The Basics of Structural Equation Modeling

Diana Suhr, Ph.D.
University of Northern Colorado

Abstract

Structural equation modeling (SEM) is a methodology for representing, estimating, and testing a network of relationships between variables (measured variables and latent constructs). This tutorial provides an introduction to SEM including comparisons between "traditional statistical" and SEM analyses. Examples include path analysis/ regression, repeated measures analysis/latent growth curve modeling, and confirmatory factor analysis. Participants will learn basic skills to analyze data with structural equation modeling.

Rationale

Analyzing research data and interpreting results can be complex and confusing. Traditional statistical approaches to data analysis specify default models, assume measurement occurs without error, and are somewhat inflexible. However, structural equation modeling requires specification of a model based on theory and research, is a multivariate technique incorporating measured variables and latent constructs, and explicitly specifies measurement error. A model (diagram) allows for specification of relationships between variables.

Purpose

The purpose of this tutorial is to provide participants with basic knowledge of structural equation modeling methodology. The goals are to present a powerful, flexible and comprehensive technique for investigating relationships between measured variables and latent constructs and to challenge participants to design and plan research where SEM is an appropriate analysis tool.

Structural equation modeling (SEM*)*

- is a comprehensive statistical approach to testing hypotheses about relations among observed and latent variables (Hoyle, 1995).
- is a methodology for representing, estimating, and testing a theoretical network of (mostly) linear relations between variables (Rigdon, 1998).
- tests hypothesized patterns of directional and nondirectional relationships among a set of observed (measured) and unobserved (latent) variables (MacCallum & Austin, 2000).

Two **goals** in SEM are

1) to understand the patterns of correlation/covariance among a set of variables and
2) to explain as much of their variance as possible with the model specified (Kline, 1998).

The **purpose** of the model, in the most common form of SEM, is to account for variation and covariation of the measured variables (MVs). Path analysis (e.g., regression) tests models and relationships among MVs. Confirmatory factor analysis tests models of relationships between latent variables (LVs or common factors) and MVs which are indicators of common factors. Latent growth curve models (LGM) estimate initial level (intercept), rate of change (slope), structural slopes, and variance. Special cases of SEM are regression, canonical correlation, confirmatory factor analysis, and repeated measures analysis of variance (Kline, 1998).

Similarities between Traditional Statistical Methods and SEM

SEM is similar to traditional methods like correlation, regression and analysis of variance in many ways. First, both traditional methods and SEM are based on linear statistical models. Second, statistical tests associated with both methods are valid if certain assumptions are met. Traditional methods assume a normal distribution and SEM assumes multivariate normality. Third, neither approach offers a test of causality.

Differences Between Traditional and SEM Methods

Traditional approaches differ from the SEM approach in several areas. First, SEM is a highly flexible and comprehensive methodology. This methodology is appropriate for investigating achievement, economic trends, health issues, family and peer dynamics, self-concept, exercise, self-efficacy, depression, psychotherapy, and other phenomenon.

Second, traditional methods specify a default model whereas SEM requires formal specification of a model to be estimated and tested. SEM offers no default model and places few limitations on what types of relations can be specified. SEM model specification requires researchers to support hypothesis with theory or research and specify relations a priori.

Third, SEM is a multivariate technique incorporating observed (measured) and unobserved variables (latent constructs) while traditional techniques analyze only measured variables. Multiple, related equations are solved simultaneously to determine parameter estimates with SEM methodology.

Fourth, SEM allows researchers to recognize the imperfect nature of their measures. SEM explicitly specifies error while traditional methods assume measurement occurs without error.

Fifth, traditional analysis provides straightforward significance tests to determine group differences, relationships between variables, or the amount of variance explained. SEM provides no straightforward tests to determine model fit. Instead, the best strategy for evaluating model fit is to examine multiple tests (e.g., chi-square, Comparative Fit Index (CFI), Bentler-Bonett Nonnormed Fit Index (NNFI), Root Mean Squared Error of Approximation (RMSEA)).

Sixth, SEM resolves problems of multicollinearity. Multiple measures are required to describe a latent construct (unobserved variable). Multicollinearity cannot occur because unobserved variables represent distinct latent constructs.

Finally, a graphical language provides a convenient and powerful way to present complex relationships in SEM. Model specification involves formulating statements about a set of variables. A diagram, a pictorial representation of a model, is transformed into a set of equations. The set of equations are solved simultaneously to test model fit and estimate parameters.

Statistics

Traditional statistical methods normally utilize one statistical test to determine the significance of the analysis. Structural Equation modeling, however, relies on several statistical tests to determine the adequacy of model fit to the data. The chi-square test indicates the amount of difference between expected and observed covariance matrices. A chi-square value close to zero indicates little difference between the expected and observed covariance matrices. In addition, the probability level must be greater than 0.05 when chi-square is close to zero.

The Comparative Fit Index (CFI) is equal to the discrepancy function adjusted for sample size. CFI ranges from 0 to 1 with a larger value indicating better model fit. Acceptable model fit is indicated by a CFI value of 0.90 or greater (Hu & Bentler, 1999).

Root Mean Square Error of Approximation (RMSEA) is related to residual in the model. RMSEA values range from 0 to 1 with a smaller RMSEA value indicating better model fit. Acceptable model fit is indicated by an RMSEA value of 0.06 or less (Hu & Bentler, 1999).

If model fit is acceptable, the parameter estimates are examined. The ratio of each parameter estimate to its standard error is distributed as a z statistic and is significant at the 0.05 level if its value exceeds 1.96 and at the 0.01 level it its value exceeds 2.56 (Hoyle, 1995). Unstandardized parameter estimates retain scaling information of variables and can only be interpreted with reference to the scales of the variables. Standardized parameter estimates are transformations of unstandardized estimates that remove scaling and can be used for informal comparisons of parameters throughout the model. Standardized estimates correspond to effect-size estimates.

If unacceptable model fit is found, the model could be revised when the modifications are meaningful. Model modification involves adjusting a specified and estimated model by either freeing parameters that were fixed or fixing parameters that were free. The Lagrange multiplier test provides information about the amount of chi-square change that results if fixed parameters are freed. The Wald test provides information about the change in chi-square that results if free parameters are fixed (Hoyle, 1995).

Considerations

The use of SEM could be ***impacted*** by

- the research hypothesis being testing
- the requirement of sufficient sample size
  A desirable goal is to have a 20:1 ratio for the number of subjects to the number of model parameters . However, a 10:1 may be a realistic target. If the ratio is less than 5:1, the estimates may be unstable.
- measurement instruments
- multivariate normality
- parameter identification
- outliers
- missing data
- interpretation of model fit indices (Schumacker & Lomax, 1996).

SEM Process

A ***suggested approach*** to SEM analysis proceeds through the following process:

- review the relevant theory and research literature to support model specification
- specify a model (e.g., diagram, equations)
- determine model identification (e.g., if unique values can be found for parameter estimation; the number of degrees of freedom, df, for model testing is positive)
- select measures for the variables represented in the model
- collect data
- conduct preliminary descriptive statistical analysis (e.g., scaling, missing data, collinearity issues, outlier detection)
- estimate parameters in the model
- assess model fit
- respecify the model if meaningful
- interpret and present results.

Definitions

A ***measured variable*** (MV) is a variables that is directly measured whereas a latent variable (LV) is a construct that is not directly or exactly measured.

A ***latent variable*** could be defined as whatever its multiple indicators have in common with each other. LVs defined in this way are equivalent to common factors in factor analysis and can be viewed as being free of error of measurement.

Relationships between variables are of three types

- Association, e.g., correlation, covariance
- Direct effect is a directional relation between two variables, e.g., independent and dependent variables
- Indirect effect is the effect of an independent variable on a dependent variable through one or more intervening or mediating variables

Variable Labels

- Independent → predictor → exogenous (external) → affect other variables in the model
- Dependent → criterion → endogenous (internal) → effects of other variables → can be represented as causes of other endogenous variables
- Latent variable → factor → construct
- Observed variable → measured variable → manifest variable → indicator → generally considered endogenous

A ***model*** is a statistical statement about the relations among variables.

A ***path diagram*** is a pictorial representation of a model.

