

chapter eight

Application 6: Testing Invariance of Latent Mean Structures

(First-Order CFA Model)

Key Concepts

- Distinguishing between observed and latent means
- Distinguishing between covariance and mean structures
- The moment matrix
- Critically important constraints regarding model identification and factor identification
- Illustrated use of the automated multigroup procedure in testing invariance
- Link between multigroup dialog box and automated labeling of Amos graphical models

In the years since the printing of my first Amos book (Byrne, 2001) there has been a steady, yet moderate increase in reported findings from tests for multigroup equivalence. A review of the SEM literature, however, reveals most tests for invariance to have been based on the analysis of covariance structures (COVS), as exemplified in Chapter 7 of this volume. Despite Sörbom's (1974) introduction of the mean and covariance structures (MACS) strategy in testing for latent mean differences over 40 years ago, only a modicum of studies have been designed to test for latent mean differences across groups based on real (as opposed to simulated) data (see, e.g., Aiken, Stein, & Bentler, 1994; Byrne, 1988b; Byrne & Stewart, 2006; Little, 1997; Marsh & Grayson, 1994; Cooke, Kosson, & Michie, 2001; Reise et al. 1993; Widaman & Reise, 1997). The aim of this chapter, then, is to introduce you to basic concepts associated with the analysis of latent mean structures, and to walk you through an application that tests for their invariance across two groups. Specifically, we test for differences in the latent means of general, academic, English, and mathematics self-concepts across high- and low-track secondary school students. The example presented here draws from two published papers—one that focuses on methodological issues related to testing for invariant covariance and mean

structures (Byrne et al., 1989), and one oriented toward substantive issues related to social comparison theory (Byrne, 1988b).

Basic Concepts Underlying Tests of Latent Mean Structures

In the usual univariate or multivariate analyses involving multigroup comparisons, one is typically interested in testing whether the *observed* means representing the various groups are statistically significantly different from each other. Because these values are directly calculable from the raw data, they are considered to be *observed* values. In contrast, the means of latent variables (i.e., latent constructs) are *unobservable*; that is, they are not directly observed. Rather, these latent constructs derive their structure indirectly from their indicator variables which, in turn, are directly observed and hence, measurable. Testing for the invariance of mean structures, then, conveys the notion that we intend to test for the equivalence of means related to each underlying construct or factor. Another way of saying the same thing, of course, is that we intend to test for differences in the latent means (of factors for each group).

For all the examples considered thus far, the analyses have been based on *covariance structures*. As such, only parameters representing regression coefficients, variances, and covariances have been of interest. Accordingly, the covariance structure of the observed variables constitutes the crucial parametric information; a hypothesized model can thus be estimated and tested via the sample covariance matrix. One limitation of this level of invariance is that while the unit of measurement for the underlying factors (i.e., the factor loading) is identical across groups, the origin of the scales (i.e., the intercepts) is not. As a consequence, comparison of latent factor means is not possible, which led Meredith (1993) to categorize this level of invariance as “weak” factorial invariance. This limitation, notwithstanding, evidence of invariant factor loadings nonetheless permits researchers to move on in testing further for the equivalence of factor variances, factor covariances, and the pattern of these factorial relations, a focus of substantial interest to researchers interested more in construct validity issues than in testing for latent mean differences. These subsequent tests would continue to be based on the analysis of covariance structures.

In contrast, when analyses are based on mean and covariance structures, the data to be modeled include both the sample means and the sample covariances. This information is typically contained in a matrix termed a *moment matrix*. The format by which this moment matrix is structured, however, varies with each particular SEM program.

In the analysis of covariance structures, it is implicitly assumed that all observed variables are measured as deviations from their means; in other words, their means (i.e., their intercepts) are equal to zero. As a consequence, the intercept terms generally associated with regression equations are not relevant to the analyses. However, when these observed means take on nonzero values, the intercept parameter must be considered thereby necessitating a reparameterization of the hypothesized model. Such is the case when one is interested in testing the invariance of latent mean structures, which necessarily includes testing first for invariance of the observed variable intercepts.

To help you in understanding the concept of mean structures, I draw on the work of Bentler (2005) in demonstrating the difference between covariance and mean structures as it relates to a simple bivariate regression equation. Consider first, the following regression equation:

$$y = \alpha + \beta x + \varepsilon(1)$$

where α is an intercept parameter. Although the intercept can assist in defining the mean of y , it does not generally equal the mean. Now, if we take expectations of both sides of this equation, and assume that the mean of ε is zero, the above expression yields

$$\mu_y = \alpha + \beta\mu_x(2)$$

where μ_y is the mean of y , and μ_x is the mean of x . As such, y and its mean can now be expressed in terms of the model parameters α , β , and μ_x . It is this decomposition of the mean of y , the dependent variable, that leads to the term *mean structures*. More specifically, it serves to characterize a model in which the means of the dependent variables can be expressed or “structured” in terms of structural coefficients and the means of the independent variables. The above equation serves to illustrate how the incorporation of a mean structure into a model necessarily includes the new parameters α and μ_x , the intercept and observed mean (of x), respectively. Thus, models with structured means merely extend the basic concepts associated with the analysis of covariance structures.

In summary, any model involving mean structures may include the following parameters:

- regression coefficients;
- variances and covariances of the independent variables;
- intercepts of the dependent variables;
- means of the independent variables;

Estimation of Latent Variable Means

As with the invariance applications presented in Chapters 7 and 9, this application of a structured means model involves testing simultaneously across two groups.¹ The multigroup model illustrated in this chapter is used when one is interested in testing for group differences in the means of particular latent constructs. This approach to the estimation of latent mean structures was first brought to light in Sörbom's (1974) seminal extension of the classic model of factorial invariance. As such, testing for latent mean differences across groups is made possible through the implementation of two important strategies—*model identification* and *factor identification*.

Model identification. Given the necessary estimation of intercepts associated with the observed variables, in addition to those associated with the unobserved latent constructs, it is evident that the attainment of an overidentified model is possible only with the imposition of several specification constraints. Indeed, it is this very issue that complicates, and ultimately renders impossible, the estimation of latent means in single-group analyses. Multigroup analyses, on the other hand, provide the mechanism for imposing severe restrictions on the model such that the estimation of latent means is possible. More specifically, because two (or more) groups under study are tested simultaneously, evaluation of the identification criterion is considered across groups. As a consequence, although the structured means model may not be identified in one group, it can become so when analyzed within the framework of a multigroup model. This outcome occurs as a function of specified equality constraints across groups. More specifically, these equality constraints derive from the underlying assumption that both the observed variable intercepts and the factor loadings are invariant across groups.

Factor identification. This requirement imposes the restriction that the factor intercepts for one group are fixed to zero; this group then operates as a reference group against which latent means for the other group(s) are compared. The reason for this reconceptualization is that when the intercepts of the measured variables are constrained equal across groups this leads to the latent factor intercepts having no definite origin (i.e., they are undefined in a statistical sense). A standard way of fixing the origin, then, is to set the factor intercepts of one group to zero (see Bentler, 2005; Jöreskog & Sörbom, 1996). As a consequence, factor intercepts (i.e., factor means) are interpretable only in a relative sense. That is to say, one can test whether the latent variable means for one group *differ* from those of another, but one *cannot* estimate the mean of each factor in a model for each group. In other words, while it is possible to test for latent mean differences between say, adolescent boys and adolescent girls, it is not possible to estimate the mean of each factor for both boys and girls

simultaneously; the latent means for one group must be constrained to zero.

Having reviewed the conceptual and statistical underpinning of the mean structures model, I now introduce you to the hypothesized model under study in this chapter.

The Hypothesized Model

The application to be examined in this chapter addresses equivalency of the latent factor means related to four self-concept (SC) dimensions (general, academic, English, mathematics) for high ($n = 582$) and low ($n = 248$) academically tracked high school students (Byrne, 1988b). The substantive focus of the initial study (Byrne, 1988b) was to test for latent mean differences in multidimensional SCs across these two ability groups. This CFA model followed from an earlier study designed to test for the multidimensionality of SC (Byrne & Shavelson, 1986) as portrayed schematically in Figure 8.1.

As you will note in the figure, except for Academic SC, the remaining dimensions are measured by three indicator variables, each of which represents a subscale score. Specifically, General SC is measured by subscale scores derived from the General SC subscale of the Self Description Questionnaire III (SDQIII; Marsh, 1992b), the Affective Perception Inventory (API; Soares & Soares, 1979), and the Self-esteem Scale (SES; Rosenberg, 1965); these indicator variables are labeled as SDQGSC, APIGSC, and SESGSC, respectively. English SC is measured by subscale scores related to the SDQIII (SDQESC), the API (APIESC), and the Self-concept of Ability Scale (SCAS; Brookover, 1962), labeled as SCAESC in Figure 8.1. Finally, Math SC is measured by subscale scores derived from the SDQIII (SDQMSC), the API (APIMSC), and the SCAS (SCAMSC). In the case of academic SC, findings from a preliminary factor analysis of the API (see Byrne & Shavelson, 1986) revealed several inadequacies in its measurement of this SC dimension. Thus, it was deleted from all subsequent analyses in the Byrne & Shavelson (1986) study and the same holds true here.

