

AN EXAMINATION OF FRESHMEN TO SENIOR GENERAL EDUCATION GAINS ACROSS A NATIONAL SAMPLE OF INSTITUTIONS WITH DIFFERENT GENERAL EDUCATION REQUIREMENTS USING A MIXED-EFFECT STRUCTURAL EQUATION MODEL

William E. Knight

.....

This study investigated differences in freshmen to senior student general education gains across institutions with varying patterns of general education requirements using a mixed-effect structural equation model. The subjects were 6,409 students at 34 nation-wide colleges and universities. Students attending institutions where less than 40 percent of undergraduate curricular requirements were devoted to general education and where there was not equal distribution of general education courses within the requirement were found to have significantly higher general education gains than did students who attended institutions where 40 percent or more of the undergraduate curriculum was devoted to general education and there was equal distribution of courses within the general education requirement.

.....

While interest in longitudinal studies of student knowledge and skill gains in general education has been widespread, numerous methodological problems exist in the traditional study of student growth and development. Several authors (Baird, 1988; Pascarella, 1989) have concluded that the difficulties associated with the classic methods of analyzing change vitiate the benefits associated with a longitudinal or "value-added" approach.

Student learning in the area of general education has been of particular interest at the University of Tennessee at Knoxville (UTK). A report by Pike, Philippi, Banta, Bensey, Milbourne, and Columbus (1991, May) documented the results of several studies carried out on freshmen to senior UTK student gains

William E. Knight, Center for the Study of Higher and Postsecondary Education, 2007 School of Education, The University of Michigan, Ann Arbor, MI 48109-1259.

Presented at the Association for Institutional Research Annual Forum, Atlanta, Georgia, May 10-13, 1992.

on the ACT COMP Objective Test of general education knowledge and skills. Pike (1991) additionally used UTK student COMP score data to examine freshmen to senior gains and their relationship to patterns of coursework through a mixed-effect structural equation modeling technique. This approach provided for an alternative method not burdened by the inadequacies of traditional methods of studying change as discussed by Baird (1988) and Pascarella (1989) and allowing for error-free measurement of student general education learning.

Pike (1991) listed relatively small sample size ($N = 704$) and the reliance on a single outcome measure as limitations of his study. However, reliance on student ACT COMP scores from a single institution was another limitation. A number of studies cited most recently by Pascarella and Terenzini (1991) suggest that student learning or value-added in general education may follow significantly different patterns in different types of institutions.

Research on the development of students' general cognitive skills (Pace, 1984, 1990; Winter, McClelland, and Stewart, 1981) suggested that students in selective liberal arts colleges experience significantly greater freshman to senior gains than do students attending other types of institutions, even when controlling for entering abilities. The greatest student gains in critical thinking were found in a study by Dressel and Mayhew (1954) at institutions with specially designated general education courses and core curricula requirements. Forrest (1982) studied general education gains at 44 institutions by analyzing ACT COMP estimated freshmen to senior gains and the relationship between gains and institutional type. Again, he found that general education gains were highest at institutions devoting relatively large portions of their curricular requirements to general education.

Pascarella and Terenzini (1991, p. 157) concluded that institutional characteristics appear to have some influence on the development of student general cognitive skills, but they caution that "more methodologically sound studies, using objective measures and/or controlling for important precollege characteristics" are necessary. Thus, an application of a structural equation modeling methodology to the study of freshmen to senior general education gains at institutions with varying approaches toward, and student requirements for, general education should be enlightening.

AN OVERVIEW OF STRUCTURAL EQUATION MODELING

Structural equation modeling, otherwise known as latent variable modeling, has become a popular method for studying student change and development over recent years (Ecob, 1987). Structural equation models are designed to analyze changes in observed variables using unmeasured theoretical constructs, and they are based on the assumption that relationships among observed variables are the products of relationships among latent constructs (James, Mulaik,

and Brett, 1982; Long, 1983). Unlike the analysis of gain scores, which tends to produce nonsensical and counterintuitive results, structural equation models are found to produce results consistent with theory and research (Hendrickson and Jones, 1987).