Specification is formulating a statement about a set of parameters and stating a model.

- A critical principle in model specification and evaluation is the fact that all of the models that we would be interested in specifying and evaluating are wrong to some degree
- We must define as an optimal outcome a finding that a particular model fits our observed data closely and yields a highly interpretable solution.
- Instead of considering all possible models, a finding that a particular model fits observed data well and yields an interpretable solution can be taken to mean only that the model provides one plausible representation of the structure that produced the observed data.

Parameters are specified as fixed of free.

Fixed parameters are not estimated from the data and their value is typically fixed to zero or one.

Values of fixed parameters are generally defined based on requirements of model specification. A critical requirement is that we ***establish a scale*** for each LV in the model, including error terms. To resolve this, we provide each LV with a scale in the model specification process in one of two ways

- Fix the variance of each LV to 1.0
- Fix the value to 1.0 of one parameter associated with an LV directional influence

Free parameters are estimated from the data.

Fit indices indicate the degree to which a pattern of fixed and free parameters specified in the model are consistent with the pattern of variances and covariances from a set of observed data. Examples of fit indices are chi-square, CFI, NNFI, RMSEA.

Components of a general structural equation model are the measurement model and the structural model. The ***measurement model*** prescribes latent variables, e.g., confirmatory factor analysis. The ***structural model*** prescribes relations between latent variables and observed variables that are not indicators of latent variables.

Identification involves the study of conditions to obtain a single, unique solution for each and every free parameter specified in the model from the observed data. In order to obtain a solution, the number of free parameters, q, must be equal to or smaller than the number of nonredundant elements in the sample covariance matrix, denoted as p* with $p^* = p(p + 1)/2$ where p is the number of measured variables in the covariance matrix ($q =< p^*$).

Types of model identification

- If a single, unique value cannot be obtained from the observed data for one or more free parameters, then the model is ***underidentified***. For example, infinite solutions may be obtained for the equation $x + y = 5$. The solution for y is totally dependent on the solution for x. When there are more unknowns (x and y) than the number of equations (1), the model is underidentified.
- If for each free parameter a value can be obtained through one and only one manipulation of the observed data, then the model is ***just identified***. With two equations, $x + y = 5$ and $2x + y = 8$, a unique solution can be obtained. When the number of linearly independent equations is the same as the number of unknown parameters, the model is just identified. This means we can get unique parameter estimates, but the model cannot be tested.
- If a value for one or more free parameters can be obtained in multiple ways from the observed data, then the model is ***overidentified***. Overidentification is the condition in which there are more equations than unknown independent

parameters. For example, x + y = 5, 2x + y = 8, and x + 2y = 9. There is no exact solution. We can define a criterion and obtain the most adequate solution as an alternative. An advantage of the overidentified model is that we can test the model.

- When df is positive, all q parameters can be estimated, df = (p* - q).
- For example, with the three equations below, find values of a and b that are positive and yield totals such that the sum of the squared difference (0.67) between the observations (6, 10, and 12) and these totals is as small as possible (a = 3.0 and b = 3.3 to one decimal place). This solution does not perfectly reproduce the observations 6, 10, and 12.
  a + b = 6, 2a + b = 10, 3a + b = 12

The purpose of ***estimation*** is to obtain numerical values for the unknown (free) parameters.

Maximum Likelihood Estimation

- ML is the default for many model-fitting programs
- ML estimation is simultaneous, estimates are calculated all at once
- If the estimates are assumed to be population values, they maximize the likelihood (probability) that the data (the observed covariances) were drawn from the population (the expected covariances).
- Maximum likelihood estimation methods are appropriate for nonnormally distributed data and small sample size.

The ***criterion selected for parameter estimation*** is know as the ***discrepancy function***.
It provides a guideline to minimize the differences between the population covariance matrix, Σ, as estimated by the sample covariance, **S**, and the covariance matrix derived from the hypothesized model, $\Sigma(\mathbf{0})$. For example, the discrepancy function for the ML method is

$$Fml = \log|\Sigma(\mathbf{0})| + \text{Trace}[\Sigma(\mathbf{0})^{-1}\mathbf{S}] - \log|\mathbf{S}| - p$$

Iterative methods involve a series of attempts to obtain estimates of free parameters that imply a covariance matrix like the observed one. Iterative means that the computer derives an initial solution and then attempts to improve these estimates through subsequent cycles of calculations. “Improvement” means the model-implied covariances based on the estimates from each step become more similar to the observed ones. The iterative process continues until the values of the elements in the residual matrix cannot be minimized any further. Then the estimation procedure has ***converged****.*

When the estimation procedure has converged, a single number is produced that summarizes the degree of correspondence between the expected and observed covariance matrices. That number is sometimes referred to as the ***value of the fitting function****.* That value is the starting point for constructing indexes of model fit.

Nonconvergence occurs when the iterative process is terminated during estimation because of criteria specified (e.g., maximum 30 iterations) or because of the practical consideration of excessive computer time. Nonconvergent results must not be trusted. Nonconvergence is usually the result of poor model specification or poor starting values.

The residual matrix contains elements whose values are the differences between corresponding values in the expected and observed matrices.

The ***sample covariance matrix is not positive definite*** – usually caused by linear dependency among observed variables; some variables are perfectly predictable by others. Because the inverse of the sample covariance matrix is needed in the process of computing estimates, solutions cannot be obtained from the estimation procedure when variables are linearly dependent. To avoid dependencies, variables should be carefully selected before the model is estimated and tested.

Evaluation of model fit

- *Chi-square* is a “badness-of-fit” index, smaller values indicate better fit
- *Other fit indices, e.g, CFI, NNFI,* are “goodness-of-fit” indices where larger values mean better fit
- *The Wald test* provides information about the change in chi-square and determines the degree to which model fit would deteriorate if free parameters were fixed.
- *LaGrange Multiplier Test* (LM) provides information about the amount of chi-square change and determines the degree to which model fit would improve if any selected subset of fixed parameters were converted into free parameters.

Model modification involves adjusting a specified and estimated model by either freeing parameters that were fixed or fixing parameters that were free. In SEM, model comparison is analogous to planned comparisons in ANOVA, and model modification is analogous to post-hoc comparisons in ANOVA. Model modification could sacrifice control over Type I error and lead to a situation where sample specific characteristics are generalized to a population

If a model is determined to have ***acceptable fit***, then the focus moves to specific elements of fit.

- The ***ratio of each parameter estimate to its standard error*** is distributed as a ***z statistic*** and is significant at the 0.05 level if its value exceeds 1.96.
- ***Unstandardized parameter estimates*** retain scaling information of variables involved and can only be interpreted with reference to the scales of the variables

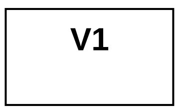
- ***Standardized parameter estimates*** are transformations of unstandardized estimates that remove scaling information and can be used for informal comparisons of parameters throughout the model. Standardized estimates correspond to effect-size estimates.

What indicates a “large” ***direct effec***t? A “small” one?

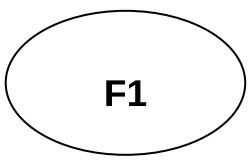
- Results of significance tests reflect not only the absolute magnitudes of path coefficients but also factors such as the sample size and intercorrelations among the variables
- Standardized path coefficients with absolute values less than .10 may indicate a “small” effect
- Values around .30, a “medium” effect
- Values greater than .50, a “large” effect

Note: SEM does nothing more than test the relations among variables as they were assessed. Researchers are often too quick to infer causality from statistically significant relations in SEM

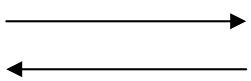
Diagram Symbols



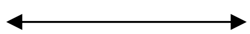
measured variable (V1), observed variable



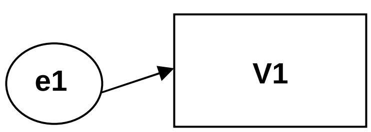
latent construct (F1), factor, unmeasured variable



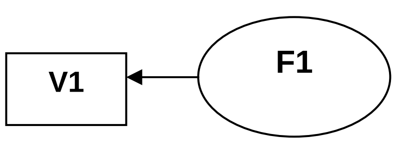
direct relationship



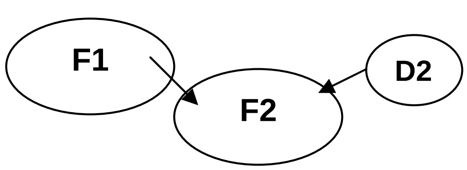
covariance or correlation



error (e1) associated with measured variable (V1)



path coefficient for regression of a latent variable (F1) on an observed variable (V1)



path coefficient for regression of one latent variable (F1) onto another latent variable (F2), residual error (D2) in prediction of F2 by F1.