In contrast to the CFA model discussed in Chapter 7, in which the items of a measuring instrument formed the units of measurement, the CFA model under study in this chapter entails subscale scores of measuring instruments as its unit of measurement. It is hypothesized that each subscale measure will have a nonzero loading on the SC factor it is designed to measure albeit a zero loading on all other factors, and that error/uniquenesses associated with each of the observed measures are uncorrelated. Consistent with theory and empirical research, the four SC factors are shown to be intercorrelated.

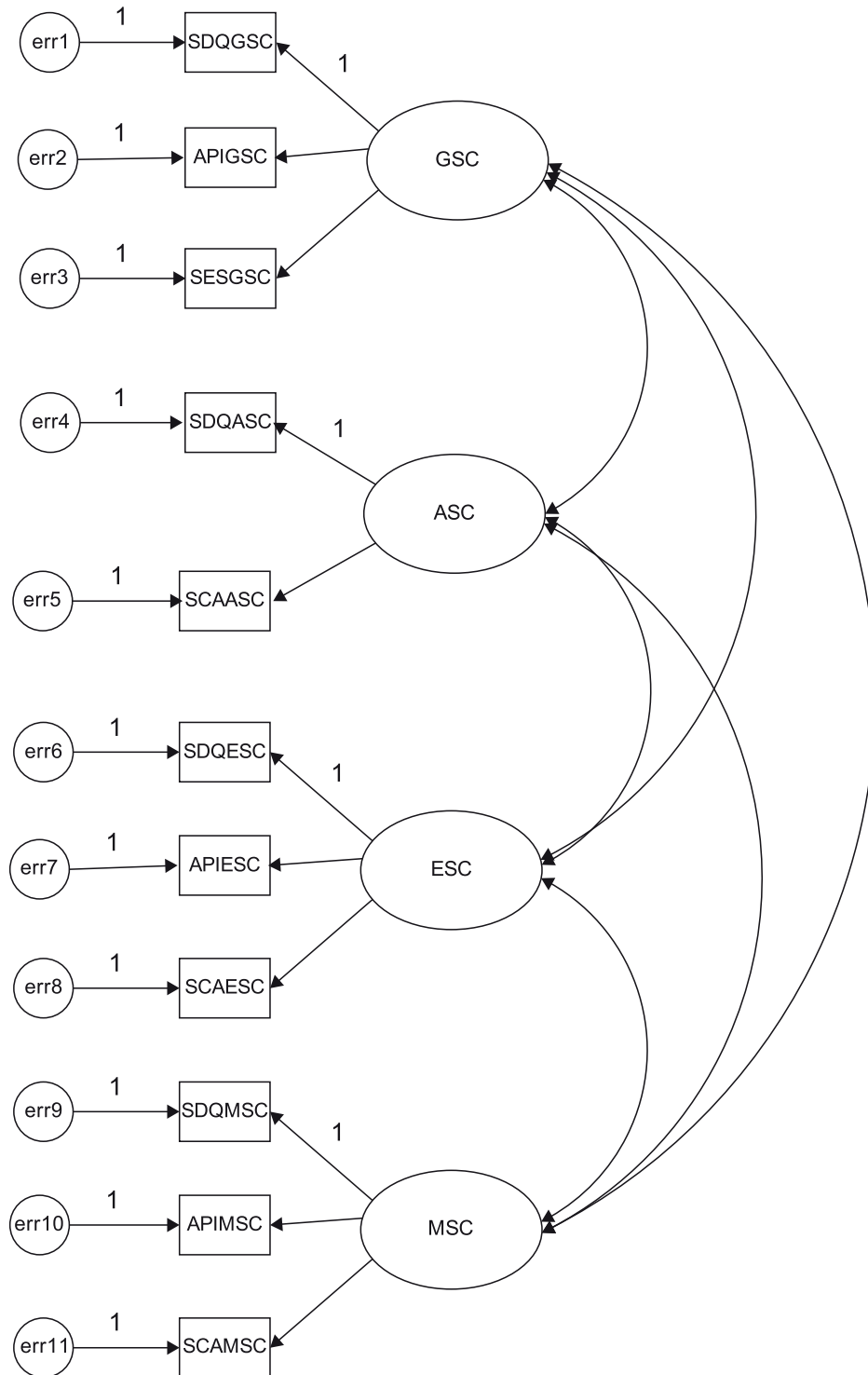


Figure 8.1 Hypothesized four-factor model of adolescent self-concept.

The Baseline Models

As with the example of multigroup invariance across independent samples in Chapter 7, goodness-of-fit related to the hypothesized model (Figure 8.1) was tested separately for high- and low-track students. Model fit statistics indicated only a modestly well-fitting model for both groups (high-track, CFI = .923, RMSEA = .128; low-track, CFI = .911, RMSEA = .114). Indeed, a review of the modification indices, for both groups, revealed substantial evidence of misspecification as a consequence of error covariances among subscales of the same measuring instrument—the SCAS. This finding of overlapping variance among the SCAS subscales is certainly not surprising and can be explained by the fact that items on the English and Math SC subscales were spawned from those comprising the Academic SC subscale. More specifically, the SCAS was originally designed to measure only Academic SC. However, in their attempt to measure the subject-specific facets of English and Math SCs, Byrne and Shavelson (1986) used the same items from the original SCAS, albeit modified the content to tap into the more specific facets of English SC and Math SC.

Given both a substantively and psychometrically reasonable rationale for estimating these three additional parameters, the originally hypothesized model was respecified and reestimated accordingly for each group. Testing of these respecified models resulted in a substantially better-fitting model for both high-track (CFI = .975; RMSEA = .076) and low-track (CFI = .974; RMSEA = .065) students. This final baseline model (which turns out to be the same for each group) serves as the model to be tested for its equivalence across high- and low-track students; it is schematically portrayed in Figure 8.2.

Modeling with Amos Graphics

The Structured Means Model

In working with *Amos Graphics*, the estimation and testing of structured means models is not much different from that of testing for invariance based on the analysis of covariance structures. It does, however, require a few additional steps. As with the multigroup application presented in Chapter 7, the structured means model requires a system of labeling whereby certain parameters are constrained to be equal across groups, while others are free to take on any value. In testing for differences in factor latent means, for example, we would want to know that the measurement model is operating in exactly the same way for both high- and low-track students. In the structured means models, this requirement

