In: Kline, R. B. 2011: Principles and Practice of Structural Equation Modeling. 3. Auflage. New York: Guilford.

APPENDIX 9.B

Constraint Interaction in Measurement Models

Suppose that a researcher specifies a standard two-factor CFA model where the indicators of factors A are X_1 and X_2 and the indicators of factor B are X_3 and X_4. The sample covariance matrix where the order of the variables is X_1 to X_4 and $N = 200$ is as shown here:

$$\begin{bmatrix} 25.00 & & & \\ 7.20 & 9.00 & & \\ 3.20 & 2.00 & 4.00 & \\ 2.00 & 1.25 & 1.20 & 4.00 \end{bmatrix} \quad (\mathrm{I})$$

It is believed that the unstandardized loadings of X_2 and X_4 on their respective factors are equal. To test this hypothesis, an equality constraint is imposed on the unstandardized estimates, or

$$A \rightarrow X_2 = B \rightarrow X_4$$

and this restricted model is compared to the one without this constraint. Ideally, the value of $\chi^2_D(1)$ for this comparison should not be affected by how the factors are scaled, but this ideal is not realized for this example. If X_1 and X_3 are the reference variables for their respective unstandardized factors, then $\chi^2_D(1) = 0$. However, if instead the factor variances are fixed to 1.0 (standardized), then $\chi^2_D(1) = 14.017$ (calculated in LISREL) for the same comparison. (Try it!)

This unexpected result is an example of constraint interaction, which means that the value of the chi-square difference statistic for the test of the equality constraint depends on how the factors are scaled. It happens in this example because the imposition of the cross-factor equality constraint has the unintended consequence of making unnecessary one of the two identification constraints that scale the factors. However, removing the unnecessary identification constraint from the model with the equality constraint would result in two nonhierarchical models with equal degrees of freedom. That is, we could not conduct the chi-square difference test.

Steiger (2002) describes this test for constraint interaction: Obtain χ^2_M for the model with the cross-factor equality constraint. If the factors are unstandardized, fix the factor loading of the reference variable to a different constant, such as 2.0. If the factors are standardized, fix the variance of one of these factors to a constant other than 1.0. Fit the model so modified to the same data. If the value of χ^2_M for the modified model is not identical to that of the original, constraint interaction is present. If so, the choice of how to scale the factors should be based on substantive grounds. If no such grounds exist, the test results for the equality constraint may not be meaningful. See Gonzalez and Griffin (2001) for a discussion about how the estimation of standard errors in SEM is not always invariant to how the factors are scaled.

10

Structural Regression Models

Structural regression (SR) models are syntheses of path models and measurement models. They are the most general of all the core types of structural equation models considered to this point. As in path analysis, the specification of an SR model allows tests of hypotheses about effect priority. Unlike path models, though, these effects can involve latent variables because an SR model also incorporates a multiple-indicator measurement model, just as in CFA. The capability to simultaneously test hypotheses about both structural and measurement relations with a single model distinguishes SEM from other techniques. Discussed next are strategies for testing SR models. Also considered is the estimation of models where some latent variables have cause (formative) indicators instead of the more usual case where all factors have effect (reflective) indicators. The advanced techniques described in the next part of this book extend the rationale of SR models to other kinds of analyses.

ANALYZING SR MODELS

A theme common to the specification and identification of SR models (Chapters 5, 6) is that a valid measurement model is needed before it makes sense to evaluate the structural part of the model. This theme carries over to the two approaches to testing SR models described next. One is based on an earlier method by Anderson and Gerbing (1988) known as two-step modeling. A more recent method by Mulaik and Millsap (2000) is four-step modeling. Both methods generally require a fully latent SR where every variable in the structural model is a factor measured by multiple indicators. Both methods deal with the problem of how to locate the source of a specification error. An example follows.

Suppose that a researcher specified the three-factor SR model presented in Figure 10.1(a).[1] The data are collected and the researcher uses **one-step modeling** to estimate

[1]Only two indicators per factor are shown in Figure 10.1 to save space, but having at least three indicators per factor is better.