Examples

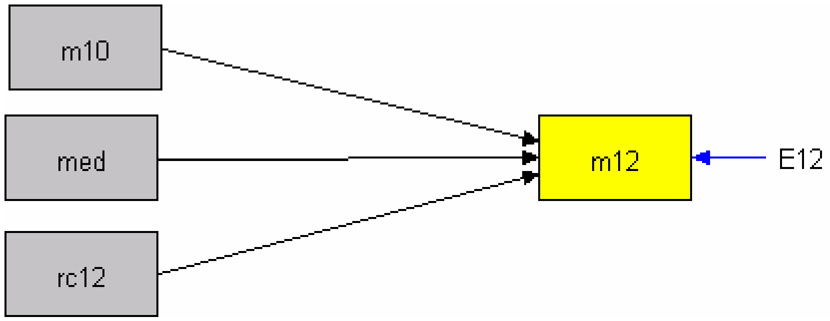


Figure 1. Regression Model (math achievement at age 10, reading comprehension achievement at age 12, and mother's educational level predicting math achievement at age 12).

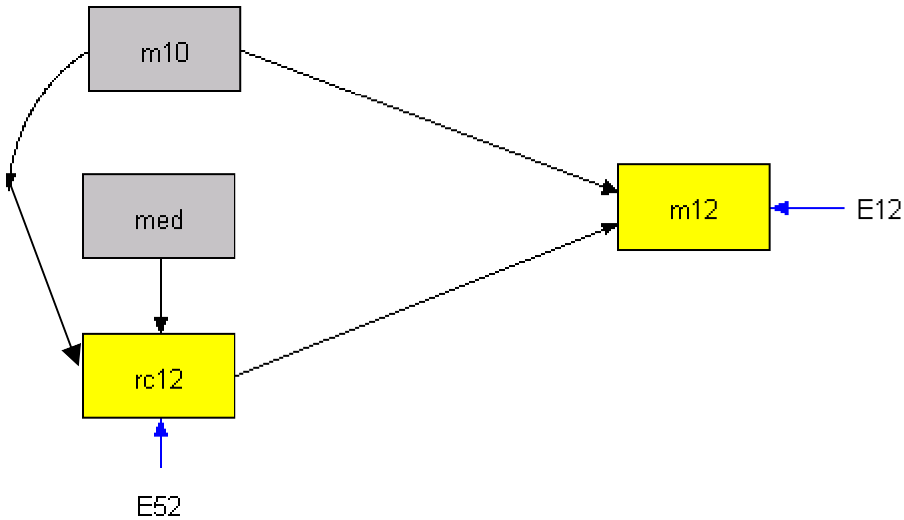


Figure 2. Revised model (math achievement at age 10, reading comprehension at age 12 predict math achievement at age 12; indirect effect of mother's educational level and math achievement at age 10).

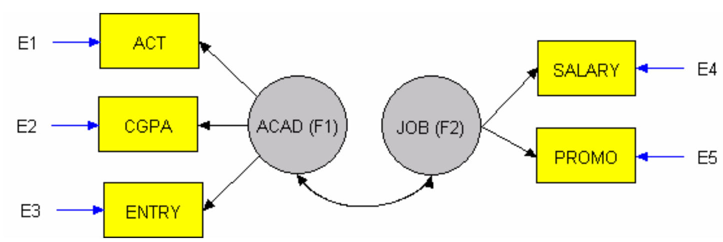


Figure 3. Structural Equation Model - Relationship between academic and job constructs

EXAMPLE 1: Regression/Path Analysis with PROC REG and PROC CALIS

SEM techniques are easily applied to analyses in the health field. Application of SEM techniques have contributed to research on illness (Roth, Wiebe, Fillingim, & Shay, 1989), on exercise (Duncan, T. & McAuley, 1993; Duncan, T., Oman, & Duncan, S., 1994) and on substance use/abuse among adults (Curran, Harford, & Muthen, 1996) and adolescents (Curran, Stice, & Chassin, 1997; Duncan, T., Duncan, S., Alpert, Hops, Stoolmiller, & Muthen, 1997).

The path analysis and confirmatory factor analysis examples reanalyze data from a study where researchers investigated the effects of hardiness, stress, fitness, and exercise on health problems (Roth, et al., 1989). College students (n=373) reported recent physical illness, recent stressful life events, current exercise participation levels, current perceived fitness levels, and hardiness components.

Multiple regression and SEM analyses examined the effects related to illness. Subjects were 163 men and 210 women enrolled in an introductory psychology course at a southern United States university. The mean age of the sugjects was 21.7 (sd = 5.5).

Assessments

Illness. Seriousness of Illness Rating Scale (Wyler, Masuda, & Holmes, 1968) is a self-report checklist of commonly recognized physical symptoms and diseases and provides a measure of current and recent physical health problems. Each item is associated with a severity level. A total illness score is obtained by adding the severity ratings of endorsed items (symptoms experienced within the last month).

Stress. Life Experience Survey (Sarason, Johnson, & Segal, 1978) is a measure used to access the occurrence and impact of stressful life experiences. Subjects indicate which events have occurred within the last month and rate the degree of impact on a 7-point scale (-3 = extremely negative impact, 0 = no impact, 3 = extremely positive impact). In the study, the total negative event score was used as an index of negative life stress (the absolute value of the sum of negative items).

Fitness. Fitness Questionnaire (Roth & Fillingim, 1988) is a measure of self-perceived physical fitness. Respondents rate themselves on 12 items related to fitness and exercise capacity. The items are on an 11-point scale of 0 = very poor fitness to 5 = average fitness to 10 = excellent fitness. A total fitness score is calculated by summing the 12 ratings. Items include questions about strength, endurance, and general perceived fitness.

Exercise. Exercise Participation Questionnaire (Roth & Fillingim, 1988) assessed current exercise activities, frequency, duration, and intensity. An aerobic exercise participation score was calculated using responses to 15 common exercise activities and providing blank spaces to write in additional activities.

Hardiness. In the study, hardiness included components of commitment, challenge, and control. A composite hardiness score was obtained by summing Z scores from scales on each component. The challenge component included one scale whereas the other components included 2 scales. Therefore, the challenge Z score was doubled when calculating the hardiness composite score.
Commitment was assessed with the Alienation From Self and Alienation From Work scales of the Alienation Test (Maddi, Kobasa, & Hoover, 1979). Challenge was measured with the Security Scale of the California Life Goals Evaluation Schedule (Hahn, 1966). Control was assessed with the External Locus of Control Scale (Rotter, Seaman, & Liverant, 1962) and the Powerlessness Scale of the Alienation Test (Maddi, Kobasa, & Hoover, 1979).

Results

Tests were conducted to determine if variables as a whole predicted a significant proportion of the variance of the illness measure and whether each individual variance uniquely accounted for a significant proportion of that variance. A main effects regression model including stress, fitness, hardiness, exercise, and gender to predict illness accounted for approximately 20% of the variance. The SEM analysis excluding gender found hardiness mediated by stress and exercise mediated by fitness (Roth, et al., 1989).