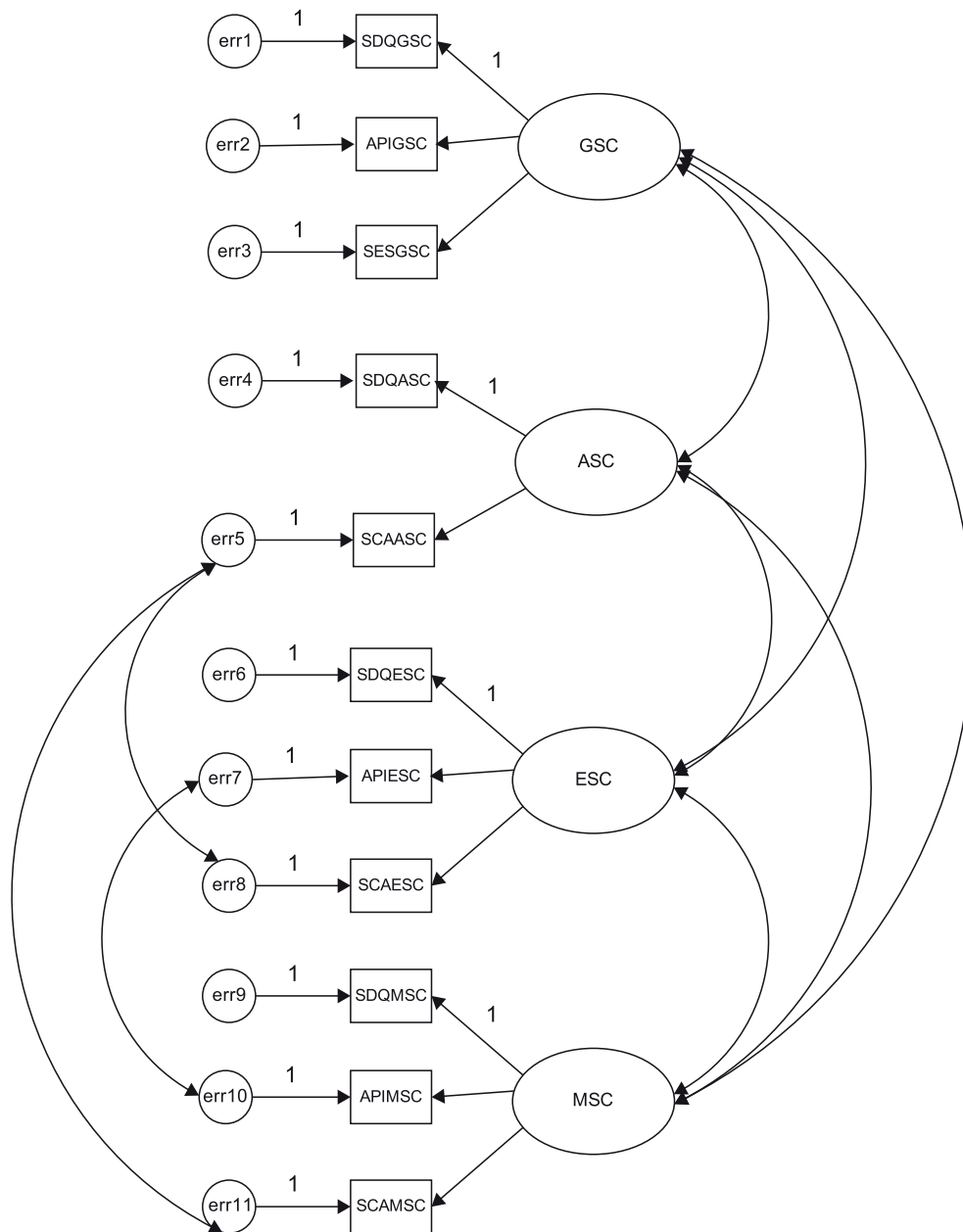


Figure 8.2 Baseline model adolescent self-concept for both the high-track and low-track groups.

includes the observed variable intercepts in addition to the factor loadings (i.e., regression weights) and, thus, both are constrained equal across both groups. We turn now to a more detailed description of this process as it relates to testing for latent mean differences.

Testing for Latent Mean Differences

The Hypothesized Multigroup Model

In testing for latent mean differences using *Amos Graphics*, the baseline model for each ability group must be made known to the program. However, in the case of our academically-tracked groups in this chapter, the final model for each was the same. Thus, the multigroup model shown in Figure 8.2 represents both groups.²

Steps in the Testing Process

Once the multigroup model (i.e., the configural model) is established, the next step (illustrated in Chapter 7) is to identify the name of each group (via the *Manage Groups* dialog box), as well as the location of the related data (via the *Data Files* dialog box). For a review of these procedures, see Chapter 7, and in particular, Figures 7.4 and 7.5. At this point Amos has all the information it requires with respect to both the model to be tested and the name and location of the data to be used. All that is needed now is to determine which analytic approach will be used in testing for differences in latent factor means across the two groups. In Chapter 7, procedures associated with only the selection of approach (manual versus automated) were illustrated. Details related to the process once this choice had been made were necessarily limited to the manual multigroup strategy as it applied to tests for invariance. In this chapter, by contrast, I focus on details related to the automated multigroup approach. Accordingly, clicking on the *Multiple-Group Analysis* tab on the *Analyze* menu, yields the labeled model shown in Figure 8.3. Although this model represents the same baseline model for the high-track and low-track groups, distinction between the two lies in the labeling of parameters. Specifically, all labeling for the high-track group is identified by “_1”, whereas labeling for the low-track group is identified by a “_2”. To review the labeling for each group, click on the selected group name listed in the left column of the model graphical file as shown in Figure 8.3. Here we see the labeled model for the high-track group (Group 1). Essentially, then, this model represents the configural model.

Testing for configural invariance. Recall from Chapter 7 that all tests for invariance begin with the configural model for which interest focuses on the extent to which the same number and pattern of factors best represents the data for both groups. As such, no equality constraints are imposed and judgment is based on the adequacy of the goodness-of-fit statistics only. In this regard, the configural model was found to be exceptionally well-fitting in its representation of the multigroup student data ($\chi^2_{(70)} = 225.298$; CFI = .975; RMSEA = .052).

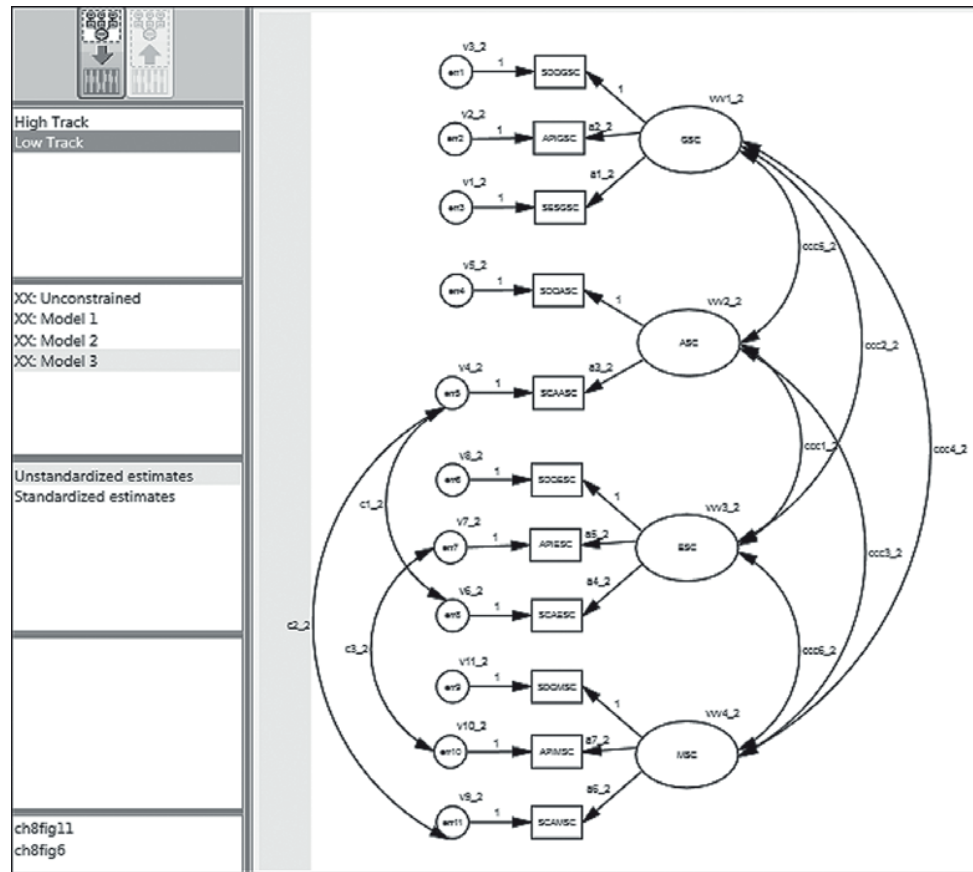


Figure 8.3 Amos Graphics: Factor loading estimates for the high-track group based on the unconstrained (i.e., configural) model.

To familiarize you with the Amos output directory tree in multigroup analyses, I include here both the unstandardized and standardized factor loading estimates for the configural model. Figure 8.4 is pertinent to the high-track group, while Figure 8.5 is pertinent to the low-track group. In reviewing both of these figures, I direct your attention to four important pieces of information. *First*, turning to the directory tree shown in the first section on the left side of Figure 8.4 you will note below that “High Track” is highlighted, thereby indicating that all results presented in this output file pertain only to this group. In contrast, if you were to click on “Low Track,” results pertinent to that group would be presented. *Second*, below this group identification section of the tree, you will see a list of four models (unconstrained + three constrained). As indicated by the cursor at the bottom of this section, results presented here (and in Figure 8.5) relate to the unconstrained model. *Third*, because there are no constraints specified in this model, the parameters are freely