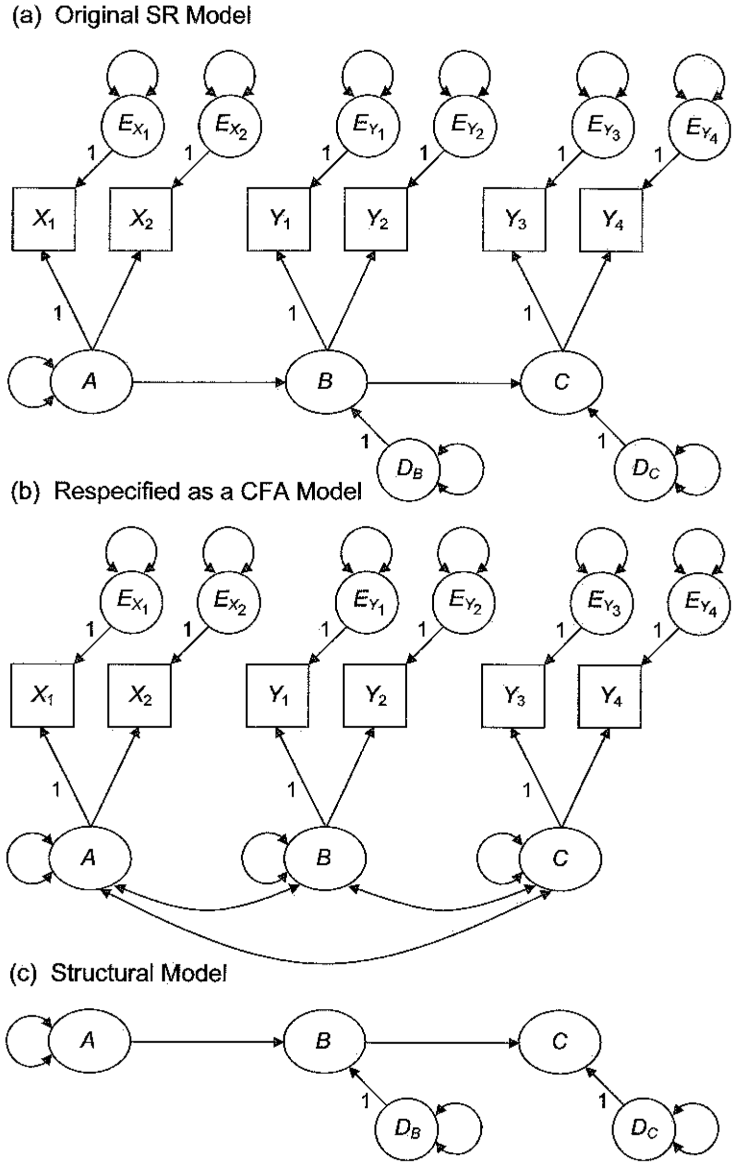

FIGURE 10.1. Two-step testing of a structural regression (SR) model.

this model, which means that the measurement and structural components of the SR model are analyzed simultaneously in a single analysis. The results indicate poor fit of the SR model. Now, where is the model misspecified? the measurement part? the structural part? or both? With one-step modeling, it is hard to precisely locate the source of poor fit. **Two-step modeling** parallels the two-step rule for the identification of SR models:

1. In the first step, an SR model is respecified as a CFA measurement model. The CFA model is then analyzed in order to determine whether it fits the data. If the fit of this CFA model is poor, then not only may the researcher's hypotheses about measurement be wrong, but also the fit of the original SR model may be even worse if its structural model is overidentified. Look again at Figure 10.1. Suppose that the fit of three-factor CFA model in Figure 10.1(b) is poor. Note that this CFA model has three paths among the factors that represent all possible unanalyzed associations (covariances, which are not directional). In contrast, the structural part of the original SR model, represented in Figure 10.1(c), has only two paths among the factors that represent direct effects. If the fit of the CFA model with three paths among the factors is poor, then the fit of the SR model with only two paths may be even worse. The first step thus involves finding an adequate measurement model. If this model is rejected, follow the suggestions in Chapter 9 about respecification of CFA models.
2. Given an acceptable measurement model, the second step is to compare the fits of the original SR model (with modifications to its measurement part, if any) and those with different structural models to one another and to the fit of the CFA model with the chi-square difference test. (This assumes that hierarchical structural models are compared.) Here is the procedure: If the structural part of an SR model is just-identified, the fits of that SR model and the CFA respecification of it are identical. These models are equivalent versions that generate the same predicted correlations and covariances. For example, if the path $A \rightarrow C$ were added to the SR model of Figure 10.1(a), then it would have just as many parameters as does the CFA measurement model of Figure 10.1(b). The SR model of Figure 10.1(a) with its overidentified structural model is thus nested under the CFA model of Figure 10.1(b). However, it may be possible to trim a just-identified portion of an SR model without appreciable deterioration in fit. Structural portions of SR models can be trimmed or built according to the same principles as in path analysis (Chapter 8).