SAS Code

```
data illfl (type=corr);
   input  _type_ $1-4     _name_ $6-13
          exercise 15-20  hardy 22-27
          fitness 29-34   stress 36-41
          illness 43-48;
cards;
n              373    373    373    373    373
mean             40.90   0.00  67.10  4.80 716.7
std              66.50   3.80  18.40  6.70 624.8
corr    exercise  1.00
corr    hardy    −0.03   1.00
corr    fitness   0.39   0.07   1.00
corr    stess    −0.05  −0.23  −0.13  1.00
corr    illness  −0.08  −0.16  −0.29  0.34  1.00
;;;;
```

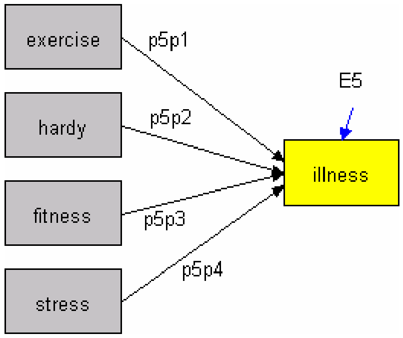


Figure 4. Regression/Path Analysis Model

Regression Model

Note: The PROC GLM procedure will not accept a correlation or covariance matrix as input. Therefore, the PROC REG procedure was run.

```
proc reg data=illfl;
   model illness = exercise hardy fitness stress /selection = backward;
```

Reanalysis
The data was reanalyzed with PROC REG (regression) and PROC CALIS (path analysis). Input data was in the form of a correlation matrix, means, and standard deviations. A correlation matrix standardizes values and loses the metric of the scales. A covariance matrix would have preserved the metric of each scale.

Regression/Path Analysis Model
```
proc calis data=illfl corr stderr;
   lineqs
      illness = p5p3 fitness + p5p4 stress + p5p1 exercise + p5p2 hardy + e5;
   std
     exercise = varex,
     hardy    = varhr,
     fitness  = varft,
     stress   =  varst,
     e5       = vare5;
  var exercise fitness hardy stress illness;
```

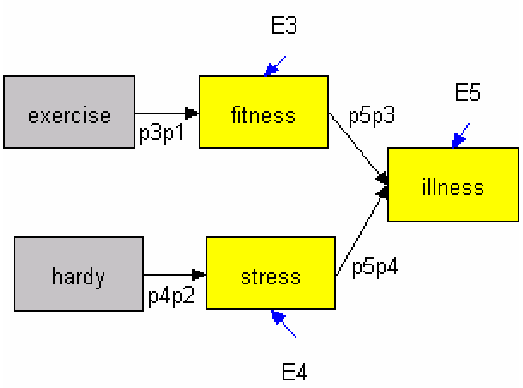


Figure 5. Structural Equation Model - Illness

Structural Equation Model
```
proc calis data=illfl corr stderr;
   lineqs
     fitness = p3p1 exercise              + e3,
     stress  = p4p2 hardy                 + e4,
     illness = p5p3 fitness + p5p4 stress + e5;
   std
     exercise = vare1,
     hardy    = vare2,
     e3       = vare3,
     e4       = vare4,
     e5       = vare5;
  var exercise fitness hardy stress illness;
```

Results
The regression analysis with backward elimination retained fitness and stress (p < 0.0001) in the model while removing exercise and then hardy. R-squared with all variables in the model is equal to 0.1835. R-squared for the revised model is equal to 0.177.

The path analysis model showed some interesting results in terms of model fit: chi-square = 0.000, df = 0, p < 0.0001 and CFI =1.000 and RMSEA = 0.000. Similar to the regression analysis, significant parameter estimates were fitness (z = -5.067) and stress (z = 6.000) while parameter estimates for hardy (z = -1.530) and exercise (z = 0.663) were not significant.

Table 1. Parameter Estimates – Path Analysis

Parameter	*Variable*	*Estimate*	*Std. Error*	*z-value*
p5p1	Exercise → illness	0.034	0.051	0.663
p5p2	Hardy → illness	-0.074	0.048	-1.530
p5p3	Fitness → illness	-0.260	0.051	-5.067
p5p4	Stress → illness	0.291	0.049	6.000
Variances				
e5	Illness - unexplained variance	0.817	0.060	13.64

Standardized and unstandardized parameter estimates were equal because input was in the form of a correlation matrix (standardized values, mean = 0 , std = 1). The standardized equation is
illness = 0.034*exercise – 0.260*fitness – 0.074*hardy
+ 0.291*stress + 0.904*e5

R-squared (1 - unexplained variance squared) from the path analysis model is equal to 0.1835 ($1 – 0.9036^2 = 1 – 0.8165$). The unexplained variance is the amount of variance that cannot be accounted for with the predictor variables.

The structural equation model with hardiness mediated by stress and exercise mediated by fitness showed acceptable fit on three measures, chi-square (11.078, df = 5, p = 0.050), CFI (0.961), and RMSEA (0.057).

Unstandardized and standardized parameter estimates are equal due to input in the form of a correlation matrix (standardized, mean = 0, std = 1). The standardized equations are
Fitness = 0.390*exercise + 0.921*e3
Stress = -0.230*hardy + 0.973*e4
Illness = =0.253*fitness + 0.311*stress + 0.917*e5

Table 2. Parameter Estimates – Structural Equation Model

Parameter	*Variable*	*Estimate*	*Std. Error*	*z-value*
p3p1	exercise → fitness	0.390	0.048	8.169
P4p2	hardy → stress	-0.230	0.051	-4.559
P5p3	fitness → illness	-0.250	0.047	-5.316
p5p4	stress → illness	0.308	0.047	6.538
Variances				
varex	Exercise	1.000	0.073	13.650
varhr	Hardy	1.000	0.073	13.650
varft	Fitness	0.848	0.062	13.640
varst	Stress	0.947	0.069	13.640
e5	Illness – unexplained variance	0.823	0.603	13.640

Discussion
The regression and SEM model specify different relationships between variables in the model. The models include the same predictor and predicted variables in different configurations. Although statistical tests of significance differ, the amount of variance explained in each model is equal.

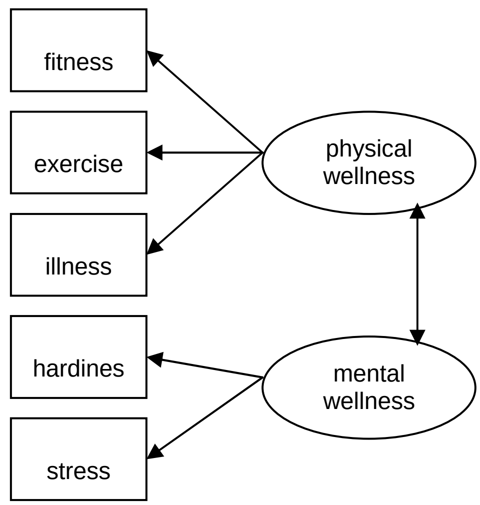
Analysis Conclusion
The development of theoretical models prior to SEM data analysis is critical. The results of the SEM analysis can serve to support or refute previous research. The direction of parameter estimates indicates effects on illness. More hardiness indicates less stress and less stress indicates less illness. More exercise indicates better fitness and less illness. The SEM analysis provides flexibility in determining the relationships between variables. Direct as well as indirect relationships between variables can be specified and estimated.

EXAMPLE 2: Confirmatory Factor Analysis with PROC CALIS

This example hypothesizes two latent constructs related to wellness, physical (fitness, exercise, illness) and mental (stress, hardiness). CFA analyzes data from a study where researchers investigated the effects of hardiness, stress, fitness, and exercise on health problems (Roth, et al., 1989). College students (n=373) reported physical illness, stressful life events, exercise participation levels, perceived fitness levels, and hardiness components. Previously, multiple regression and SEM analyses examined the effects related to illness. Subjects were 163 men and 210 women enrolled in an introductory psychology course at a southern United States university. The mean age of the subjects was 21.7 (sd = 5.5).