The power in structural equation modeling lies in the ability of the researcher to create hierarchies of models by successively freeing or constraining parameters (Pike, 1991). The hierarchies allow the comparison of theoretically acceptable models so that the model that gives the best explanation of the observed data can be determined (Joreskog and Sorbom, 1976). The fact that the same construct is being measured by pre- and posttests is a prerequisite for change research. Structural equation modeling allows the stability or invariance of latent variables to be tested over time. The one requirement of the measurement model for assuming that latent variables represent true scores is that the observed variables must be congeneric, meaning they measure the same construct and are linearly related (Alwin and Jackson, 1981; Joreskog, 1971).

A limitation of structural models is that they provide information exclusively about group effects (Muthén, 1989). Some relatively simple modifications have been suggested (Muthén, 1989) that change structural equation models from fixed-effect models appropriate for analyzing group differences to mixed-effect models appropriate for studying interactions between groups and individuals. Estimating latent variable means using a mixed-effect model with multiple groups allows for the identification of effects specific to individuals and effects representing a conditional relationship or interaction between groups and individuals.

PROCEDURES

The procedures paralleled those used by Pike (1991), with two differences: (1) ACT Assessment Composite scores, rather than English and mathematics subscores, were used as the observed measures with the Precollege Ability latent variable, and (2) groups were based on data from students attending institutions with different general education requirements as opposed to students with different course-taking patterns from the same institution. ACT Assessment Composite scores were used because they were the only manifest variable associated with Precollege Ability available in the data set received from ACT. Observed measures of both Freshmen Achievement and Senior Achievement were the three COMP Objective Test content subtest scores: Functioning Within Social Institutions (FSI), Using Science and Technology (US), and Using the Arts (UA). The research model for the present study is shown in Figure 1.

The models were suggested by Pike (personal communication, February 18 and 24, 1991) following the guidelines of Marsh and Grayson (1990). They