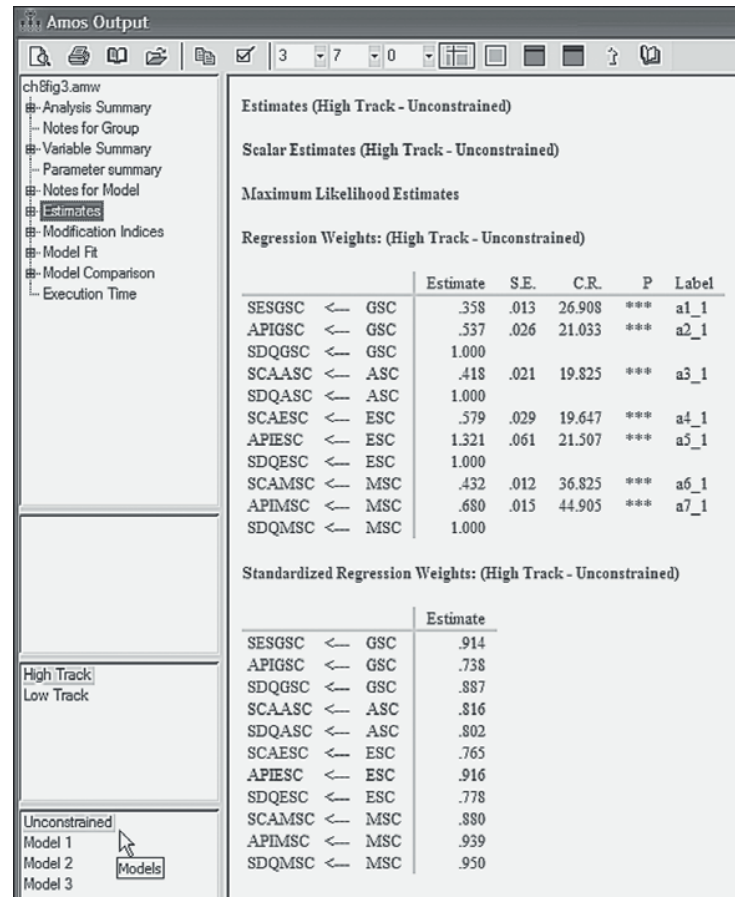


Figure 8.4 Amos Graphics: Factor Loading estimates for the high-track group based on the unconstrained model.

estimated and, thus, vary across high- and low-track students. *Finally*, although equality constraints were not assigned to the factor loadings, the program automatically assigned labels that can be used to identify both the parameters and the groups in subsequent tests for invariance. These labels appear in the last column of the unstandardized estimates related to the *Regression weights*. For example, the label assigned to the factor loading of SESGSC on GSC is a1_1 for high-track students and a1_2 for low-track students.³

Testing for measurement invariance. Subsequent to the configural model, all tests for invariance require the imposition of equality constraints across groups. Implementation of this process, using the automated approach, begins with the hypothesized model open followed by selection of the *Multiple-Group Analysis* tab from the *Analysis* drop-down menu as illustrated in the previous chapter (Figure 7.13). Clicking on this selection

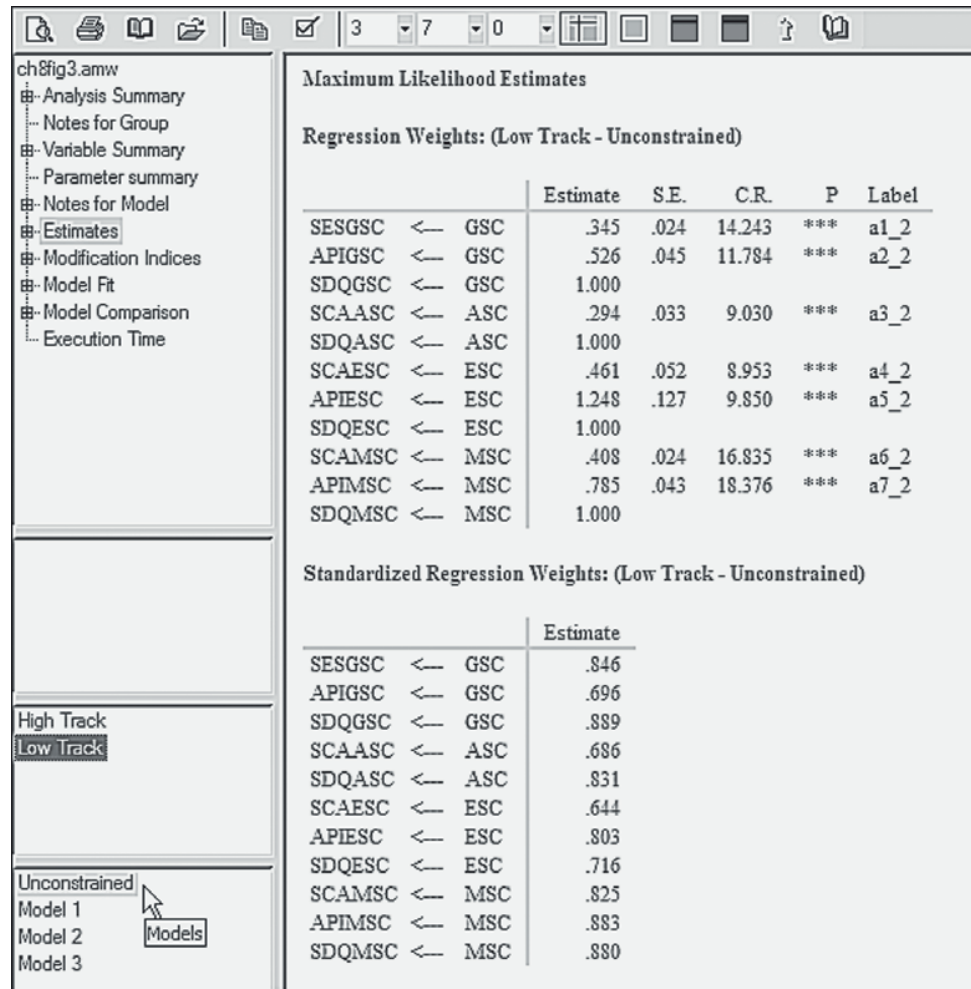


Figure 8.5 Amos Graphics: Factor loading estimates for the low-track group based on the unconstrained model.

will present you with the warning message shown overlapping the *Amos Graphics*-labeled configural model in Figure 8.6. However, once you click on the OK button, you will immediately be presented with the *Multiple-Group Analysis* dialog box in which the first five columns show default parameter subsets that are checked. Some of the boxed checkmarks will be in bolded print, while others will appear faded, indicating that they are inactive, but cannot manually be deleted. In particular, the latter marks appear in columns 3 to 5 and represent structural weights, structural intercepts, and structural residuals, respectively, which are default parameters pertinent to full SEM models only. As I emphasized in Chapter 7, in testing for the invariance of all models, I consider it prudent to test first for invariance