Given an acceptable CFA measurement model, one should observe only slight changes in the factor loadings as SR models with alternative structural components are tested. If so, then the assumptions about measurement may be invariant to changes in the structural part of the SR model. But if the factor loadings change markedly when different structural models are specified, the measurement model is not invariant. This phenomenon may lead to **interpretational confounding** (Burt, 1976), which here means that the empirical definitions of the constructs (factor loadings) change depending on the structural model. It is generally easier to avoid interpretational confounding in two-step modeling than in one-step modeling.

Four-step modeling is basically an extension of the two-step method that is intended to even more precisely diagnose measurement model misspecification. In this strategy, the researcher specifies and tests a sequence of at least four hierarchical models. In order for these nested models to be identified, each factor in the original SR model should have at least four indicators. As in two-step modeling, if the fit of a model in four-step modeling with fewer constraints is poor, then a model with even more constraints should not even be considered. The steps are outlined next:

1. The least restrictive model specified at the first step is an EFA model—one based on a principal (common) factor analysis, *not* a principal components analysis—that allows each indicator to load on every factor and where the number of factors is the same as that in the original SR model. This EFA model should be analyzed with the same method of estimation, such as maximum likelihood (ML), as used to analyze the final SR model (at the fourth step). This first step is intended to test the provisional correctness of the hypothesis regarding the number of factors, but it cannot confirm that hypothesis if model fit is adequate (Hayduk & Glaser, 2000).
2. The second step of four-step modeling corresponds to the first step of two-step modeling: a CFA model is specified where some factor loadings (pattern coefficients) are fixed to zero. These specifications reflect the prediction that the indicator does not depend on that factor, not that the indicator is uncorrelated with that factor (i.e., the structure coefficient is not expected to equal zero). If the fit of the CFA model at the second step is acceptable, one goes on to test the original SR model. Otherwise, the measurement model should be revised.
3. The third step involves testing the SR model with the same set of zero pattern coefficients as represented in the measurement model from the second step but where at least one unanalyzed association from the second step is respecified as a direct effect or reciprocal effect and some of the factors are specified as endogenous. That is, the CFA measurement model of the second step is respecified as an SR model.
4. The last step involves tests of a priori hypotheses about parameters free from the outset of model testing. These tests typically involve the imposition of zero constraints, or dropping a path from the structural model. The third and fourth steps of four-step modeling are basically a more specific statement of activities that would fall under the second step of two-step modeling.

Which approach to analyzing SR models is better, two-step or four-step modeling? Both methods have their critics and defenders (e.g., Hayduk, 1996; Herting & Costner, 2000), and both capitalize on chance variation when hierarchical models are tested and respecified using the same data. The two-step method is simpler, and it does not require ≥ 4 indicators per factor. Both two-step and four-step modeling are better than one-step modeling, where there is no separation of measurement issues from structural issues. Neither method is a "gold standard" for testing SR models, but there is no such thing in SEM (Bentler, 2000). Bollen (2000) describes additional methods for testing SR models.

ESTIMATION OF SR MODELS

Discussed next are issues in the estimation of SR models.

