of this paper is to explore the relationship between changes in the scores from MSPAP's science performance assessment from 1993 to 1998 and classroom instructional and assessment practices, student learning and motivation, students' and teachers' beliefs about and attitude towards the assessment, and finally, school characteristics. Using growth models estimated within a structural equation modeling (SEM) framework, several factors from each of these dimensions were observed to explain a significant amount of the variability in school performance. The paper discusses these factors as well as the design of evaluations that hope to study the impact of assessment programs on students, teachers, and schools.

MSPAP Performance Gains From 1993-98 and their Relationship to "MSPAP Impact" and School Characteristic Variables

The Maryland State Performance Assessment Program (MSPAP) is a performance assessment program designed to measure school performance for grades 3, 5, and 8 and provide information for school accountability and improvement (Maryland State Board of Education, 1995). Implemented in the early 1990's, MSPAP requires students to develop written responses to interdisciplinary tasks that require the application of skills and knowledge to real life problems, and is intended to promote performance-based instruction and classroom assessments. The purpose of this paper is to explore the relationship between changes in MSPAP test scores and classroom instructional and assessment practices, student learning and motivation, professional development, students' and teachers' beliefs about and attitude towards MSPAP, and finally, school characteristics. Although information was collected in a variety of content areas (science, language arts, math, and social studies), the focus of this paper is on the science content area. Across the number of schools that were assessed, 116 schools provided both teacher and student information and formed the basis for the analyses presented herein. The language arts data was based on a different and smaller set of schools which prohibited analyses similar to those conducted on the science data. Results involving the math content area were presented previously (Lane, Parke, & Stone, 1998); subsequent analyses are planned for the social studies data.

Modeling Differences in School Performance Over Time

Random coefficient or growth models were used to examine science performance on MSPAP from 1993 to 1998 in relation to variables derived from the teacher and student questionnaires, and the school characteristic, percent free or reduced lunch which served as a proxy for socioeconomic status. The advantages of using growth curve methodologies to analyze change has been discussed in the literature (c.f., Rogosa & Willet, 1985; Willet & Sayer, 1994; Rogosa, 1987). These methodologies are particularly well suited for studying processes that consider change as continuous with individual differences in the pattern of change (e.g., initial level and rate of change). Further, these methodologies allow for studying individual differences and identifying factors that affect the trajectory of change. This type of analysis can not be modeled by time-specific comparisons involving group-level (e.g., means) differences.

Variables from questionnaires administered to teachers and students from the schools in the sample were hypothesized to explain individual differences in school performance over time. A subset of variables from the questionnaires was used because of the relatively small school sample size. In addition, the dimensions that were used were considered to be more relevant than the other dimensions for examining the relationship between change and teachers' perceptions. From the teacher questionnaire, two dimensions were examined: MSPAP Impact and Current Science Instruction. These two dimensions were derived from subsets of items that asked questions about the direct impact MSPAP has had on classroom activities, and questions related to the extent to which classroom activities focus on the science learning outcomes and reform-oriented problem types. From the student questionnaire, the Current Instruction dimension and two Likert-scaled items were analyzed: 1) In science class this year, how often did you work on tasks like those on MSPAP? And, 2) How important is it for you to do well on MSPAP? The Current Instruction dimension was similar to the teacher-level dimension, that is, it was derived from questions about the type of instructional activities engaged in, but from the student's perspective.