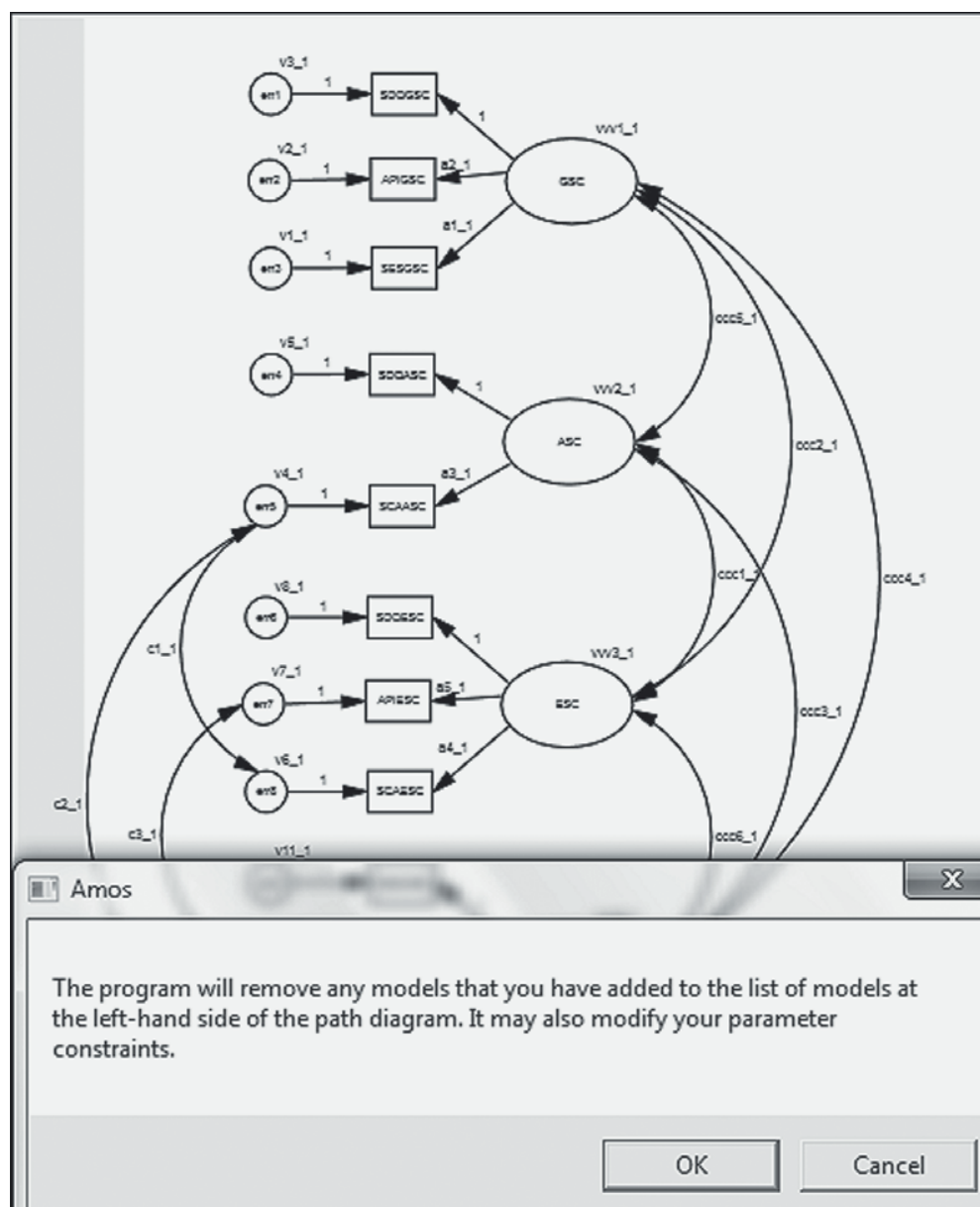


Figure 8.6 Warning message associated with selection of the *Multiple-Group Analysis* dialog box.

related to only the factor loadings. In this way, should the results yield evidence of noninvariance, you can then conduct follow-up analyses to determine which factor loading parameters are not operating the same way across groups and can exclude them from further analyses. Thus, in testing for invariance related only to the factor loadings, you will need

to delete all bolded checkmarks from the *Multiple-Group Analysis* dialog box, except for the one in Column 1 (see Figure 8.7) representing only the “measurement weights.”

Once you click on the OK button of this dialog box, you will then see the assignment of constraint labels to your model. As shown in Figure 8.3, our hypothesized multigroup model with the appropriate factor loading estimates labeled (a2_2 to a6_2), in addition to those representing the factor variances (vvv1_2 to vvv4_2), factor covariances (ccc1_2 to ccc6_2), and error variances (v1_2 to v11_2) as they pertain to Group 2, the low-track group. Likewise, labeling for the high-track (Group 1), is retrieved by clicking on the related group tab in the Amos output directory tree. As such the labeling system uses the number “1” to indicate its relevance to this group of students (e.g., a2_1; a3_1).

Goodness-of-fit results from this test of invariant factor loadings again provided evidence of a well-fitting model ($\chi^2_{(77)} = 245.571$; CFI = .972; RMSEA = .052). Although the difference in χ^2 from the configural model was statically significant ($\Delta\chi^2_{(7)} = 20.273$), the difference between the CFI values met the recommended cutoff criterion of .01 ($\Delta\text{CFI} = .003$). Using the CFI difference test as the criterion upon which to determine evidence of invariance, I concluded the factor loadings to be operating similarly across high-track and low-track students.

Parameter Subsets	1	2	3	4	5	6	7	8
Measurement weights	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Measurement intercepts	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Structural weights	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Structural intercepts	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Structural means	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Structural covariances	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Structural residuals	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Measurement residuals	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Help Default OK Cancel

Figure 8.7 Amos Graphics: Multiple group analysis dialog box with equality constraints indicated for factor loadings only.

Testing for latent mean differences. As noted earlier, in testing for differences in latent factor means, it is necessary to constrain both the factor loadings and the observed variable intercepts equal across the groups. Consistent with the case for factor loadings, given findings of noninvariance pertinent to the intercepts, it becomes necessary to then conduct additional analyses to determine exactly which intercept or combination of intercepts is contributing to the inequality of intercepts across groups. As with the factor loadings, implementation of partial measurement invariance can be used provided that there is more than one invariant intercept over and above the reference variable (fixed to 1.00) for each latent factor (see Byrne et al., 1989; Muthén & Muthén, 1998–2012, p. 486). Admittedly, there has been, and still is much debate (see Chapter 7) and confusion concerning the issue of partial measurement invariance, particularly as it applies to the intercepts in the ultimate task of testing for latent mean differences. However, clearly it makes little sense to maintain equality constraints on intercepts that are not, in fact, operating equivalently across the groups under test.

Our first step in testing for latent mean differences, then, is to constrain the intercepts equal across groups. This task is easily accomplished by activating the *Analysis Properties* dialog box, either by clicking on its related icon or by selecting it from the *Analysis* drop-down menu. Once the *Analysis Properties* dialog box is open, we click on the *Estimation* tab and then select the *Estimate means and intercepts* option as shown in Figure 8.8.

The next step in the process is to once again select the *Multiple-Group Analysis* option from the *Analysis* drop-down menu. This time, however, in addition to placing a checkmark in Column 1 for only the measurement weights, we additionally check off Column 2, which incorporates both the measurement weights and the measurement intercepts. Once these choices have been initiated, Amos automatically assigns a zero followed by a comma (0,) to each factor. Figure 8.9 captures this labeling action by showing both the options checked in the *Multiple-Group Analysis* dialog box, together with the resulting assignment of “0,” to the General SC (GSC) and Academic SC (ASC) factors.

The fixed zero values shown in Figure 8.9 are assigned to the model relevant to each group. However, as discussed at the beginning of this chapter, in testing for latent mean differences, one of the groups is freely estimated while the other is constrained equal to some fixed amount. In the case here, Amos automatically fixes this zero value for *both* groups. Thus, the next step in the process is to remove these fixed factor values for one of the groups. The decision of which group will be fixed to zero is an arbitrary one and has no bearing on the final estimated mean values; regardless of which group is chosen, the results will be identical. In the present case, I elected to use the low-track group as the reference group

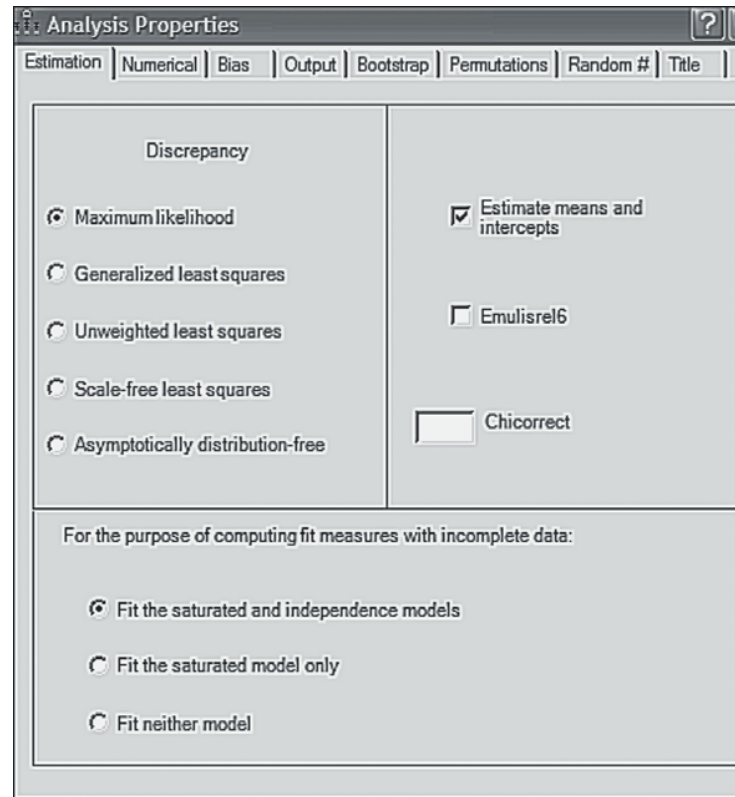


Figure 8.8 Amos Graphics: Analysis Properties dialog box with means and intercept estimates requested.

(i.e., the latent means were fixed to a value of 0.0) and thus, moved on to removing the mean constraints for the high-track group. With the model open, this process begins by first highlighting the factor to be re-labeled by means of a left click of the mouse and then right-clicking on this factor, which will activate the *Object Properties* tab as shown in Figure 8.10.