Methods

The same estimation methods described in the previous chapters for path models and CFA models can be used with SR models. Briefly, standard ML estimation would normally be the method of choice for SR models with continuous indicators that are normally distributed. If the distributions are severely non-normal or the indicators are discrete with a small number of categories (e.g., Likert scale items as indicators), then one of the alternative methods described in Chapters 7 or 9 should be used instead.

Interpretation of Parameter Estimates and Problems

Interpretation of parameters estimates from the analysis of an SR model should not be difficult if one knows something about path analysis and CFA (and you do by now). For example, path coefficients are interpreted for SR models as regression coefficients for effects on endogenous variables from other variables presumed to directly cause them. Total effects among the factors that make up the structural model can be broken down into direct and indirect effects using the principle of effect decomposition (Chapter 7). Factor loadings are interpreted for SR models as regression coefficients for effects of factors on indicators, just as in CFA (Chapter 9).

Some SEM computer programs print estimated squared correlations (R^2_{smc}) for each endogenous variable. For SR models this includes the indicators and endogenous factors. Values of R^2_{smc} are usually computed for indicators in the unstandardized solution as one minus the ratio of the estimated measurement error variance over the sample variance of that indicator. Although variances of endogenous factors are not model parameters, they nevertheless have model-implied variances. Therefore, values of R^2_{smc} are usually calculated for endogenous factors as one minus the ratio of the estimated disturbance variance over the model-implied variance for that factor. Look out for Heywood cases, such as negative variance estimates, that suggest a problem with the data, specification, sample size, number of indicators per factor, or identification status of the model. If iterative estimation fails due to poor start values set automatically by the computer, the guidelines in Appendix 7.A can be followed for generating your own start values for the structural model or in Appendix 9.A for the measurement model.

Most SEM computer programs calculate a standardized solution for SR models by first finding the unstandardized solution with unit loading identification (ULI) constraints for endogenous factors and then transforming it to standardized form. Steiger (2002) notes that this method assumes that the ULI constraints function only to scale the endogenous variables. In other words, there is no constraint interaction. See Appendix 10.A for more information about constraint interaction in SR models.

DETAILED EXAMPLE

This example of the two-step analysis of an SR model of factors of job satisfaction was introduced in Chapter 5. Briefly reviewed, Houghton and Jinkerson (2007) measured within a sample of 263 full-time university employees three indicators each of constructive thinking, dysfunctional thinking, subjective well-being, and job satisfaction. They hypothesized that constructive thinking reduces dysfunctional thinking, which leads to an enhanced sense of well-being, which in turn results in greater job satisfaction. They also predicted that dysfunctional thinking directly affects job satisfaction. Their SR model with a standard four-factor, 12-indicator measurement model and an overidentified recursive structural model (i.e., the whole SR model is identified) is presented in Figure 5.9. We will first evaluate whether its measurement model is consistent with the data summarized in Table 10.1. All results described next are from converged, admissible solutions.

I submitted to Mplus 5.2 the correlations and standard deviations presented in Table 10.1, and Mplus converted these statistics to a sample covariance matrix. The first model I analyzed with ML estimation was a standard one-factor CFA model with 12 indicators. Values of selected fit statistics for this initial measurement model are reported in Table 10.2. It is clear that the fit of the one-factor measurement model is poor. For example, this model fails the chi-square test (χ^2_M (54) = 566.797, $p < .001$), the RMSEA with its 90% confidence interval is .190 (.176–.204), and the CFI is only .498. Next, I specified the measurement portion of the Houghton–Jinkerson SR model (Figure 5.9) as a standard four-factor CFA model. Values of selected fit statistics for this four-factor CFA model are also listed in Table 10.2. The model chi-square is not statistically significant—χ^2_M (48) = 62.468, p = .078—so the exact-fit hypothesis is not rejected. The relative improvement in fit of the four-factor CFA model over that of the one-factor CFA model is statistically significant—χ^2_D (6) = 504.329, $p < .001$—and values of approximate fit indexes for the four-factor model are generally favorable (e.g., RMSEA = .034; CFI = .986; Table 10.2).