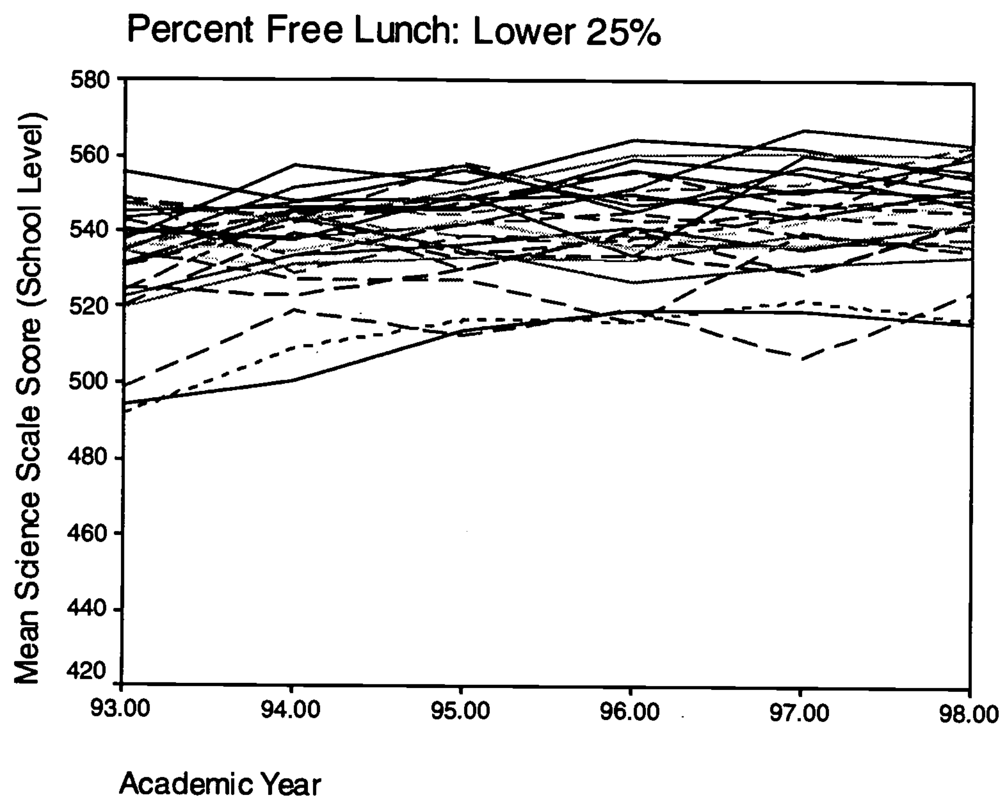
Figure 1 illustrates the differences in initial mean MSPAP performance and changes in mean MSPAP performance from 1993 to 1998 for the sample of schools in the present study. Since percent free or reduced lunch was found to correlate significantly with MSPAP performance, the plots are presented for two subgroups of this variable (i.e., lower and upper quartiles) to reduce the number of lines in any one graph. As can be seen, there are differences among the schools in terms of their initial MSPAP science performance and their change over time. Schools in the lower quartile (Higher SES) were concentrated in the range of 520-550 in 1993 whereas schools in the upper quartile (Lower SES) were concentrated in the range of 480-500 in 1993. In addition, the rate of change for schools in the lower quartile exhibited a more consistent increase over time whereas considerably more variability was observed for schools in the upper quartile. In both cases, the rate of change appears modest from 1993 to 1998.

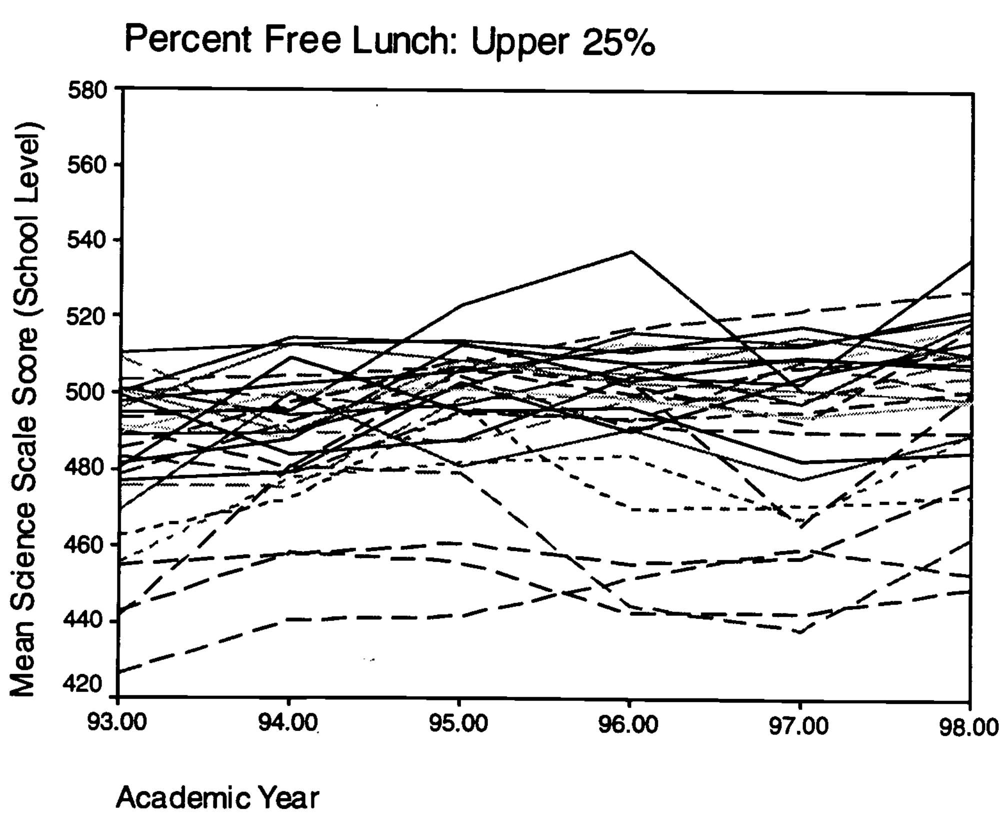
Table 1 summarizes the mean performance across the set of schools. As can be seen, mean performance is increasing over time if the 1995 time-point is disregarded. In addition, mean scores appear to level off from 1995 to 1997 at which point there is an increase in performance. From the graphs and the figure, non-linearity in performance changes over time is apparent.