Clicking on this tab opens the *Object Properties* dialog box, after which we click on the *Parameters* tab; this action enables the re-labeling of the mean parameters. Our interest here is in removing the fixed zero value assigned to each of the factor means and replacing them with a label that allows these parameters to be freely estimated for the high-track group; Figure 8.11 captures this process. More specifically, when the dialog box was first opened, the label seen in the space below “mean” was 0. I subsequently replaced this number with an appropriate label (e.g., *mn_gsc*) that could be modified in accordance with its related factor as can be seen on the model behind. Circled within the dialog box, you can see assignment of the label to be assigned to the first factor (GSC). Note also, that

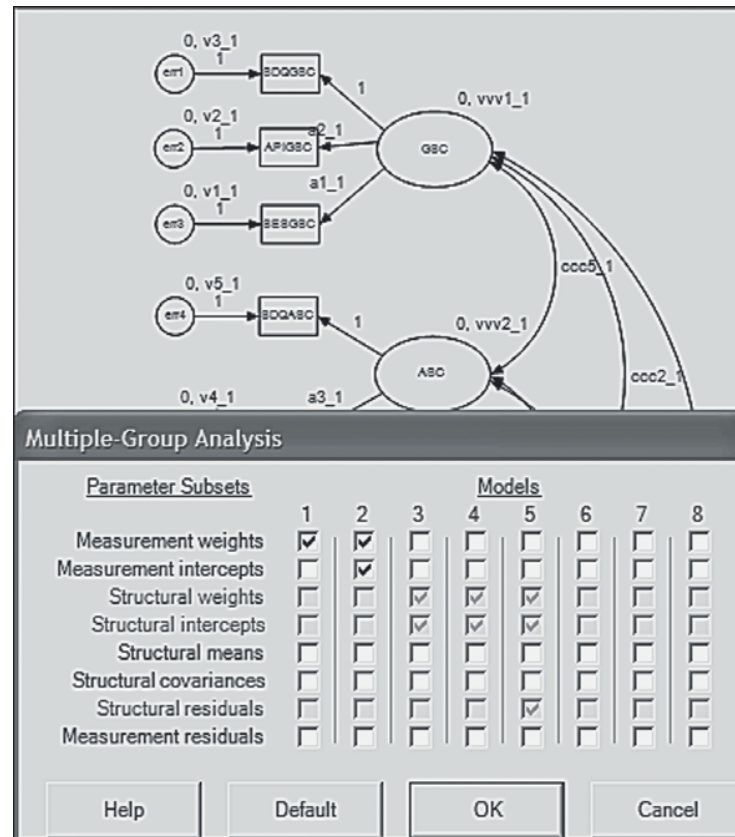


Figure 8.9 Amos Graphics: Multiple-Group Analysis dialog box with equality constraints indicated for factor loadings and intercepts.

the square beside “All groups” is empty and remains empty as we do not wish these re-labeling changes to be applied to both groups.⁴

The final model to be tested for latent mean differences is shown in Figure 8.12 as it relates to the high-track group. However, clicking on the low-track label in the directory tree presents the same model, albeit with the zero values assigned to each of the factors. A mini version of both models is displayed in Figure 8.13.

Selected Amos Output: Model Summary

Of primary interest in analyses related to structured means models are (a) the latent mean estimates, and (b) the goodness-of-fit between the hypothesized model and the multigroup data. Before turning to these results for our analysis of high and low-track students, let's once again take a few minutes to review a summary of the model in terms of the number of estimated parameters and resulting degrees of freedom. This summary is presented in Figure 8.14.

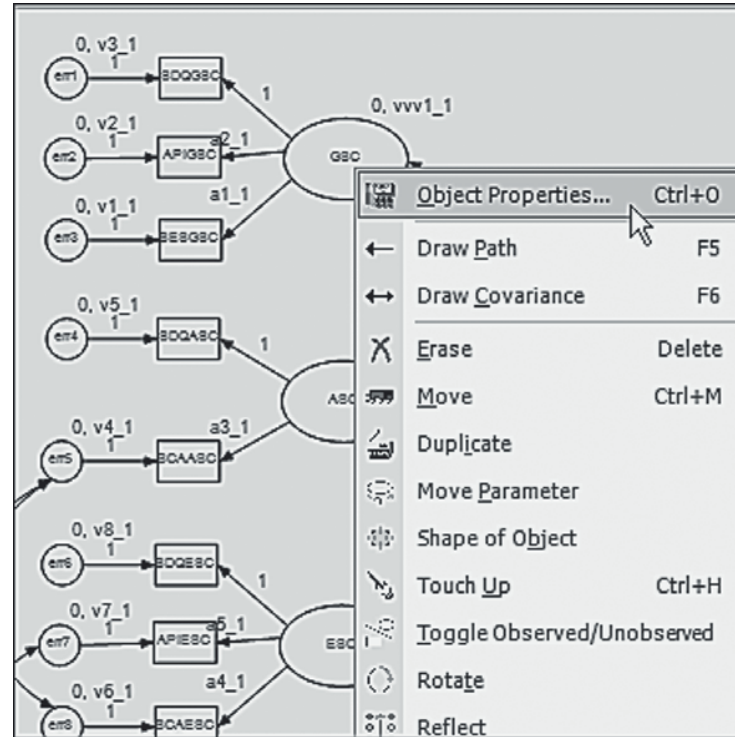


Figure 8.10 Amos Graphics: Clicking on the *Object Properties* tab to open the dialog box.

As you know by now, in order to calculate the degrees of freedom associated with the test of a model, you need to know two pieces of information: (a) the number of sample moments, and (b) the number of estimated parameters. In Figure 8.14, we see that there are 154 distinct sample moments and 70 estimated parameters. Let's turn first to the number of sample moments. If you had calculated the number of moments in the same manner used with all other applications in this book, you would arrive at the number "132" $[(11 \times 12)/2] = 66$; given two groups, the total is 132). Why this discrepancy? The answer lies in the analysis of covariance versus mean structures. All applications prior to the present chapter were based on covariance structures and, thus, the only pieces of information needed for the analyses were the covariances among the observed variables. However, in the analysis of structured means models, information related to both the covariance matrix and the sample means is required which, as noted earlier, constitutes the moment matrix. With 11 observed variables in the model, there will be 11 means; 22 means for the two groups. The resulting number of moments, then, is $132 + 22 = 154$.

Turning next to the reported 70 estimated parameters, let's look first at the number of parameters being estimated for the high-track group.

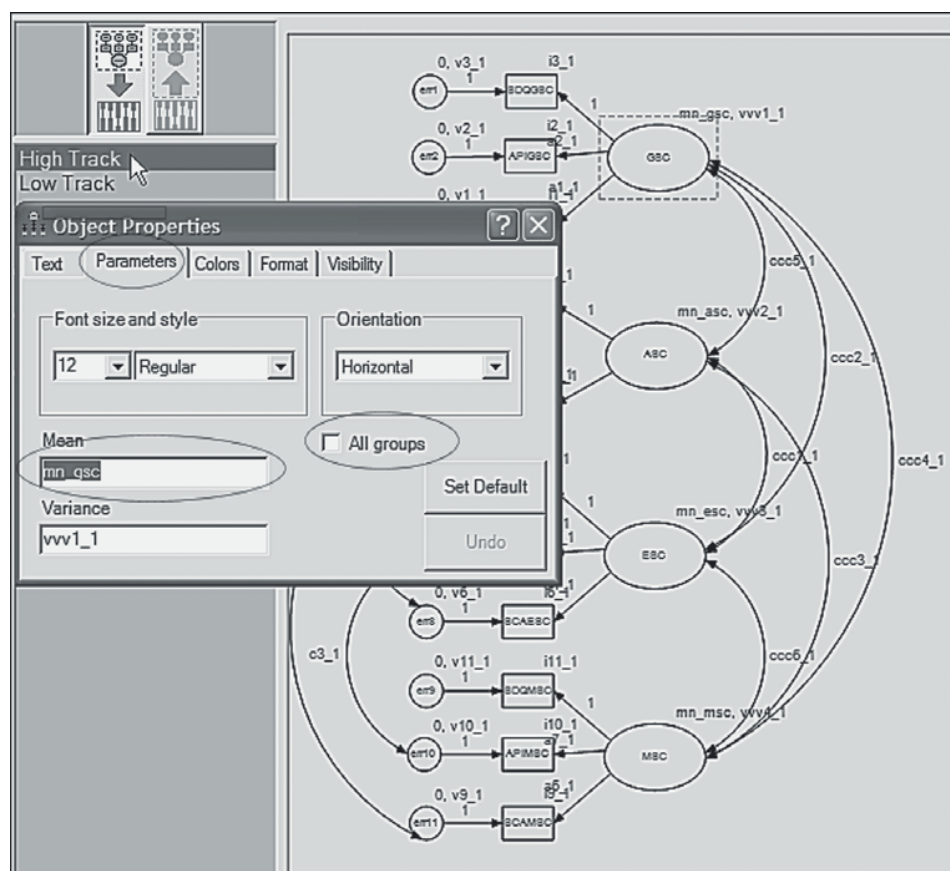


Figure 8.11 Amos Graphics: Object Properties dialog box showing modification of the means labeling protocol.