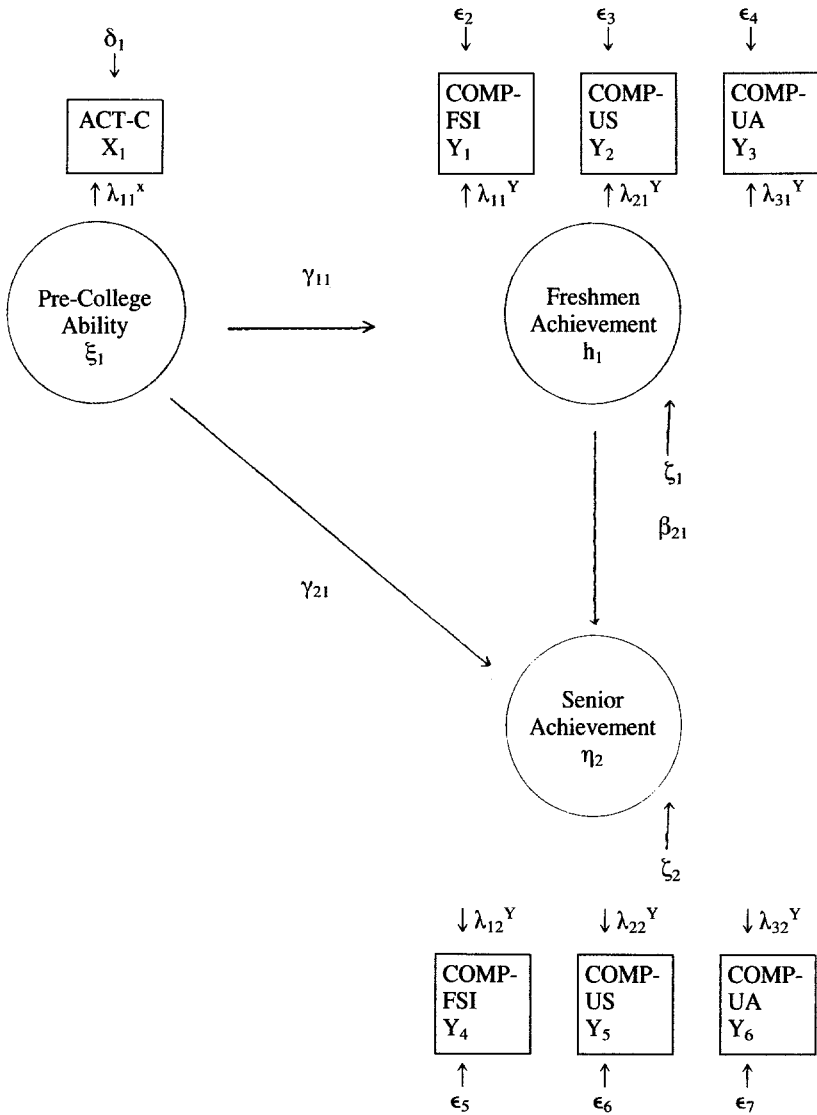


FIG. 1. The research model.

assumed Freshmen Achievement was influenced by Precollege Ability and Senior Achievement was influenced by both Precollege Ability and Freshmen Achievement. ACT Assessment and COMP content scores were found to be normally distributed in preliminary analyses. Lambda parameters were fixed at unity for ACT Composite and COMP Functioning Within Social Institutions

scores. Effects parameters for Precollege Ability were fixed at unity. The effects parameter for Freshmen Achievement was free to vary and the structural disturbance for Senior Achievement was set at zero.

In each of the sets of analyses, hierarchical structural equation models with various degrees of constraints were compared against a baseline model using Tucker Lewis Indices (and Type-2 Relative Normed Fit Indices for the mixed-effect models) in order to assess the goodness-of-fit for each model. Given that more than one model in each set met a minimum goodness-of-fit criterion (.90 for TLI and RNFI₂), the most parsimonious model was then selected as best.

Three sets of analyses paralleling those of Pike (1991) were performed using LISREL VII (Joreskog and Sorbom, 1989) based on procedures suggested by Hayduk (1987). The first set of analyses examined the invariance of the measurement model across time (freshmen and senior scores) and the second examined the invariance of the measurement model across groups. The third set of analyses tested the observed data against several mixed-effect models (where more than one grouping factor is involved). Due to the large sample size, a Tucker-Lewis Index was used in addition to the traditional chi-squared test and a Type-2 Relative Normed Fit Index was also used for the mixed-effect models.

The final procedures involved computing latent variable means using parameter estimates from the chosen mixed-effect model and structural equations (1) and (2) at given Precollege Ability levels for the four groups and then testing for significant differences among the four groups in Freshmen to Senior Achievement gains. The structural equations used to compute latent variable means are shown below:

$$E(\eta_1) = \alpha_1 + \gamma_{11}E(\epsilon_1) \quad (1)$$

$$E(\eta_2) = \alpha_2 + \beta_{21}E(\eta_1) + \gamma_{11}E(\epsilon_1) \quad (2)$$

Invariance Across Time

The first set of analyses used scores combined from the four groups based on general education programs to test the invariance of the measurement model over time. The first model constrained six measured variables, which were the six COMP content subtest scores, to be uncorrelated; this served as a standard against which other models could be evaluated (Marsh, Balla, and McDonald, 1988). The second model constrained Freshmen Achievement and Senior Achievement latent variables to be perfectly correlated. The third model was the same as the second except that the requirement of perfect correlation was relaxed. The fourth model kept the perfect correlation between latent variables but allowed error terms to be correlated across times.

Invariance Across Groups

The second set of analyses involved evaluating the invariance of the measurement model across groups. The first model constrained six measured variables (freshmen and senior COMP content scores) to be uncorrelated and provided a standard for evaluating the other models. Model Two included the two latent parameters with lambda and epsilon coefficients allowed to vary across groups. The third model was similar to the second except that lambda parameters were constrained to be equal across groups. Model Four had both lambda and epsilon parameters constrained to be the same across groups.