SAS Code – CFA

```
data illfl (type=corr);
   input  _type_ $1-4     _name_ $6-13
          exercise 15-20  hardy 22-27
          fitness 29-34   stress 36-41
          illness 43-48;
cards;
n             373    373    373    373    373
mean            40.90   0.00  67.10  4.80 716.7
std             66.50   3.80  18.40  6.70 624.8
corr    exercise  1.00
corr    hardy    –0.03   1.00
corr    fitness   0.39   0.07   1.00
corr    stess    –0.05  –0.23  –0.13  1.00
corr    illness  –0.08  –0.16  –0.29  0.34  1.00
;;;;
```



(error terms not shown on diagram)

Figure 6. Confirmatory Factor Analysis – Health Data

```
proc calis data=illfl corr;
  lineqs
    hardy   = p1 F1 + e1,
    stress  = p2 F1 + e2,
    illness = p3 F1 + e3,
    fitness = p4 F2 + e4,
    exercise= p5 F2 + e5;
  std
    e1-e5 = vare1-vare5,
    F1 = 1,
    F2 = 1;
  cov
    F1 F2 = covf1f2;
VAR exercise fitness hardy stress illness;
```

Results

PROC CALIS procedure provides the number of observations, variables, estimated parameters, and informations (related to model specification). Notice measured variables have different scales.

```
The CALIS Procedure
Covariance Structure Analysis: Maximum Likelihood Estimation

Observations         373    Model Terms             1
Variables              5    Model Matrices          4
Informations          15    Parameters             11

Variable          Mean       Std Dev
exercise      40.90000      66.50000
fitness       67.10000      18.40000
hardy                0       3.80000
stress         4.80000       6.70000
illness      716.70000     624.80000
```

Fit Statistics

Determine criteria a priori to access model fit and confirm the factor structure. Some of the criteria indicate acceptable model fit while other are close to meeting values for acceptable fit.

- Chi-square describes similarity of the observed and expected matrices. Acceptable model fit. Is indicated by a chi-square probability greater than or equal to 0.05. For this CFA model, the chi-square value is close to zero and p = 0.0478, almost the 0.05 value.
- RMSEA indicates the amount of unexplained variance or residual. The 0.0613 RMSEA value is larger than the 0.06 or less criteria.
- CFI (0.9640), NNI (0.9101), and NFI (0.9420) values meet the criteria (0.90 or larger) for acceptable model fit.

For purposes of this example, 3 fit statistics indicate acceptable fit and 2 fit statistics are close to indicating acceptable fit. The CFA analysis has confirmed the factor structure. If the analysis indicates unacceptable model fit, the factor structure cannot be confirmed, an exploratory factor analysis is the next step.

```
The CALIS Procedure
Covariance Structure Analysis: Maximum Likelihood Estimation

Fit Function . . .
Chi-Square                                            9.5941
Chi-Square DF                                              4
Pr > Chi-Square                                       0.0478
. . .
RMSEA Estimate                                        0.0613
. . .
Bentler's Comparative Fit Index                       0.9640
. . .
Bentler & Bonett's (1980) Non-normed Index            0.9101
Bentler & Bonett's (1980) NFI                         0.9420
. . .
```

Parameter Estimates

When acceptable model fit is found, the next step is to determine significant parameter estimates.

- A t value is calculated by dividing the parameter estimate by the standard error, 0.3213 / 0.1123 = 2.8587.
- Parameter estimates are significant at the 0.05 level if the t value exceeds 1.96 and at the 0.01 level if the t value exceeds 2.56.
- Parameter estimates for the confirmatory factor model are significant at the 0.01 level.

```
Manifest Variable Equations with Estimates
exercise =   0.3212*F2       +  1.0000 e5
Std Err      0.1123 p5
t Value      2.8587

fitness  =   1.2143*F2       +  1.0000 e4
Std Err      0.3804 p4
t Value      3.1923

hardy    =   0.2781*F1       +  1.0000 e1
Std Err      0.0673 p1
t Value      4.1293

stress   =  -0.4891*F1       +  1.0000 e2
Std Err      0.0748 p2
t Value     -6.5379

illness  =  -0.7028*F1       +  1.0000 e3
Std Err      0.0911 p3
t Value     -7.7157
```

Variances are significant at the 0.01 level for each error variance except vare4 (error variance for fitness).

Variances of Exogenous Variables

Variable	Parameter	Estimate	Standard Error	t Value
F1		1.00000		
F2		1.00000		
e1	vare1	0.92266	0.07198	12.82
e2	vare2	0.76074	0.07876	9.66
e3	vare3	0.50604	0.11735	4.31
e4	vare4	-0.47457	0.92224	-0.51
e5	vare5	0.89685	0.09209	9.74

Covariance between latent constructs is significant at the 0.01 level.

Covariances Among Exogenous Variables

Var1	Var2	Parameter	Estimate	Standard Error	t Value
F1	F2	covf1f2	0.30404	0.11565	2.63

Correlation between latent constructs is 0.30. Latent constructs are uncorrelated.

Correlations Among Exogenous Variables

Var1	Var2	Parameter	Estimate
F1	F2	covf1f2	0.30404

Standardized Estimates
Report equations with standardized estimates when measured variables have different scales

```
Manifest Variable Equations with Standardized Estimates
exercise =   0.3212*F2       +  0.9470 e5
                    p5
fitness  =   1.2143*F2       +  1.0000 e4
                    p4
hardy    =   0.2781*F1       +  0.9606 e1
                    p1
stress   =  -0.4891*F1       +  0.8722 e2
                    p2
illness  =  -0.7028*F1       +  0.7114 e3
                    p3
```

EXAMPLE 3: Repeated Measures Analysis with PROC GLM and PROC CALIS

This example investigates the change in reading achievement for 7- through 13-year old girls. Comparisons of repeated measures anova using PROC GLM and a latent growth curve model using PROC CALIS are shown.

Participants
Participants were part of the National Longitudinal Survey of Youth (NLSY79). The original NLSY79 sample design enabled researchers to study longitudinal experiences of different age groups as well as analyze experiences of women, Hispanics, Blacks, and economically disadvantaged. The NLSY79 is a nationally representative sample of 12,686 young men and women who were 14- to 22-years old when first surveyed in 1979 (Baker, Keck, Mott, & Quinlan, 1993).

As part of the NLSY79, mothers and their children have been surveyed biennially since 1986. Although the NLSY79 initially analyzed labor market behavior and experiences, the child assessments were designed to measure academic achievement as well as psychological behaviors. The child sample consisted of all children born to NLSY79 women respondents participating in the study biennially since 1986.

The number of children interviewed for the NLSY79 from 1988 to 1994 ranged from 4,971 to 7,089. Due to attrition, the number of children who completed four assessments in reading recognition achievement was 1,188. A cohort-sequential design in a previous study permitted estimation of growth curves for 5- through 14-year old boys and girls (Suhr, 1999). For simplicity, the 7-year old girls cohort was used as the example in this paper (n = 135). Assessments were collected at ages 7-, 9-, 11-, and 13-years old for the 7-year old cohort.

Measurement Instrument
The PIAT (Peabody Individual Achievement Test) Reading Recognition Subtest, provided an achievement measure of word recognition and pronunciation ability, essential components of reading (Dunn & Markwardt, 1970). Test-retest reliability of PIAT Reading Recognition ranged between 0.81 for kindergartners to a high of 0.94 for third graders (median = 0.89).

Models
The default model for the repeated measures anova using PROC GLM assumes change is linear, constant across time, and occurs at one unit between each measurement period. In addition, traditional statistical approaches assume measurement occurs without error. Therefore, the PROC GLM model specifies no measurement error. SEM using PROC CALIS explicitly specifies measurement error.

The SEM model shown in Figure 1 specifies latent variables, initial level (intercept) and rate of change (slope). A constant is regressed on the latent variables to determine the mean initial level and mean rate of change across the time period shown. The latent variables are regressed on the measured variables to estimate structural slopes and determine rate of change. The error terms specified indicate the amount of variance in each measured variable (reading achievement). The disturbance terms (error terms of latent variables) estimate the amount of unexplained variance for each latent variable. The correlation between the disturbance terms indicates the relationship between the unexplained variances.

PROC GLM for Repeated Measures Anova

The SAS Code below provides a repeated measures analysis of variance with four repeated measured variables (rr7, rr9, rr11, rr13). The repeated measures represent the assessment of reading recognition achievement for girls ages 7-, 9-, 11-, and 13-years old.

```
PROC GLM DATA = COH7F;
  CLASS GENDER;
  MODEL RR7 RR9 RR11 RR13 =   /NOUNI;
  REPEATED RR 4 / SUMMARY;
```

The NOUNI command requests no univariate (anova) models be printed. The REPEATED RR 4 command indicates four repeated measurements be renamed RR.