As such, we have: 7 factor loadings, 9 covariances (6 factor covariances; 3 error covariances), 15 variances (11 error variances; 4 factor variances), 11 intercepts, and 4 latent means, yielding a total of 46 parameters to be estimated. For the low-track group, on the other hand, we have: only 9 covariances and 15 variances, resulting in a total of 24 parameters to be estimated; the factor loadings and intercepts have been constrained equal to those of the high-track group, and the latent means have been fixed to 0.0. Across the two groups, then, there are 70 (46 + 24) parameters to be estimated. With 154 sample moments, and 70 estimated parameters, the number of degrees of freedom will be 84.⁵

Selected Amos Output: Goodness-of-fit Statistics

To provide you with a basis of comparison between the structured means model (Model 2 in the output) and both the configural model (i.e.,

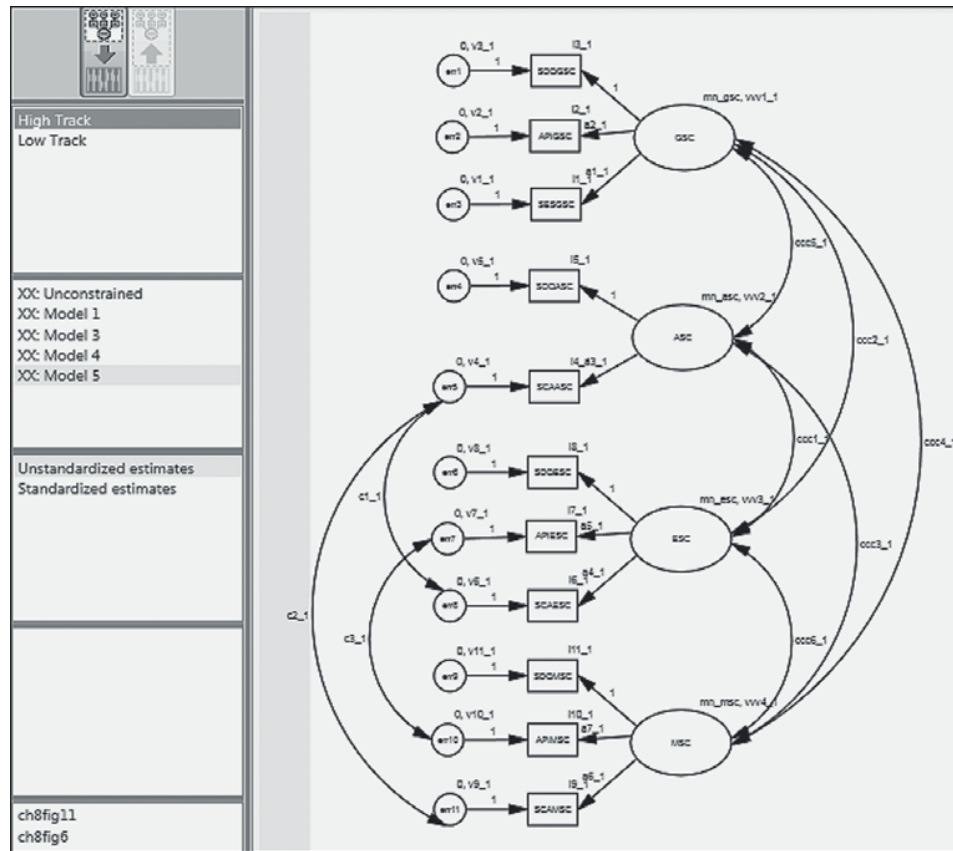


Figure 8.12 Structured means model as represented for the high-track group.

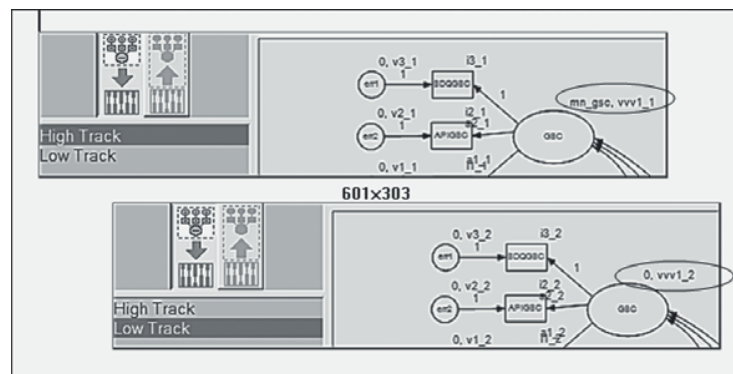


Figure 8.13 Micro view of structured means model for high-track and low-track groups.

ch8fig6.amw	
Analysis Summary	
Notes for Group	
Variable Summary	
Parameter summary	
Notes for Model	Notes for Model (Model 2)
Estimates	
Modification Indices	
Model Fit	
Execution Time	
	Computation of degrees of freedom (Model 2)
	Number of distinct sample moments: 154
	Number of distinct parameters to be estimated: 70
	Degrees of freedom (154 - 70): 84
	Result (Model 2)
	Minimum was achieved
	Chi-square = 309.456
	Degrees of freedom = 84
	Probability level = .000

Figure 8.14 Amos Graphics: Summary information related to structured means model.

unconstrained model) and measurement model in which only the factor loadings were group invariant (Model 1 in the output), goodness-of-fit statistics related to each are reported in Table 8.1. In each case, the fit statistics indicated well-fitting models. As reported earlier, although comparisons of Model 1 and 2 with the unconstrained model results in χ^2 -difference tests that were statistically significant ($p < .01$ and $p < .001$, respectively), the CFI-difference tests met the Cheung and Rensvold (2002) cutoff criteria of $<.01$ (with rounding in the case of Model 2). Indeed, despite the equality constraints imposed on both the factor loadings and the observed variable intercepts across the two groups, the structured means model fitted the data exceptionally well (e.g., CFI = .963) and demonstrated an adequate approximation to the two adolescent ability track populations (RMSEA = .057). Given these findings, then, we can feel confident in interpreting the estimates associated with the current solution.

Selected Amos Output: Parameter Estimates

High-track students. We turn first to parameter estimates for the high-track group, which are reported in Table 8.2. In the interest of space, factor and error variances are not included. A brief perusal of the critical ratios (C.R.s) associated with these estimates reveals all, except the covariance between the factors ESC and MSC, to be statistically significant. This nonsignificant finding, however, is quite consistent with self-concept theory as it relates to these two academic dimensions and therefore is no cause for concern.

Of major interest here are the latent mean estimates reported for high-track students as they provide the key to the question of whether the latent factor means for this group are significantly different from those for

Table 8.1 Goodness-of-fit Statistics for Configural and Measurement Models

Model	NPAR	CMIN	DF	P	CMIN/DF
Unconstrained	62	225.298	70	.000	3.219
Model 1	55	245.571	77	.000	3.189
Model 2	70	309.456	84	.000	3.684
Model 3	62	225.298	70	.000	3.219
Saturated model	132	.000	0		
Independence model	22	6213.470	110	.000	56.486

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Unconstrained	.964	.943	.975	.960	.975
Model 1	.960	.944	.973	.961	.972
Model 2	.950	.935	.963	.952	.963
Model 3	.964	.943	.975	.960	.975
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Unconstrained	.052	.044	.059	.338
Model 1	.051	.044	.059	.361
Model 2	.057	.050	.064	.045
Model 3	.052	.044	.059	.338
Independence model	.259	.253	.264	.000

low-track students. Given that the low-track group was designated as the reference group and thus their factor means were fixed to zero, the values reported here represent latent mean differences between the two groups. Reviewing these values we see that whereas the latent factor means related to the more specific facets of academic, English, and mathematics self-concepts were statistically significant (as indicated by the critical ratio values >1.96), this was not the case for general self-concept ($C.R.=.304$).