Table 1: Means and Standard Deviations of MSPAP Science Scores

	N	Mean	Std. Deviation
Science 93	116	509.6	25.3
Science 94	116	514.6	23.0
Science 95	116	519.0	21.4
Science 96	116	518.3	23.8
Science 97	116	518.9	24.9
Science 98	116	523.6	22.9

Figure 1: Change in Mean MSPAP Science Scores Over Time by Percent Free Lunch Percentiles





The models used to capitalize on the information contained in multiwave data appear in the literature under a variety of labels, including random-effects models or random coefficient models (e.g., Laird & Ware, 1982) and hierarchical linear models (Bryk & Raudenbush, 1992). In order to model individual differences in change and assess the correlates or predictors of change, two levels of statistical modeling are required: Level 1 - within individual schools, trends across the repeated measurements are modeled; and Level 2 - across schools, the parameters from the model of individual differences in change at Level 1 are modeled in relation to other factors. At Level 1, growth models analyze the repeated measurements of test scores, analyze the relationship between time (year) and test score levels, and estimate a reference status (intercept) and rate of change (slope) for each school. It would be expected that schools would differ with regard to their initial levels MSPAP performance (measured at time 1), their rates of change over time, and the shape or pattern of change (e.g., linear, nonlinear).

A linear growth model with a single outcome variable y measured for each school at each timepoint is:

$$y_{it} = \alpha_i + \beta_i x_{it} + \varepsilon_{it}, \qquad (1)$$

where α_i is an intercept parameter for each ith school, x_{it} is the time-related variable for the i^{th} school at time t, β_i is a slope parameter reflecting the linear rate of change over time for the i^{th} school, and ε_{it} is a residual reflecting both random measurement error and unspecified time-specific effects.

The parameters from the model at Level 1 (intercepts and slopes) are then modeled in relation to factors that are introduced to explain variation in the parameters across schools (Level 2). For example, the school-specific parameters, α_i and β_i from the Level 1 model, are incorporated into the Level 2 model with one school-specific explanatory variable (z_i) as follows:

$$\begin{aligned} \alpha_i &= \mu_\alpha + \gamma_\alpha z_i + \varepsilon_{\alpha i} \\ \beta_i &= \mu_\beta + \gamma_\beta z_i + \varepsilon_{\beta i} \end{aligned} \qquad (2)$$

where μ_α and μ_β are parameters reflecting group-level means of the intercepts and slopes, respectively, and the variance of these factors reflects the individual differences or random effects that exist around these group level parameters. Larger variances reflect increased variability (less similar patterns) in intercepts and slopes; z_i is a time-invariant covariate introduced to explain variation in these parameters (e.g., SES level); γ_α and γ_β are regression parameters reflecting the effects of the covariate on the Level 1 intercept and slope parameters; and, $\varepsilon_{\alpha i}$ and $\varepsilon_{\beta i}$ are residual terms. It is assumed that the ε_{it} are uncorrelated with $\varepsilon_{\alpha i}$ and $\varepsilon_{\beta i}$, but $\varepsilon_{\alpha i}$ and $\varepsilon_{\beta i}$ may be correlated. It should be noted that it is straightforward to increase the number of explanatory variables in the Level 2 model and consider time-varying covariates as well as non-linear growth rates in the Level 1 model. In the present study, various dimensions from the teacher and student questionnaires, and the variable percent free or reduced lunch were introduced to explain variation in the intercepts and slopes.

Growth models can be estimated using a variety of software. Recently, Singer (1999) illustrated the estimation of such models in SAS PROC MIXED. Specialized software is also available (e.g., HLM: Bryk & Raudenbush, 1992). In addition, several researchers have discussed how growth models can be estimated within a structural equation modeling (SEM) framework by considering the intercept and slope factors as latent variables (e.g., McArdle & Epstein, 1987; Meredith & Tisak, 1990; Muthen, 1991; Willet & Sayer, 1994). Muthen and Curen (1997) have further discussed the flexibility in modeling that is afforded by estimating growth models using SEM. In the present study, the growth models were estimated using the SEM program AMOS (Arbuckle, 1997).