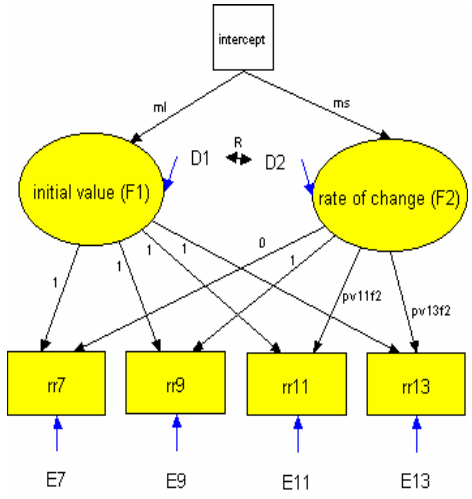


Figure 7. Latent Growth Model – Reading Readiness Achievement

Note: The intercept is a constant equal to 1. Measured variables (reading recognition achievement) are rr7, rr9 rr11, and rr13 measured at ages 7, 9, 11, and 13 respectively. Variances of measured variables are estimated with variances of E7, E9, E11, and E13. Structural slopes are fixed at 0 and 1 for ages 7 and 9 respectively. Structural slopes are estimated by pv11f2 and pv13f2 for ages 11 and 13. Unmeasured variables (latent variables) are initial value (F1) and rate of change (F2). Variances of latent variables are denoted and estimated by D1 and D2. The relationship between latent variable variances is denoted and estimated with R. The intercept is regressed on the latent variables to estimate mean initial value (ml) and mean rate of change (ms).

PROC CALIS for Repeated Measures Anova

A graphical representation of the Latent Growth Curve model is shown in Figure 1. Repeated measures analysis of variance is a special case of a latent growth curve model.

The PROC CALIS code specifies the equations that will be solved simultaneously. Measured variables are rr7, rr9, rr11, and rr11 (four measures of reading achievement). Parameters to be estimated are the structural slopes (pv11f2, pv13f2), the mean initial level (ml), the mean rate of change (ms), error and disturbance variances (vare7, vare9, vare11, vare13, varD1, varD2) and covariance of the disturbance terms (CD1D2).

```
PROC CALIS DATA = COH7F UCOV AUG ALL;
  LINEQS
    RR7   =  F1  +              +  E7,
    RR9   =  F1  +           F2 +  E9,
    RR11  =  F1  +  PV11F2  F2 +  E11,
    RR13  =  F1  +  PV13F2  F2 +  E13,
    F1    =  ML  INTERCEPT      +  D1,
    F2    =  MS  INTERCEPT      +  D2;
  STD
    E7    =  VARE7,
    E9    =  VARE9,
    E11   =  VARE11,
    E13   =  VARE13,
    D1-D2 =  VARD1-VARD2;
  COV
    D1 D2 = CD1D2;
  VAR RR7 RR9 RR11 RR13;
```

Using parameter estimates from PROC CALIS, the repeated measures anova with PROC GLM was respecified.

PROC GLM with Respecified Repeated Measures Anova

Estimates of structural slopes (0 1 1.648 2.172) from PROC CALIS replace the original linear contrasts (1 2 3 4) in PROC GLM code. The repeated measures anova model was respecified using PROC GLM to test for nonlinear change.

```
PROC GLM DATA = COH7F;
  CLASS GENDER;
  MODEL RR7 RR9 RR11 RR13 =      /NOUNI;
  REPEATED RR 4 (0 1 1.648 2.172)
     POLYNOMIAL / SUMMARY PRINTM;
```

Results

Results from PROC GLM and PROC CALIS analyses for the repeated measures anova were similar. However, the SEM analysis using PROC CALIS proved more flexible and provided parameter estimates for rate of change, initial level, and variances that were not included in the PROC GLM analysis.

PROC GLM tested for significant mean differences in reading achievement between repeated measures with a linear default model (levels 1, 2, 3, 4). The SEM model estimated nonlinear growth with PROC CALIS by anchoring the rate of change at 0 and 1 and estimating parameters for the third and fourth measurements. The estimates for the third and fourth levels of change with PROC CALIS were 1.648 and 2.172. The growth steps (0, 1, 1.648, 2.172) estimated by PROC CALIS indicated a nonlinear growth trajectory. The SEM model does not assume change is constant and linear whereas the default PROC GLM repeated measures anova model assumes change is constant and equal at each time interval.

PROC GLM, however, was used to test the respecified anova model for a nonlinear trend. Significant differences between the means were found with the linear trend and with the nonlinear trend ($F(3,132) = 429.38$), $p < 0.0001$).

PROC CALIS has no specific tests of significance. The chi-square probability indicated an unacceptable model fit (chi-square = 9.101, df = 3, p = 0.028). The chi-square value is a measure of the difference between the observed and expected covariance matrices. Acceptable model fit is determined by a chi-square value close to zero and a probability value greater than 0.05. RMSEA indicated unacceptable model fit (RMSEA = 0.123). RMSEA is a measure similar to a residual indicating error. An RMSEA less than 0.06 indicates acceptable model fit (Hu and Bentler, 1999). However, CFI (0.997) and NNFI (0.989) indicated acceptable model fit (values greater than 0.90, Hu and Bentler, 1999). Table 1 illustrates the parameter estimates and statistical tests (z-values) included in the PROC CALIS procedure.

Table 3. Estimates for 7-year old girls cohort (n = 135)

	Estimate	SE	z-value
Mean Initial Value (F1)	29.671	0.862	34.410
Mean Growth Rate(F2)	14.156	0.641	22.103
Variance of Initial Value (D1)	78.836	12.832	6.140
Variance of Growth Rate (D2)	10.912	3.387	3.220
Covariance of Initial Value and Growth Rate (R)	8.298	4.972	1.670
Growth Scores			
Age 7	0.000*	---	---
Age 9	1.000*	---	---
Age 11 (pv11f2)	1.648	0.057	28.862
Age 13 (pv13f2)	2.172	0.083	26.259
Variances			
Age 7 (e7)	21.618	6.957	3.110
Age 9 (e9)	23.695	3.898	6.080
Age 11 (e11)	16.902	3.956	4.270
Age 13 (e13)	47.935	8.204	5.840

*indicates fixed value

Analysis Conclusion

Repeated measures anova using PROC GLM tested a default model to find significant differences between the means of the repeated measures. The SEM method, using PROC CALIS, estimated rates of change, variances of measured variables and latent variables, mean initial value and mean rate of change. The SEM analysis with PROC CALIS proved flexible, allowed for explicit representation of measurement error, and provided more information than the repeated measures anova with PROC GLM.

EXAMPLE 4: Repeated Measures Analysis including covariate with PROC CALIS

Few studies have applied Latent Growth Modeling (LGM) to investigate cognitive development or acquisition of academic skills. This example will contrast a baseline LGM model to an LGM model with gender as a covariate. The models show development of reading recognition achievement for 6- through 12-year old children.

Participants

Participants were part of the National Longitudinal Survey of Youth (NLSY79). A cohort-sequential design in a previous study permitted estimation of growth curves for 5- through 14-year old boys and girls (Suhr, 1999). For simplicity, the 6-year old cohort (both boys and girls) was used as the example in this paper (n = 352, 163 boys and 189 girls). Assessments were collected at ages 6-, 8-, 10-, and 12-years old for the 6-year old cohort.

Measurement Instrument
The PIAT (Peabody Individual Achievement Test) was developed following principles of item response theory (IRT). The PIAT Reading Recognition Subtest, provides an achievement measure of word recognition and pronunciation ability, essential components of reading (Dunn & Markwardt, 1970). Test-retest reliability of PIAT Reading Recognition ranged between 0.81 for kindergartners to a high of 0.94 for third graders (median = 0.89).

Models
The SEM models shown in Figures 4 and 5 specify latent variables, initial level (intercept) and rate of change (slope). A constant is regressed on the latent variables to determine the mean initial level and mean rate of change across the time period shown. The latent variables are regressed on the measured variables to estimate structural slopes and determine rate of change. The error terms specified indicate the amount of variance in each measured variable (reading achievement). The disturbance terms (error terms of latent variables) estimate the amount of unexplained variance for each latent variable. The correlation between the disturbance terms indicates the relationship between the unexplained variances.

A covariate (male) has been added to the baseline model. Figure 5 illustrates the covariate model. A constant is regressed on the covariate. The covariate is regressed on latent constructs, initial value and growth rate.