Testing of Mixed-Effect Models

The third set of analyses involved testing the data against several mixed-effect models proposed by Pike (personal communications, February 18 and 24, 1992). Model One was comprised of seven uncorrelated measured variables; it was the same as the standard model used earlier to test for invariance of measurement across groups except that ACT scores were also included. The second model was a mixed-effect model whereby students in the four groups with Precollege Ability scores constrained to be equal had different Freshmen and Senior Achievement scores. Model Three constrained the parameters in the structural equation for Freshmen Achievement to be the same for all groups, implying that students with the same Precollege Ability scores would have the same Freshmen Achievement scores but different Senior Achievement scores. The fourth model constrained the parameters in the structural equations for both Freshmen and Senior Achievement to be the same for the four groups. This meant students with the same Precollege Ability scores would have the same Freshmen and Senior Achievement scores. A fifth model, where all parameters were constrained to be invariant across groups, was used to serve as a baseline for computing Type-2 Relative Normed Fit Indices for the third and fourth models.

Testing for Significant Differences in Latent Freshmen to Senior Achievement Gains Among the Four Groups

Maximum likelihood estimates for the selected mixed-effect model were used along with equations (1) and (2) to compute latent freshmen and senior achievement means for the four groups. Differences in achievement gains for the four groups were tested for significance using a latent variable multivariate analysis of variance (MANOVA) procedure suggested by Bagozzi and Yi (1991). Planned comparisons of latent achievement gains were used to detect where significant differences occurred among the four groups.

SAMPLE

The study sample consisted of 6,409 students from 34 institutions for whom ACT Assessment and freshmen and senior ACT COMP Objective Test scores were available in a data set compiled by ACT. Sixty-three percent of students in the data set attended public institutions. Thirty-two percent of institutions in the data set were research universities as classified by Carnegie type; 9 percent were doctoral-granting universities; 32 percent were comprehensive institutions; and 27 percent were liberal arts colleges.

Fifty-nine percent of students in the data set were female and students' average age at the time of taking the COMP as seniors was 21.5. Their mean ACT Assessment Composite score was 21.1. Mean freshmen total COMP Objective Test score was 176.0. Mean senior COMP Objective Test score was 184.6, representing a mean score gain of 8.6 for this group.

A final variable in the ACT data set corresponded to type of general education program in place at the institution students attended when taking the Objective Test as freshmen and again as seniors. Institutions were grouped according to Forrest's (1982) classification scheme based on general education curricular characteristics. The four groups included (1) programs where general education distribution requirements comprised 40 percent or more of the curriculum and were evenly divided between courses in written and oral communication, social sciences and history, natural sciences and mathematics, and humanities and fine arts (five institutions, $N = 218$); (2) programs with 39 percent or less of the curriculum devoted to general education and no equal distribution (four institutions, $N = 2,319$); (3) programs with 40 percent or more of the curriculum devoted to either general education or equal distribution requirements but not both (23 institutions, $N = 3,281$); and (4) programs with core general education curricula (two institutions, $N = 591$).

RESULTS

Invariance Across Time

The results of hierarchical models used to test for invariance across time are shown in Table 1. While all of the models had significant chi squares, this was not taken conclusively to mean that the models did not represent the data due to the very large sample size. Models 2, 3, and 4 all provided an acceptable fit to the data because they had Tucker-Lewis Index values greater than .90 using the procedures recommended by Marsh, Balla, and McDonald (1988). Having established that all three models provided an acceptable fit to the data, Model 2, the time invariance model, which suggested the ACT COMP was a congeneric test that measured the same underlying construct for freshmen and seniors in

TABLE 1. Goodness-of-Fit Results for Models of Invariance Over Time

Model	<i>df</i>	χ^2	TLI
Baseline Model	15	12,212.66*	
Time Invariance	10	83.11*	0.99
Common Factors	10	83.11*	0.99
Correlated Errors	7	36.09*	0.99

* $p < .001$

the sample, was employed throughout subsequent analyses since it represented the most parsimonious model.

Invariance Across Groups

Results of the use of hierarchical models to test for invariance across groups are shown in Table 2. The second, third, and fourth models were all found to provide adequate representations of the observed data since they showed Tucker-Lewis indices in excess of .90. The epsilon or pattern invariance model (2), suggesting the same constructs were being measured across groups, was employed for the final set of analyses because it represented the most parsimonious model.

Testing of Mixed-Effect Models

The final set of LISREL analyses was used to distinguish between hierarchical mixed-effect models. Model 2 allowed students in the four groups who had the same Precollege Ability scores to have different Freshmen and Senior Achievement scores. Model 3 held Freshmen Achievement scores constant for students with identical Precollege Ability scores but allowed them to have different Senior Ability scores. Model 4 held both Freshmen and Senior Achievement scores constant for students with the same Precollege Ability scores.