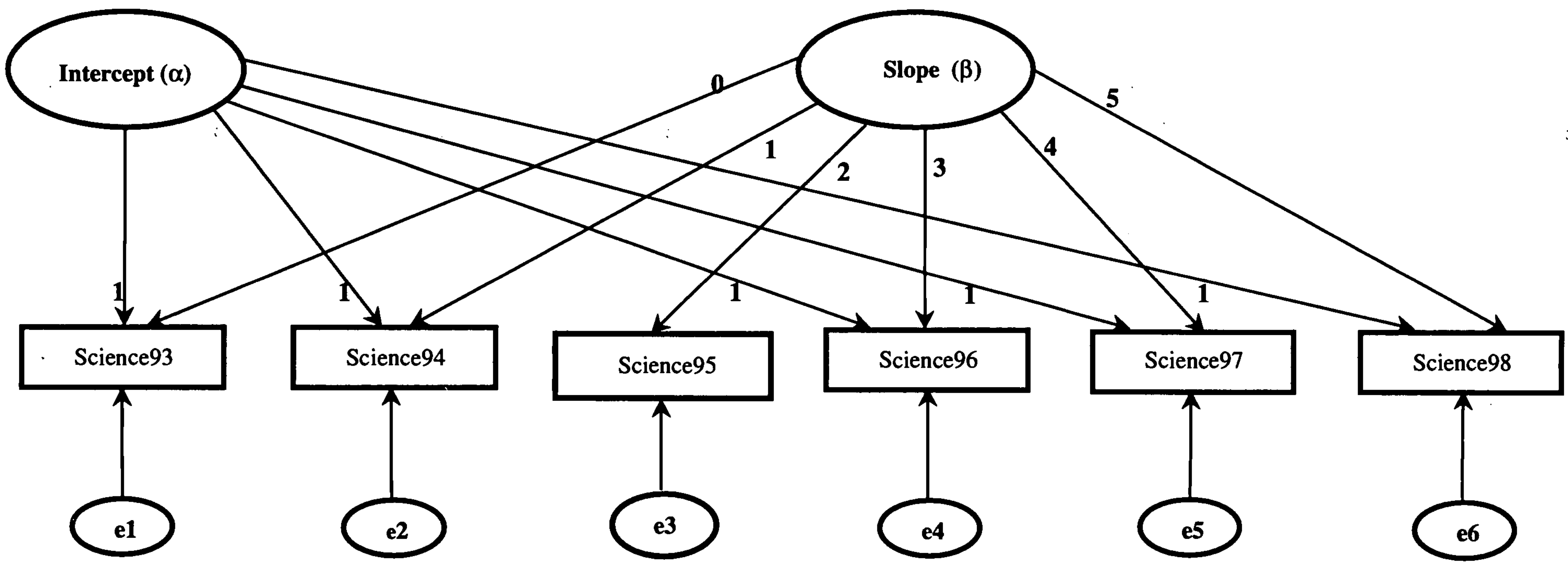
Figure 2 presents a Level 1 (Unconditional) growth model for the present study. This model involves the outcome variable, MSPAP science standard score, measured at six timepoints. In order to translate the growth model into the framework of structural equation modeling, the school-specific random coefficients (intercepts and slopes from Level 1) are each modeled using two latent factors: 1) a factor representing a reference status of MSPAP performance (intercept or α), and 2) a factor which corresponds to the rate of change in MSPAP performance over time (slope or β). The mean of these factors represent group level estimates (Level 2) of the intercepts and slopes, respectively, and the variance of these factors reflects the school differences or random effects that exist around these group level parameters. Larger variances reflect increased variability or less similarity in intercept and slopes among the schools.

As can be seen from the figure, the Level 1 model has the format of a measurement or confirmatory factor analysis model in SEM with restrictive loadings: $\mathbf{Y} = \Lambda\eta + \varepsilon$

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \\ Y_{i3} \\ Y_{i4} \\ Y_{i5} \\ Y_{i6} \end{pmatrix} = \begin{pmatrix} 1 & x_t \\ 1 & x_t \\ 1 & x_t \\ 1 & x_t \\ 1 & x_t \\ 1 & x_t \end{pmatrix} * \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{pmatrix} \tag{3}$$

where $\mathbf{Y}$ is a vector of original measurements over time, η is a vector of latent variables (intercept and slope parameters), Λ is a matrix of regression coefficients relating the slope and intercept factors to the Y measurements, and ε is a vector of residuals representing variance not accounted for due to time specific factors not included in the model or random error. In addition, an association between the intercept and slope factors may be specified and indicated through a curved bi-directional arrow in the figure. Note that, in order to specify these models in SEM, it is necessary to assume that $x_{it} = x_t$, which means that all individuals are measured at the same point in time at each time-point. In this study as well as other state-wide testing situations, tests are typically administered at the same time.

Figure 2. Level 1 Unconditional Growth Model



The regression coefficients relating the intercept factor to the measurements are fixed at 1 since the intercepts reflect a constant contribution to the measurements over time. The scaling of the slope factor is determined by the pattern in the x_t coefficients that relate the time variable to the observed measurements. To reflect a simple linear growth pattern with one unit of change between time points the coefficients (x_t) would be set to 0, 1, 2, 3, 4, and 5. Note that in the framework of SEM, it is possible to freely estimate coefficients or constrain parameters to any other specified pattern. Thus, there is no constraint that time points be equally spaced or that all x_t be specified.