Both models were estimated and evaluated for model fit and significance of estimated parameters. Comparisons will be made to contrast the differences between the two models.

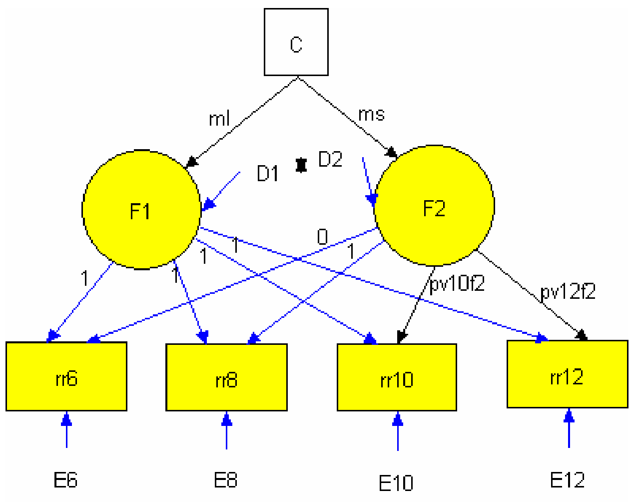


Figure 8. Latent Growth Model - Achievement

Note: The intercept is a constant equal to 1. Measured variables (reading recognition achievement) are rr6, rr8 rr10, and rr12 measured at ages 6, 8, 10, and 12 respectively. Variances of measured variables are estimated with variances of E6, E8, E10, and E12. Structural slopes are fixed at 0 and 1 for ages 6 and 8 respectively. Structural slopes are estimated by pv10f2 and pv12f2 for ages 10 and 12. Unmeasured variables (latent variables) are initial value (F1) and rate of change (F2). Variances of latent variables are denoted and estimated by D1 and D2. The relationship between latent variable variances is denoted and estimated with R. The intercept is regressed on the latent variables to estimate mean initial value (ml) and mean rate of change (ms).

SAS Code

Baseline Model

```
proc calis data=coh6 ucov aug;
  lineqs
    rr6  = F1               + e6,
    rr8  = F1  +         F2  + e8,
    rr10 = F1  +  pv10f2 F2  + e10,
    rr12 = F1  +  pv12f2 F2  + e12,
    F1   = ml intercept      + D1,
    F2   = ms intercept      + D2;
  std
    e6    = vare6,
    e8    = vare8,
    e10   = vare10,
    e12   = vare12,
    D1-D2 = varD1-varD2;
  cov
    D1 D2 = CD1D2;
  var rr6 rr8 rr10 rr12;
```

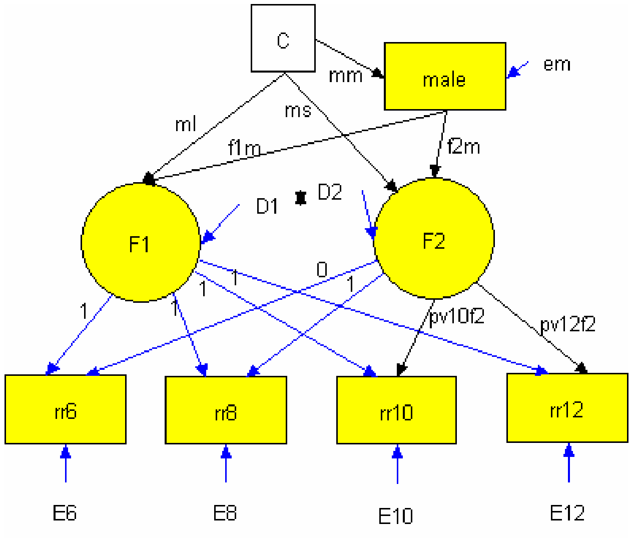


Figure 9. Covariate Latent Growth Model
Note: A categorical covariate (male) is regressed on initial value (F1) and growth rate (F2).

Model with Gender as a Covariate
The following LGM model includes the variable male as a covariate. The variable gender has been recoded into a variable "male" with the following code:
if gender eq 1 then male=1; else if gender eq 2 then male=0;

```
proc calis data = coh6 ucov aug;
  lineqs
    rr6   = F1                + e6,
    rr8   = F1  +          F2  + e8,
    rr10  = F1  +  pv10f2 F2  + e10,
    rr12  = F1  +  pv12f2 F2  + e12,
    male  = mm intercept      + em,
    F1    = f1m male + ml intercept + D1,
    F2    = f2m male + ms intercept + D2;
  std
    e6    = vare6,
    e8    = vare8,
    e10   = vare10,
    e12   = vare12,
    em    = varem,
    D1-D2 = varD1-varD2;
  cov
    D1 D2 = CD1D2;
  var rr6 rr8 rr10 rr12 male;
```

Results
The baseline model showed acceptable fit on CFI (0.992). Unacceptable fit was found with chi-square (40.235, df = 3, p < 0.0001) and RMSEA (0.188). The covariate model found acceptable fit for CFI (0.993) and unacceptable fit for chi-square (40.423, df = 5, p < 0.0001) and RMSEA (0.142). Values for fit indices changed a small amount with the addition of the gender covariate. However, differences between boys and girls in the parameter estimates could be investigated.

Discussion
Structural slopes in the baseline and covariate models were 0, 1, 1.609, and 2.082. Growth steps can be determined by finding differences between structural slopes. The growth steps were equal to 1, 0.609, and 0.473 (1-0, 1.609 – 1, and 2.082 – 1.609 respectively). Growth steps and structural slopes provide information about the shape of the developmental curve, e.g., nonlinear and a negatively accelerating function of time). Decomposition of effects to determine gender differences in structural slopes was not examined in this paper.

Differences between parameter estimates for boys and girls can be determined from the unstandardized equations of the covariate model. If male is equal to 1 (boys), initial value is equal to 19.544

(19.929 – 0.385) and growth rate is equal to 16.074 (17.762 – 1.688). For girls, male is equal to 0 and initial value is equal to 19.929, growth rate is equal to 17.762. Parameter estimates are shown in Table 5.

Table 4. Parameter Estimates – Baseline Model

	Estimate	*SE*	*z-value*
Mean Initial Value (F1)	19.750	0.319	62.012
Mean Growth Rate (F2)	16.980	0.392	43.376
Variance of Initial Value (D1)	23.181	3.640	6.370
Variance of Growth Rate (D2)	21.591	2.630	8.210
Covariance of Initial Value and Growth Rate(R)	9.462	2.389	3.960
Growth Scores			
Age 6	0.000*	--	--
Age 8	1.000*	--	--
Age 10 (pv10f2)	1.609	0.027	60.123
Age 12 (pv12f2)	2.082	0.039	53.578
Variances			
Age 6 (e6)	12.587	3.094	4.070
Age 8 (e8)	21.045	2.092	10.060
Age 10 (e10)	17.153	2.645	6.490
Age 12 (e12)	55.532	5.743	9.670

*fixed value

Standardized Equations for Baseline Model

```
rr6   = 0.985*F1              + 0.172*e6
rr8   = 0.537*F1   + 0.465*F2 + 0.121*e8
rr10  = 0.420*F1   + 0.585*F2 + 0.085*e10
rr12  = 0.357*F1   + 0.643*F2 + 0.131*e12
F1    = 0.972*V999 + 0.237*D1
F2    = 0.965*V999 + 0.264*D2
```

***T*able 5. Parameter Estimates – Covariate Model**

	Estimate	*SE*	*z-value*
Mean Initial Value (F1)	19.929	0.432	46.186
Mean Growth Rate (F2)	17.762	0.479	37.093
Mean Male	0.463	0.027	17.399
Variance of Initial Value (D1)	23.183	3.636	6.380
Variance of Growth Rate (D2)	20.924	2.579	8.110
Covariance of Initial Value and Growth Rate(R)	9.264	2.372	3.900
Growth Scores			
Age 6	0.000*	--	--
Age 8	1.000*	--	--
Age 10 (pv10f2)	1.609	0.027	60.123
Age 12 (pv12f2)	2.082	0.039	53.578
Variances			
Age 6 (e6)	12.521	3.088	4.050
Age 8 (e8)	21.049	2.091	10.070
Age 10 (e10)	17.074	2.630	6.490
Age 12 (e12)	55.719	5.740	9.710
Male (em)	0.249	0.018	13.250
Covariate-Initial Value (F1 ← Male) (f1m)	-0.385	0.628	-0.612
Covariate-Growth Rate (F2 ← Male) (f2m)	-1.688	0.579	-2.915