The results are shown in Table 3. All three models had Tucker-Lewis indices

TABLE 2. Goodness-of-Fit Results for Models of Invariance Across Groups

Model	<i>df</i>	χ^2	TLI
Baseline Model	60	11,755.81*	
Pattern Invariance	40	132.61*	0.99
Lambda Invariance	46	197.04*	0.98
Error Invariance	64	317.35*	0.98

* $p < .001$

TABLE 3. Goodness-of-Fit Results for Mixed-Effect Models

Model	<i>df</i>	χ^2	TLI	RNFI ₂
Baseline Model	84	16,959.61*		
F/S Ach. Free	82	426.60*	0.98	
Fr. Ach. Constant	85	495.58*	0.98	0.89
F/S Ach. Constant	91	502.79*	0.98	0.89
Baseline Model for RNFI ₂	127	1,048.70*	0.96	

**p* < .001

above .90. Neither Model 3 nor Model 4 had Type-2 Relative Normed Fit Indices at or above .90, so they were not selected as superior to Model 2. The most parsimonious model (2), which allowed Freshmen and Senior Achievement to vary for students with the same levels of Precollege Ability, was used with equations (1) and (2) to calculate scores for Freshmen and Senior Achievement given various Precollege Ability scores. The maximum likelihood estimates for this model are shown in Table 4. Latent variable means for Freshmen and Senior Achievement at various levels of Precollege Ability are shown in Table 5.

Testing for Significant Differences in Freshmen to Senior Latent Achievement Gains Among the Four Groups

A multivariate analysis of variance (MANOVA) procedure suggested by Bagozzi and Yi (1991) was used to test for significant differences among freshmen to senior latent difference scores for the four groups by comparing the chi square and degrees of freedom resulting from the chosen mixed-effect model to that of a second model where latent variable means were constrained to be equal across groups. The results are shown in Table 6. The results showed a significant difference occurred between the two models, indicating evidence to

TABLE 4. Selected Maximum Likelihood Estimates for the Mixed-Effect Model

Parameter	Group 1	Group 2	Group 3	Group 4
Alpha ₁	39.099	37.559	38.271	37.513
Gamma _{1,1}	1	1	1	1
Alpha ₂	45.735	44.618	46.199	43.028
Gamma _{2,1}	1	1	1	1
Beta _{2,1}	-0.075	-0.066	-0.086	-0.044

TABLE 5. Latent Variable Means for Freshmen and Senior Achievement at Precollege Ability Levels of 10, 20, and 30

Precollege ability = 10				
Latent Variable	Group 1	Group 2	Group 3	Group 4
Freshmen Achievement	49.099	47.559	48.271	47.513
Senior Achievement	52.053	51.479	52.048	50.937
Precollege ability = 20				
Latent Variable	Group 1	Group 2	Group 3	Group 4
Freshmen Achievement	59.099	57.559	58.271	57.513
Senior Achievement	61.303	60.819	61.188	60.497
Precollege ability = 30				
Latent Variable	Group 1	Group 2	Group 3	Group 4
Freshmen Achievement	69.099	67.559	68.271	67.513
Senior Achievement	70.553	70.159	70.328	70.057

reject the null hypothesis that latent achievement gain scores were equal across groups.

Finally, planned comparisons were carried out to determine which groups differed in gain scores among latent achievement levels. To accomplish this, the results of the chosen mixed-effect model were compared with the results of

TABLE 6. Latent Variable MANOVA Results Used to Test the Null Hypothesis of No Difference in Achievement Gains Among the Four Groups

Model	<i>df</i>	χ^2	Δdf	$\Delta \chi^2$
Mixed-Effect Model	82	426.60		
Null Model (Latent Means Equal Across Groups)	88	497.22	6	70.62*

* $p < .01$

TABLE 7. Results of Planned Comparisons Between Groups One and Two and Groups Two and Four

Model	<i>df</i>	χ^2	Δdf	$\Delta\chi^2$
Mixed-Effect Model	82	426.60		
Group One and Two Means Equal	88	497.22	6	70.62*
Group Two and Four Means Equal	84	426.97	2	0.37

* $p < .01$

two models where latent variable means were constrained to be equal across (1) group one and group two, and (2) group two and group four. The results (Table 7) showed there were significant differences in achievement gains between groups one and two, but not between groups two and four. Gains were significantly greater for students in group two institutions than for students in group one institutions.