The meaning of the intercept factor depends on the scaling of the time variable for the slope factor, and the scaling of the slope factor is determined by the factor loadings or regression coefficients relating the slope factor to the observed measurements. Under the scaling in Figure 2, the intercept could be interpreted as MSPAP initial status of schools since time 0 corresponds to 1993 performance. However, it is also possible to estimate coefficients or constrain the parameters to some other pattern. In this study, the pattern adopted was 5, 4, 3, 2, 1, and 0. Since time 0 is associated with 1998 MSPAP performance, the intercept factor is interpreted as 1998 MSPAP status and a decrease in performance would be expected from 1998 to 1993. This scaling was adopted because other school related information was collected in 1998 and introduced into the analysis to explain variations in the 1998 MSPAP performance and rates of change among schools. The intercept factor will be referred to as 1998 MSPAP performance hereafter.

The structure or distribution of the residuals (Level 1 error models) is defined through constraints on the parameters of the error variance-covariance matrix. The classical assumption of homoscedastic independent errors can be defined by constraining the diagonal elements (variances) of the error variance covariance matrix to be equal over time and off-diagonal elements (covariances) fixed at 0. This assumption can be relaxed by allowing the variances to vary over time and/or estimating a certain pattern to the error variances and covariances (e.g., compound symmetry or adjacent error covariances estimated). In addition, all error variances and covariances can be estimated as in a fully parameterized or unstructured error matrix. In Figure 2, independent but unequal error variances are assumed.

In order to estimate group level estimates of the intercept and slope latent variables for the Level 2 model, means for the latent variable intercepts and slope factors must be estimated. The general covariance structure model accommodates such a parameterization and is often used when analyzing longitudinal data or multiple populations. In order to estimate these types of models, the general covariance structure model includes an intercept term as follows: $\mathbf{Y} = \tau + \Lambda\eta + \varepsilon$, where τ is a vector of intercepts and is the E[Y] when $\eta = 0$, and all other model parameters are defined as before. Note that $\tau = 0$ when deviations from means are analyzed.

Table 2 presents results from estimating the Level 1 model for the 116 schools. The chi-square statistic for model-data-fit was 13.9 with 7 *df* (p=.053) indicating that the null hypothesis that the variance-covariance matrix implied by the model equals the observed variance-covariance matrix could not be rejected. It should be noted that the 1995 time-point, which represented an anomaly in Table 1, was deleted from the analysis in order to attain an acceptable model-data-fit. As can be seen, the 1998 MSPAP performance (intercept factor) across the schools was 523.6 with a significant mean rate of change (slope factor) over time of –2.77, although the rate of change was modest given the scale of the test scores. Recall that the rate of change is associated with a decrease in performance from 1998 to 1993. Thus, this result suggests that there was a significant increase in performance from 1993 to 1998. Also, the results in the table indicate that a non-linear rate of change was estimated in the model. The chi-square difference between a model assuming linear change and the non-linear rate of change model was significant (χ^2 difference equal to 8.12 with 2 *df*; p<.01) and is consistent with the results in Table 1. A larger than average change was apparent between 1993 and 1994 (estimated coefficient of 3.3 versus a fixed coefficient of 4), followed by a leveling off, and then a larger than average change between 1997 and 1998 (estimated coefficient of 1.8 versus a fixed coefficient of 1).

The variances for 1998 MSPAP performance and rate of change indicate significant variability in these parameters across the school. However, the covariance between 1998 MSPAP performance and rate of change was not significant and was thus fixed at 0. In order to investigate this last finding further, an analysis in which 1993 MSPAP performance was the reference point was examined. This analysis revealed a significant negative covariance between 1993 MSPAP performance and rate of change (r = -.46). This indicated that higher rates of change were associated with lower initial performance in 1993. This suggests that the rate of change is more similar for schools in 1998 than in 1993 which may be due to the observed decrease in variability in 1998 school performance as compared to 1993. Finally, note that although a fully unstructured error model was not required, two covariances between errors for the 1997 and 1998 MSPAP scores were significant and required estimation.

These results are very consistent with the previously presented results for the math content area (Lane, Parke, and Stone, 1998). A similar modest (-2.70) but significant and non-linearity rate of change was observed. In addition, the variances in the intercepts and slopes were significant and a similiar significant correlation between 1993 MSPAP performance and rate of change variable was observed (r = -.40).