Given that the latent mean parameters were estimated for the high-track group, and that they represent positive values, we interpret these findings as indicating that, on average, high-track students in secondary school

Table 8.2 Selected Amos Output: Parameter Estimates for Means Structures
Model—High-track Students

			Estimate	S.E.	C.R.	P	Label
Regression Weights:							
SESGSC	<---	GSC	.356	.012	30.458	***	a1_1
APIGSC	<---	GSC	.535	.022	24.124	***	a2_1
SDQGSC	<---	GSC	1.000				
SCAASC	<---	ASC	.445	.018	25.030	***	a3_1
SDQASC	<---	ASC	1.000				
SCAESC	<---	ESC	.587	.026	22.252	***	a4_1
APIESC	<---	ESC	1.307	.054	24.158	***	a5_1
SDQESC	<---	ESC	1.000				
SCAMSC	<---	MSC	.428	.010	41.502	***	a6_1
APIMSC	<---	MSC	.698	.014	49.706	***	a7_1
SDQMSC	<---	MSC	1.000				
Standardized Regression Weights:							
SESGSC	<---	GSC	.913				
APIGSC	<---	GSC	.738				
SDQGSC	<---	GSC	.888				
SCAASC	<---	ASC	.826				
SDQASC	<---	ASC	.784				
SCAESC	<---	ESC	.769				
APIESC	<---	ESC	.913				
SDQESC	<---	ESC	.781				
SCAMSC	<---	MSC	.876				
APIMSC	<---	MSC	.943				
SDQMSC	<---	MSC	.947				
Means:							
GSC			.296	.978	.302	.762	mn_gsc
ASC			10.397	.781	13.315	***	mn_asc
ESC			3.424	.571	6.001	***	mn_esc
MSC			7.639	1.062	7.193	***	mn_msc
Intercepts:							
SESGSC			31.317	.291	107.449	***	i1_1
APIGSC			76.653	.474	161.804	***	i2_1
SDQGSC			75.638	.813	93.086	***	i3_1
SCAASC			25.424	.283	89.745	***	i4_1
SDQASC			47.894	.675	70.991	***	i5_1
SCAESC			26.459	.289	91.679	***	i6_1
APIESC			57.401	.606	94.746	***	i7_1

(continued)

Table 8.2 Selected Amos Output: Parameter Estimates for Means Structures
Model—High-track Students (*Continued*)

			Estimate	S.E.	C.R.	P	Label
Intercepts:							
SDQESC			54.382	.492	110.595	***	i8_1
SCAMSC			22.958	.364	63.034	***	i9_1
APIMSC			41.851	.585	71.518	***	i10_1
SDQMSC			41.444	.829	49.979	***	i11_1
Covariances:							
ASC	<-->	ESC	37.423	4.020	9.309	***	ccc1_1
GSC	<-->	ESC	23.971	4.773	5.022	***	ccc2_1
ASC	<-->	MSC	90.511	8.140	11.119	***	ccc3_1
GSC	<-->	MSC	54.373	9.556	5.690	***	ccc4_1
GSC	<-->	ASC	53.028	6.246	8.490	***	ccc5_1
ESC	<-->	MSC	.359	5.556	.065	.948	ccc6_1
err5	<-->	err8	5.159	.588	8.773	***	c1_1
err5	<-->	err11	4.546	.554	8.201	***	c2_1
err7	<-->	err10	8.590	1.406	6.111	***	c3_1
Correlations:							
ASC	<-->	ESC	.537				
GSC	<-->	ESC	.240				
ASC	<-->	MSC	.628				
GSC	<-->	MSC	.263				
GSC	<-->	ASC	.453				
ESC	<-->	MSC	.003				
err5	<-->	err8	.499				
err5	<-->	err11	.439				
err7	<-->	err10	.487				

appear to have significantly higher perceptions of self than their low-track peers with respect to perceived mathematics and English capabilities, as well as to overall academic achievement in general. On the other hand, when it comes to a global perception of self, there appears to be little difference between the two groups of students. Readers interested in a more detailed discussion of these results from a substantive perspective are referred to the original article (Byrne, 1988b). The rationale for this interpretation derives from the fact that the observed variable intercepts were constrained equal across groups and thus, the latent factors have an arbitrary origin (i.e., mean). As a consequence, the latent factor means are not estimated in an absolute sense, but rather, reflect average differences in the level of self-perceptions across these two groups (see Brown, 2006).

Table 8.3 Selected Amos Output: Parameter Estimates for Means Structure Model—Low-track Students

			Estimate	S.E.	C.R.	P	Label
Regression Weights:							
SESGSC	<---	GSC	.356	.012	30.458	***	a1_1
APIGSC	<---	GSC	.535	.022	24.124	***	a2_1
SDQGSC	<---	GSC	1.000				
SCAASC	<---	ASC	.445	.018	25.030	***	a3_1
SDQASC	<---	ASC	1.000				
SCAESC	<---	ESC	.587	.026	22.252	***	a4_1
APIESC	<---	ESC	1.307	.054	24.158	***	a5_1
SDQESC	<---	ESC	1.000				
SCAMSC	<---	MSC	.428	.010	41.502	***	a6_1
APIMSC	<---	MSC	.698	.014	49.706	***	a7_1
SDQMSC	<---	MSC	1.000				
Standardized Regression Weights:							
SESGSC	<---	GSC	.854				
APIGSC	<---	GSC	.698				
SDQGSC	<---	GSC	.880				
SCAASC	<---	ASC	.767				
SDQASC	<---	ASC	.694				
SCAESC	<---	ESC	.692				
APIESC	<---	ESC	.785				
SDQESC	<---	ESC	.684				
SCAMSC	<---	MSC	.851				
APIMSC	<---	MSC	.849				
SDQMSC	<---	MSC	.892				
Intercepts:							
SESGSC			31.317	.291	107.449	***	i1_1
APIGSC			76.653	.474	161.804	***	i2_1
SDQGSC			75.638	.813	93.086	***	i3_1
SCAASC			25.424	.283	89.745	***	i4_1
SDQASC			47.894	.675	70.991	***	i5_1
SCAESC			26.459	.289	91.679	***	i6_1
APIESC			57.401	.606	94.746	***	i7_1
SDQESC			54.382	.492	110.595	***	i8_1
SCAMSC			22.958	.364	63.034	***	i9_1
APIMSC			41.851	.585	71.518	***	i10_1
SDQMSC			41.444	.829	49.979	***	i11_1

(continued)

Table 8.3 Selected Amos Output: Parameter Estimates for Means Structure Model—Low-track Students (*Continued*)

			Estimate	S.E.	C.R.	P	Label
Covariances:							
ASC	<-->	ESC	31.828	5.002	6.363	***	ccc1_2
GSC	<-->	ESC	23.400	5.860	3.993	***	ccc2_2
MSC	<-->	ASC	49.353	8.598	5.740	***	ccc3_2
MSC	<-->	GSC	49.941	10.717	4.660	***	ccc4_2
GSC	<-->	ASC	41.082	8.136	5.050	***	ccc5_2
MSC	<-->	ESC	3.018	5.764	.524	.601	ccc6_2
err5	<-->	err8	5.051	1.050	4.808	***	c1_2
err5	<-->	err11	3.882	.863	4.499	***	c2_2
err7	<-->	err10	17.425	3.273	5.323	***	c3_2
Correlations:							
ASC	<-->	ESC	.620				
GSC	<-->	ESC	.317				
MSC	<-->	ASC	.496				
MSC	<-->	GSC	.349				
GSC	<-->	ASC	.428				
MSC	<-->	ESC	.040				
err5	<-->	err8	.432				
err5	<-->	err11	.396				
err7	<-->	err10	.510				

Low-track students. Turning now to the results for low-track students reported in Table 8.3, we find that, consistent with the high-track group, all parameter estimates are statistically significant except for the factor covariance between English and Math SC. Upon further scrutiny, however, you may think that something is amiss with this table as both the factor loadings (i.e., regression weights) and the observed variable intercepts, as well as their related labels are identical to those reported for the high-track group. These results are correct, of course, as these parameters (for the low-track group) were constrained equal to those estimated for the high-track group. Finally, it is important to note that no estimates for the factor means are reported. Although these values were constrained to zero, Amos does not report these fixed values in the output.

In the example application illustrated in this chapter, model comparisons based on CFI-difference values $<.01$ (Cheung & Rensvold, 2002) were used in determining measurement invariance of both the factor loadings and observed variable intercepts. However, if these model comparisons had resulted in CFI-difference values that had exceeded the cutpoint of

.01, it would have been necessary to determine which individual factor loadings and/or intercepts were noninvariant. This process was demonstrated in Chapter 7 with respect to factor loadings; the same approach is used for the intercepts. In walking you through both the manual and automated approaches to tests for invariance in Chapter 7 and 8, respectively, I hope that I have succeeded in helping you to feel sufficiently comfortable and confident in using either approach in analyses of your own data.

Notes

- 1 Of course, the procedure can also include more than two groups.
- 2 For an example of testing a multigroup model that involves two different baseline models, readers are referred to Byrne (2004).
- 3 Selection of which group serves as “Group 1” is purely arbitrary.
- 4 Although it was necessary to check this box when using the manual approach to invariance (see Chapter 7), this is not needed when the automated multigroup approach is used as Amos programs these constraints automatically.
- 5 It is important to note a serious error in the *Parameter Summary* table for this model as it relates to the Low-track Group, in the current version of the program (Amos 23). Turning first to the *Means* column, the correct values should report four *fixed* means, thus yielding a total of four means. Switching to a horizontal view of the *Parameter Summary* table, the *Fixed* row should show a total of 19 parameters, and the *Total* row, a grand total of 61 parameters. Calculation of these values is scheduled to be corrected in the next version of the program (J. Arbuckle, personal communication, April 3, 2015).