*fixed value

Standardized equations for Covariate Model

```
rr6   =  0.985*F1               + 0.171*e6
rr8   =  0.537*F1   + 0.465*F2   + 0.121*e8
rr10  =  0.420*F1   + 0.585*F2   + 0.085*e10
rr12  =  0.357*F1   + 0.643*F2   + 0.131*e12
male  =  0.681*V999             + 0.643*em
F1    = -0.013*male + 0.980*V999 + 0.237*D1
F2    = -0.065*male + 1.009*V999 + 0.260*D2
```

The covariate model adjusts baseline parameter estimates for initial value and growth rate. The change from the baseline to covariate model can be calculated to give a rough estimate of adjusted values. Initial value baseline parameter estimate plus the product of mean male (proportion male subjects) and parameter estimate for the regression of the covariate (male) on the latent construct (initial level) to calculate adjusted value, 19.750 + (0.464)(-0.385) = 19.750 - 0.178 = 19.572. Values used in the calculation are shown in Tables 4 and 5. Note: adjusted value is slightly different due to rounding.

Follow the same procedure to determine adjusted growth rate. Growth Rate baseline parameter estimate plus the product of mean male (proportion) and parameter estimate for regression of covariate (male) on latent construct (growth rate) to find adjusted value, 16.980 + (0.463)(-1.688) = 16.980 - 0.782 = 16.198.

Calculation of a rough estimate of adjusted values for girls can be calculated by using the opposite direction (sign) for the parameter estimates. Initial value baseline parameter estimate plus the product of proportion of girls and parameter estimate for the regression of the covariate (male) with reverse direction on the latent construct (initial level) to calculate adjusted value, 19.750 – (0.537)(0.385) = 19.750 + 0.207 = 19.957. Values used in the calculation are shown in Tables 4 and 5. Note: adjusted value is slightly different due to rounding.

Follow the same procedure to determine adjusted growth rate. Growth Rate baseline parameter estimate minus the product of mean male (proportion) and parameter estimate for regression of covariate (male) on latent construct (growth rate) to find adjusted value, 16.980 – (0.463)(-1.688) = 16.980 + 0.960 = 17.886.

Analysis Conclusion

LGM analysis illustrated, for this sample, that the magnitude of initial value and growth rate is greater for 6-year old girls than 6-year old boys in reading achievement. Including covariates in a SEM analysis allows for the investigation of group differences (e.g., gender) that may be overlooked in analysis of the group as a whole.

Conclusion

This tutorial has covered the basics of Structural Equation Modeling. There are differences and similarities between "traditional" statistical techniques and SEM. With SEM techniques, models are specified a priori, measurement error is specified explicitly, and models are tested for acceptable fit with chi-square and several fit indices. SEM gives you the power not available with "traditional" statistical procedures. You are challenged to design and plan research where SEM is an appropriate analysis tool.

WAM

Structural Equation Modeling is a flexible and powerful statistical methodology used to examine the relationships between measured variables and latent constructs.

References

Baker, P. C., Keck, C. K., Mott, F. L. & Quinlan, S. V. (1993). NLSY79 child handbook. A guide to the 1986-1990 National Longitudinal Survey of Youth Child Data (rev. ed.) Columbus, OH: Center for Human Resource Research, Ohio State University.

Chou, C. P. & Bentler, P. M. (1995). Estimates and tests in structural equation modeling. In Structural equation modeling: Concepts, issues, and applications, R. H. Hoyle (editor). Thousand Oaks, CA: Sage Publications, Inc..

Curran, P. J., Harford, T. C., & Muthen, B. O. (1996). The relation between heavy alcohol use and bar patronage: A latent growth model. *Journal of Studies on Alcohol,* 57, 410-418.

Curran, P. J., Stice, E., & Chassin, L. (1997). The relation between adolescent alcohol use and peer alcohol use: A longitudinal random coefficients model. *Journal of Consulting and Clinical Psychology,* 68(1), 130-140.

Duncan, T. E., Duncan, S. C., Alpert, A., Hops, H., Stoolmiller, M., & Muthen, B. (1997). Latent variable modeling of longitudinal and multilevel substance use data. *Multivariate Behavioral Research,* 32(3), 275-318.

Duncan, T. E., Duncan, S. C., Strycker, L. A., Li, F., & Alpert, A. (1999). An introduction to latent variable modeling: Concepts, issues, and applications. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Duncan, T. E., & McAuley, E. (1993). Social support and efficacy cognitions in exercise adherence: A latent growth curve analysis. *Journal of Behavioral Medicine,* 16(2), 199-218.

Duncan, T.E., Oman, R., & Duncan, S. C. (1994). Modeling incomplete data in exercise behavior research using structural equation methodology. *Journal of Sport and Exercise Psychology,* 16, 187-203.

Dunn, L. M. & Markwardt, J. C. (1970). Peabody individual achievement test manual. Circle Pines, MN: American Guidance Service.

Hahn, M. E. (1966). *California Life Goals Evaluation Schedule.* Palo Alto, CA: Western Psychological Services.

Hoyle, R. H. (1995). The structural equation modeling approach: Basic concepts and fundamental issues. In Structural equation modeling: Concepts, issues, and applications, R. H. Hoyle (editor). Thousand Oaks, CA: Sage Publications, Inc., pp. 1-15.

Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling, 6(1), 1-55.

Kline, R. B. (1998). *Principles and Practice of Structural Equation Modeling.* New York: The Guilford Press.

MacCallum, R. C. (1995). Model specification: Procedures, strategies, and related issues. In Structural equation modeling: Concepts, issues, and applications, R. H. Hoyle (editor). Thousand Oaks, CA: Sage Publications, Inc., pp. 16-29.

MacCallum, R. C. & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. Annual Review of Psychology, 51, 201-226.

Maddi, S. R., Kobasa, S. C., & Hoover, M. (1979) An alienation test. *Journal of Humanistic Psychology,* 19, 73-76.

Rigdon, E. E. (1998). Structural equation modeling. In Modern methods for business research, G. A. Marcoulides (editor). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers, pp. 251-294.

Roth, D. L., & Fillingim , R. B. (1988). *Assessment of self-reported exercies participation and self-perceived fitness levels.* Unpublished manuscript.

Roth, D. L., Wiebe, D. J., Fillingim, R. B., & Shay, K. A. (1989). "Life Events, Fitness, Hardiness, and Health: A Simultaneous Analysis of Proposed Stress-Resistance Effects". *Journal of Personality and Social Psychology,* 57(1), 136-142.

Rotter, J. B., Seeman, M., & Liverant, S. (1962). Internal vs. external locus of control: A major variable in behavior theory. In N. F. Washburne (Ed.), *Decisions, values, and groups* (pp. 473-516). Oxford, England: Pergamon Press.

Sarason, I. G., Johnson, J. H., & Siegel, J. M. (1978) Assessing the impact of life changes: Development of the Life Experiences Survey. *Journal of Consulting and Clinical Psychology,* 46, 932-946.

SAS® STAT, Version 6. Cary, N.C.: SAS Institute.

Schumacker, R. E. & Lomax, R. G. (1996). *A Beginner's Guide to Structural Equation Modeling*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

Suhr, D. D. (2000). An investigation of mathematics and reading achievement of 5- through 14-year olds using latent growth curve methodology. *Dissertation Abstracts International, A (Humanities and Social Sciences). Vol 60(12-A),* 4323, US: Univ Microfilms International.

Wyler, A. R., Masuda, M., & Holmes, T. H. (1968). Seriousness of Illness Rating Scale. *Journal of Psychosomatic Research,* 11, 363-374.

Contact Information

Diana Suhr is a Statistical Analyst in the Office of Budget and Institutional Analysis at the University of Northern Colorado. She holds a Ph.D. in Educational Psychology. She has been a SAS programmer since 1984.
email: diana.suhr@unco.edu