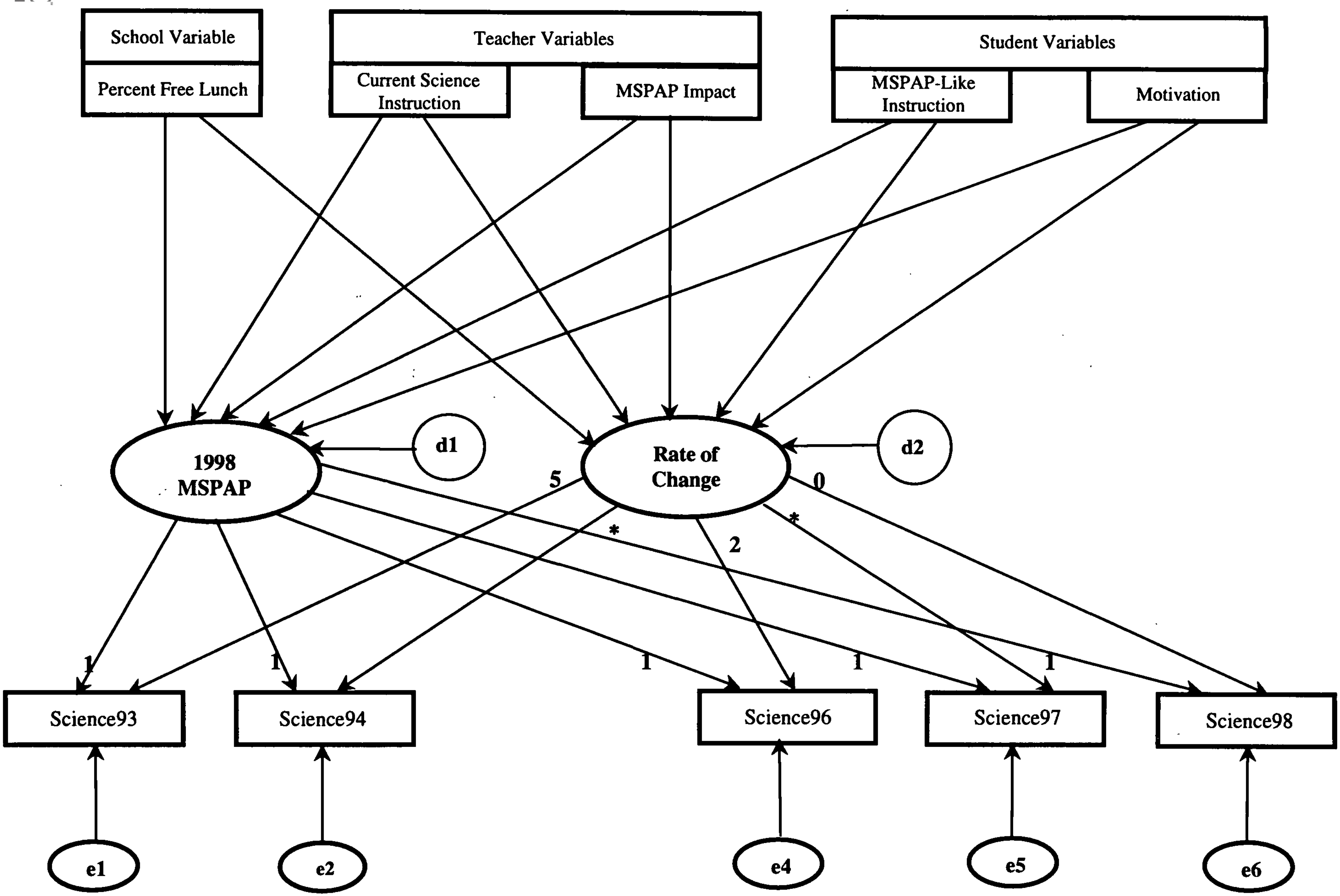
The structural component of the structural equation model is used to reflect factors which are hypothesized to explain variability in 1998 MSPAP performance (intercepts) and rates of change (slopes): $\eta = \alpha + \beta\eta + \zeta$; where, η is defined as above, α is a vector of population means for the latent variables, β is a matrix of structural slopes for the effects among endogenous and exogenous η variables (e.g., variables included to explain variablility in intercepts and slopes), and ζ are structural residuals.

Table 2: Results for the Level 1 Growth Model

Measure and variable	Estimates	SE	T
Regression Coefficients:			
Science93← 1998 Performance	1		
Science94← 1998 Performance	1		
Science96← 1998 Performance	1		
Science97← 1998 Performance	1		
Science98← 1998 Performance	1		
Science93← Rate of Change	5		
Science94← Rate of Change	3.34	.31	10.78
Science96← Rate of Change	2		
Science97← Rate of Change	1.80	.25	7.26
Science98← Rate of Change	0		
Latent Variable Means:			
1998 Performance	523.6	2.10	248.85
Rate of Change	-2.77	.24	-11.50
Variances/Covariances:			
1998 Perform, Rate of Change	0		
1998 Performance	473.95	65.68	7.22
Rate of Change	3.18	1.15	2.77
e1	44.24	15.19	2.91
e2	75.45	12.61	5.99
e4	52.39	9.76	5.37
e5	73.90	13.81	5.35
e6	45.17	14.34	3.15
e4, e5	25.13	10.41	2.41
e5, e6	16.65	7.81	2.13

Figure 3 presents a Level 2 (Conditional) growth model for the present study. A school variable (percent free lunch), and a limited number of variables from the teacher and student questionnaires were introduced into the growth model to explain variability in 1998 MSPAP performance and rate of change across the schools. The structural residuals are specified by the latent variables d1 and d2 in the figure, and the relationship between 1998 MSPAP performance and rate of change is estimated through these two residual parameters. Note that, in theory, it would be possible to incorporate the confirmatory factor analysis model for the questionnaires directly within the growth model rather than use the derived variables. However, given the sample size in the present study, such a model was overly complex to be estimated.