DISCUSSION

The results showed that structural equation modeling represented a useful method of representing the interactions between student characteristics and general education program effects. The theoretical constructs of Precollege Ability, Freshmen Achievement, and Senior Achievement and their interrelationships first noted by Pike (1991) were found to adequately represent the data collected from a nation-wide group of colleges and universities. The important abilities of structural equation models to demonstrate invariance over time and across groups proved useful in this study.

The best (most parsimonious) model that adequately represented the data was one where both Freshmen and Senior Achievement varied with given fixed levels of Precollege Ability. This is an interesting finding because it says that given the same ACT Assessment composite scores, freshmen students' COMP scores are not the same across institutions. This speaks well for the COMP's ability to detect different general education knowledge and skill levels given the same ACT Composite scores but leads to the question of why students with the same Precollege Ability levels show different levels of freshmen general education achievement across institutions in the four groups. Perhaps there is some self-selection of students into particular types of institutions even given the same levels of high school ability.

The study showed that students in group two institutions (where less than 40 percent of the minimum undergraduate curricular requirements were devoted to general education courses and where there was not equal distribution of the curriculum among the general education requirement) showed significantly

greater COMP gains than students in group one institutions (where 40 percent or more of undergraduate students' requirements were in general education and there was equal distribution of communication, mathematics, social sciences, natural sciences, and humanities courses within the general education requirement).

This finding is exactly the opposite of that of Forrest (1982), whose study of the effects of institutional general education requirements on students' COMP score gain showed an average COMP total score gain of 8.9 points for students from institutions with the same characteristics as the present study's group one and an average gain of 3.8 points for students from institutions with the same general education requirements as the present study's group two.

Three possible reasons are suggested for the contradictory results of Forrest's (1982) study and the present research. The provision of error-free measurement afforded by the methodology of the present study overcame a limitation of Forrest's (1982) effort in which the results were obscured due to problems in the study of gain scores. A second explanation for the two different research results may be that the grouping of institutions in the present study may not be optimal. The third explanation offered for differences in the results of the two studies is that intervening demographic and educational characteristics may have influenced the present research. Several variables were compared between institutions and students included in groups one and two and considerable differences existed between institutional type, control, and enrollment, and somewhat of a difference in student gender, between group one and two institutions.

Curricular implications resulting from the present study are that the definitions, course content, and instruction pertaining to general education may be considerably different between institutions in the sample and that which is expressed in the COMP Objective Test. Policy implications include the idea that a relatively large percentage of the undergraduate curricula devoted to general education and equal distribution of course content within the general education requirement is not necessarily positively related to greater gains in student achievement (at least as measured by the COMP). This suggests changes in student requirements or at least a more critical study of the effects of general education programs.

One limitation was that there was some subjectivity in the grouping of the institutions (i.e., what constitutes equal distribution of courses in some cases). Another limitation may have been that the timing of the pretesting at each institution varied and that more than one freshmen to senior cohort may be included for each institution.

Other useful applications of this methodology to research in general education gains might include examining the effects of some of the demographic and institutional variables noted above, in addition to institutional general education requirements, on general education gains and using the same methodology with

students at institutions where general education assessments other than the COMP are utilized.