TABLE 10.1. Input Data (Correlations, Standard Deviations) for Analysis of a Structural Regression Model of Thought Strategies and Job Satisfaction

Variable	1	2	3	4	5	6	7	8	9	10	11	12
Job satisfaction												
1. Work_1	1.00											
2. Work_2	.668	1.00										
3. Work_3	.635	.599	1.00									
Subjective well-being												
4. Happy	.263	.261	.164	1.00								
5. Mood_1	.290	.315	.247	.486	1.00							
6. Mood_2	.207	.245	.231	.251	.449	1.00						
Dysfunctional thinking												
7. Perform_1	–.206	–.182	–.195	–.309	–.266	–.142	1.00					
8. Perform_2	–.280	–.241	–.238	–.344	–.305	–.230	.753	1.00				
9. Approval	–.258	–.244	–.185	–.255	–.255	–.215	.554	.587	1.00			
Constructive thinking												
10. Beliefs	.080	.096	.094	–.017	.151	.141	–.074	–.111	.016	1.00		
11. Self-Talk	.061	.028	–.035	–.058	–.051	–.003	–.040	–.040	–.018	.284	1.00	
12. Imagery	.113	.174	.059	.063	.138	.044	–.119	–.073	–.084	.563	.379	1.00
SD	.939	1.017	.937	.562	.760	.524	.585	.609	.731	.711	1.124	1.001

Note. Input data are from Houghton and Jinkerson (2007); N = 263.

TABLE 10.2. Values of Selected Fit Statistics for Two-Step Testing of a Structural Regression Model of Thought Strategies and Job Satisfaction

Model	χ^2_M	df_M	χ^2_D	df_D	RMSEA (90% CI)	CFI	SRMR
Measurement model							
1-factor standard CFA	566.797[a]	54	—	—	.190 (.176–.204)	.498	.143
4-factor standard CFA	62.468[b]	48	504.329[a]	6	.034 (0–.056)	.986	.040
4-factor CFA with $E_{\text{Ha}} \leftrightarrow E_{\text{Mo}_2}$	56.662[c]	47	5.806[d]	1	.028 (0–.052)	.991	.037
Structural regression model							
Just-identified structural model (6 paths)	56.662[c]	47	—		.028 (0–.052)	.991	.037
Overidentified structural model (4 paths)	60.010[e]	49	3.348[f]	2	.029 (0–.052)	.989	.043

Note. CI, confidence interval.
[a]$p < .001$; [b]p = .078; [c]p = .158; [d]p = .016; [e]p = .135; [f]p = .188.

Close inspection of diagnostic information about the fit of the standard four-factor measurement indicated few apparent problems. For example, only two absolute correlation residuals (calculated in EQS) just exceeded .10, which is not a bad result in a larger model. There were a total of four standardized residuals (z statistics) with absolute values of about 2.00. Two of these larger residuals were for two different pairs among the three indicators of subjective well-being. One of the largest modification indexes (about 5.40) was for an error covariance between the indicators "Happy" (percent time happy) and "Mood_2" of this factor (see Figure 5.9).

Because it seems reasonable that shared item content across the two indicators just mentioned could be the basis for a common omitted cause, I respecified the four-factor measurement model by allowing the error covariance $E_{\text{Ha}} \leftrightarrow E_{\text{Mo}_2}$ to be freely estimated in a third analysis. Reported in Table 10.2 are values of selected fit statistics for this modified measurement model. Its fit to the data is statistically better than that of the standard four-factor model with no correlated errors (χ^2_D (1) = 5.806, p = .016). The exact-fit hypothesis is not rejected for the respecified measurement model (χ^2_M (47) = 56.662, p = .158). Values of other fit statistics are generally favorable (e.g., RMSEA = .028, SRMR = .037). Finally, no absolute correlation residuals exceeded .10.

Based on the results just described, the four-factor measurement model with an error correlation presented in Figure 10.2 was retained. In contrast, Houghton and Jinkerson's (2007) final measurement model was a standard four-factor model, so my conclusion differs somewhat from theirs. Reported in Table 10.3 are estimates of factor loadings and error variances for the measurement model in Figure 10.2. Values of the standardized factor loadings for indicators of some factors are uniformly high, which suggests convergent validity. For example, the range of these loadings for the job satisfaction factor is .749–.839. A few other standardized loadings are somewhat low, such as .433 for the self-talk indicator of constructive thinking, so evidence for convergent validity is mixed. Values of R^2_{smc} for indicators range from .188 to .817. (You should verify this statement based on the information in Table 10.3.)