Figure 3. Level 2 Growth Model with School Level Covariates



14

Table 3 presents the unstandardized regression coefficients for the variables introduced to explain variation in 1998 MSPAP performance and changes in performance over time. The chi-square statistic for model-data-fit was 37.13 with 26 *df* ($p=.07$) indicating that the null hypothesis that the variance-covariance matrix implied by the model equals the observed variance-covariance matrix could not be rejected. Other measures of model-data-fit included: RMSEA statistic = .06 which is within the acceptable range (Browne and Cudeck, 1993) and NFI was .99. Note that for any effect that was not significant or borderline significant, the parameter for the effect was fixed at 0.

Table 3 Results for the Level 2 Growth Model with School Level Covariates

Measure and Variable	Estimates	SE	t
Regression Coefficients			
Effects on 1998 Performance:			
Percent Free Lunch	-.79	.05	-15.19
Current Science Instruction	5.85	3.07	1.91
MSPAP Impact	0		
MSPAP-Like Instruction	-4.04	1.58	-2.55
Student Motivation	0		
Effects on Rate of Change:			
Percent Free Lunch	0		
Current Science Instruction	0		
MSPAP Impact	-1.07	.40	-2.68
MSPAP-Like Instruction	0		
Student Motivation	-2.09	.63	-3.33
Variances:			
1998 Performance	97.91	17.25	5.68
Rate of Change	.97	.81	5.68

As can be seen, the variable Percent Free Lunch is significantly related to 1998 MSPAP performance. Thus, increases in the percentage of students receiving free or reduced lunch is associated with lower levels of MSPAP performance in 1998. The regression coefficients can be interpreted as any unstandardized regression coefficient. For example, in the case of the Percent Free Lunch variable, one unit change in this variable corresponds to a decrease of .79 units in 1998 MSPAP science scores. Other variables explaining a significant or borderline significant amount of the variability in 1998 science scores included the Current Science Instruction described by teachers and the Students indication of how often they worked on MSPAP-like tasks in class. Note that Current Science Instruction as described by students is not represented in the analysis. Both Current Science Instruction variables (teacher and student perceptions) predicted significantly 1998 MSPAP science performance when included separately in the model. However, when they were included simultaneously, the effects were attenuated.

Therefore, the student level variable was excluded since it was not as inclusive with regard to the classroom instructional and assessment activities.

From the table, there is an apparent paradox between the direction of the relationship for Current Science Instruction and student's perception of MSPAP-Like Instruction. With regard to the Current Science Instruction, as teachers' instruction more closely reflected the Maryland Learning Outcomes and reform-oriented problem types, higher 1998 MSPAP performance was observed. On the other hand, students' perceptions of the degree to which they worked on MSPAP-like tasks were negatively related with 1998 MSPAP performance. Given the question "...how often did you work on tasks like those on MSPAP", students may have been focusing on the format of MSPAP tasks and not on the learning outcomes reflected in the tasks. Thus, this may reflect a greater likelihood of schools with lower performance using more MSPAP-like formatted tasks than schools performing at higher levels. Schools performing at higher levels may be more successful at reflecting the science learning outcomes in a variety of reform-oriented problem formats.

Two factors were also found to significantly explain variability in rates of change: MSPAP Impact and the students' motivational level (How important is it for you to do well on MSPAP?). This indicates that higher levels of teacher reports of MSPAP having a direct impact on instruction are associated with greater rates of decrease in performance from 1998 to 1993 (or higher levels of rate of change in MSPAP school performance from 1993 to 1998). In addition, greater levels in student motivation are associated with greater rates of increase from 1993 to 1998. Finally, it is interesting to note that, although increases in the percentage of students receiving free lunch is associated with lower levels of MSPAP performance in 1998, corresponding increases were not significantly associated with rate of change in MSPAP performance over time.

Finally, the variances in the table, in comparison with those in Table 2, can be used to determine how much variability is explained by the factors. With regard to the variability in 1998 Science performance, approximately 80% of the variance is accounted for by the three variables (1 - 97.91/473.95). With regard to the variability in the rates of change, approximately 70% of the variance is accounted for by the two variables (1 - .97/3.18).

It is important to note that the school sample size for this analysis was relatively small (n=116), and therefore, the results should be interpreted cautiously and additional studies should be conducted. However, it is interesting to note that very similar pattern in the findings were observed for the math content area (Lane et. al.,1998). Although student level information was unavailable, the Percent Free Lunch variable had a similarly significant negative effect (-.78) on math performance. For the Current Math instruction factor, although not significant or borderline significant as in the present case, the regression coefficient was similar in magnitude (6.9). Since the sample was smaller (n=86), the difference in the significance of the findings could be due to lack of power. With regard to the factors introduced to explain variability in rates of change, the MSPAP impact variable was also found to significantly explain variability (coefficient = -1.58) for changes in math MSPAP performance over time.