REFERENCES

- Alwin, D.F., and Jackson, D.J. (1981). Applications of simultaneous factor analysis to the issues of factorial invariance. In D.J. Jackson and E.F. Borgatta (eds.), *Factor Analysis and Measurement in Sociological Research: A Multi-dimensional Perspective* (pp. 249–279). Studies in International Sociology, No. 21. Beverly Hills: Sage.
- Bagozzi, R.P., and Yi, Y. (1991). On the use of structural equation models in experimental designs: Two extensions. *International Journal of Research in Marketing* 8:125–140.
- Baird, L.L. (1988). Value added: Using student gains as yardsticks of learning. In C. Adelman (ed.), *Performance and Judgement: Essays on Principles and Practice in the Assessment of College Student Learning* (pp. 205–216). Washington, DC: United States Government Printing Office.
- Dressel, P., and Mayhew, L. (1954). *General Education: Explorations in Evaluation*. Westport, CT: Greenwood Press.
- Ecob, R. (1987). Applications of structural equation modeling to longitudinal educational data. In P. Cuttance and R. Ecob (eds.), *Structural Equation Modeling by Example: Applications in Educational, Sociological, and Behavioral Research* (pp. 138–159). Cambridge: Cambridge University Press.
- Forrest, A. (1982). *Increasing Student Competence and Persistence: The Best Case for General Education*. Iowa City, IA: American College Testing Program.
- Hayduk, L.A. (1987). *Structural Equation Modeling with LISREL: Essentials and Advances*. Baltimore: Johns Hopkins University Press.
- Hendrickson, L., and Jones, B. (1987). A study of longitudinal causal models comparing gain score analysis with structural equation approaches. In P. Cuttance and R. Ecob (eds.), *Structural Equation Modeling by Example: Applications in Educational, Sociological, and Behavioral Research* (pp. 86–106). Cambridge: Cambridge University Press.
- James, L. R., Mulaik, S. A., and Brett, J. M. (1982). *Causal Analysis: Assumptions, Models, and Data*. Beverly Hills, CA: Sage Publishers.
- Joreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika* 36:109–133.
- Joreskog, K. G., and Sorbom, D. (1976). Statistical models and methods for test-retest situations. In D. N. M. DeGrujter and L. J. van der Kamp (eds.), *Advances in Psychological and Educational Measurement* (pp. 135–157). New York: Wiley.
- Joreskog, K. G., and Sorbom, D. (1989). *LISREL 7: A Guide to the Program and Applications* (2nd ed.). Chicago: SPSS.
- Long, J. S. (1983). *Covariance Structure Models: An Introduction to LISREL*. Beverly Hills, CA: Sage.
- Marsh, H. W., Balla, J. R., and McDonald, R. P. (1988). Goodness-of-fit indices in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin* 103:91–410.
- Marsh, H. W., and Grayson, D. (1990). Public/Catholic differences in the high school and beyond data: A multigroup structural equation modeling approach to testing mean differences. *Journal of Educational Statistics* 15:199–236.
- Muthén, B. O. (1989). *Analysis of Longitudinal Data Using Latent Variable Models*

- with Varying Parameters*. Unpublished manuscript. Los Angeles: University of California Los Angeles, Graduate School of Education.
- Pace, C. (1984). *Measuring the Quality of College Student Experiences*. Los Angeles: University of California, Higher Education Research Institute.
- Pace, C. (1990). *The Undergraduates: A Report on Their Activities and Progress in College in the 1980s*. Los Angeles: University of California, Center for the Study of Evaluation.
- Pascarella, E. T. (1989). Methodological issues in assessing the outcomes of college. In C. Fincher (ed.), *Assessing Institutional Effectiveness: Issues, Methods, and Management* (pp. 19–32). Athens, GA: University of Georgia Press.
- Pascarella, E. T., and Terenzini, P. T. (eds.) (1991). *How College Affects Students: Findings and Insights from Twenty Years of Research*. San Francisco: Jossey-Bass.
- Pike, G. R. (1991). Using mixed-effect structural equation models to study student growth and development. *Research in Higher Education* 32: 499–524.
- Pike, G. R., Phillippi, R. H., Banta, T. W., Bensey, M. W., Milbourne, C. C., and Columbus, P. J. (1991, May). *Freshmen to Senior Gains at the University of Tennessee, Knoxville*. Paper presented at the Association for Institutional Research Forum, San Francisco.
- Sorbom, D. (1974). A general method for studying differences in factor means and factor structure between groups. *British Journal of Mathematical and Statistical Psychology* 28: 229–239.
- Winter, D., McClelland, D., and Stewart, A. (1981). *A New Case for the Liberal Arts: Assessing Institutional Goals and Student Development*. San Francisco: Jossey-Bass.