Discussion

The purpose of this paper was to explore the relationship between changes in MSPAP science scores from 1993 to 1998 and classroom instructional and assessment practices, student learning and motivation, students' and teachers' beliefs about and attitude towards MSPAP, and finally, school characteristics. Several factors from each of these dimensions were observed to explain a significant amount of the variability in 1998 performance of schools and rates of change over time. Thus, there is some correlational evidence for the impact of the assessment program on classroom instructional and assessment practices. As noted above, the results should be interpreted cautiously since the sample was relatively small although cross validation of the results in other content areas provides some degree of generalizability of the findings. Further, it should be emphasized that the cross validation of results in the math content area involves a different set of schools.

In addition to increasing the sample size, the design of such a validity study could be improved by measuring the outcomes in the present study concurrently with assessment performance over time. Thus, changes in classroom instructional and assessment practices, student learning and motivation, professional development, students' and teachers' beliefs about and attitude toward the assessment program could be examined in connection with changes in assessment performance over time. Although school characteristics may or may not change appreciably over time, these could be measured at one time-point or considered to be constant. One of the advantages of estimating growth curve models in a SEM framework is that more general analyses can be conducted, such as models with multiple outcome variables with different growth processes.

Finally, the present study could be improved by examining the growth processes in a three level model. The present study involved a two level model – the unit of analysis involved measurements at the school level and variability in the schools was examined. In a three level model, measurements at the class level provide the repeated measurements at Level 1, variation in classes within schools is modeled at Level 2, and finally, variation among schools is modeled in Level 3. It would be expected that teachers would vary within schools and variables could be introduced to explain differences between teachers as well as variables introduced to explain variation in schools.

References

Arbuckle, J.L. (1997). AMOS User's Guide Version 3.6. Chicago: SmallWaters Corporation.

Bryk, A.S., & Raudenbush, S.W. (1992). Hierarchical linear models: Applications and data analysis. Beverly Hills, CA: Sage.

Browne, M.W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In Bollen, K.A. & Long, J.S. (Eds.). Testing structural equation models. Newbury Park, California: Sage, 136-162.

Laird, N.M. & Ware, J.H. (1982). Random-effects models for longitudinal data. Biometrics, 38, 963-974.

Lane, S., Parke, C.S., & Stone, C.A. (1998). Consequences of the Maryland school performance assessment program. Paper presented at the annual meeting of the National Council on Measurement in Education, San Diego, CA.

Maryland State Board of Education (1995). Maryland school performance report: State and school systems. Baltimore, MD.

McArdle, J.J. & Epstein, D. (1987). Latent growth curves within developmental structural equation models. Child Development, 58, 110-133.

Meredith, W. & Tisak, J. (1990). Latent curve analysis, Psychometrika, 55, 107-122.

Muthen, B.O. (1991). Analysis of longitudinal data using latent variable models with varying parameters. In L. Collins & J. Horn (Eds), Best methods for the analysis of change. Recent advances, unanswered questions, future directions (pp. 1-17). Washington, D.C.: American Psychological Association.

Muthen, B.O. & Curran, P.J. (1997). General growth modeling in experimental designs: A latent variable framework for analysis and power estimation. Psychological Methods, 2, 371-402.

Rogosa, D.R. (1987). Causal models do not support scientific conclusions: A comment in support of Freedman. Journal of Educational Statistics, 12, 185-195.

Rogosa, D.R. & Willet, J.B. (1985). Understanding correlates of change by modeling individual differences. Psychometrika, 50, 203-228.

Singer, J.D. (1999). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth curve models. Journal of Educational and Behavioral Statistics, 23, 323-356.

Willet, J.B. & Sayer, A.G. (1994). Using covariance structure analysis to detect correlates and predictors of change. Psychological Bulletin, 116, 363-381.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

TM030181

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: 1) Impact of MSPAP: Evidence from P, T, & S Questionnaires . . . .
2) MSPAP Performance Gains and their Relationship . . .
3) Impact of MSPAP: Evidence from Classroom Inst. & Asst Activities . . . .
4) Consequences of the MSPAP

Author(s): Lane et. al . . . . .

Corporate Source: University of Pittsburgh

Publication Date:

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1 ↑ ☑	Level 2A ↑ ☐	Level 2B ↑ ☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here,→ please

Signature: [signature]

Printed Name/Position/Title: Suzanne Lane Professor

Organization/Address: University of Pittsburgh. 5C01 Forbes Quad University of Pittsburgh, PA 15260

Telephone: 412-648-7095

FAX: 412-624-~~835~~ 7231

E-Mail Address: SL+@pitt.edu

Date: Sept 15, 1999

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

University of Maryland
ERIC Clearinghouse on Assessment and Evaluation
1129 Shriver Laboratory
College Park, MD 20742
Attn: Acquisitions

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
1100 West Street, 2nd Floor
Laurel, Maryland 20707-3598

Telephone: 301-497-4080
Toll Free: 800-799-3742
FAX: 301-953-0263
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com