Estimates of factor variances and covariances and of the sole measurement error covariance for the model of Figure 10.2 are listed in Table 10.4. Two-factor covariances are not statistically significant, including one for the pair of factors about thinking styles (constructive, dysfunctional) and the other for the association between constructive thinking and subjective well-being. Estimated factor correlations range from −.480 to .466. These moderate factor intercorrelations suggest discriminant validity. The sole error covariance (−.043) is statistically significant, and the corresponding correlation

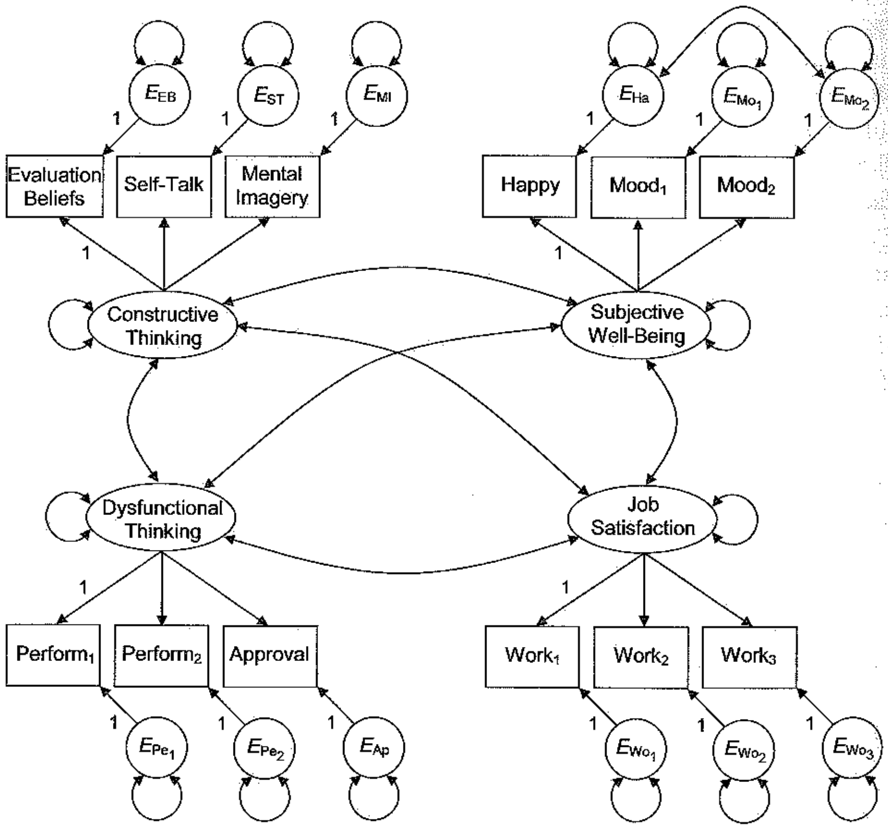

FIGURE 10.2. Measurement model for a structural regression model of thought strategies and job satisfaction.

TABLE 10.3. Maximum Likelihood Estimates of Factor Loadings and Residuals for a Measurement Model of Thought Strategies and Job Satisfaction

	Factor loadings			Measurement errors		
Indicator	Unst.	*SE*	St.	Unst.	*SE*	St.
Job satisfaction						
$Work_1$	1.000[a]	—	.839	.260	.042	.297
$Work_2$	1.035	.081	.802	.368	.050	.357
$Work_3$	.891	.073	.749	.384	.044	.439
Subjective well-being						
Happy	1.000[a]	—	.671	.173	.025	.550
$Mood_1$	1.490	.219	.739	.261	.044	.453
$Mood_2$	.821	.126	.591	.178	.022	.651
Dysfunctional thinking						
$Perform_1$	1.000[a]	—	.830	.106	.016	.311
$Perform_2$	1.133	.080	.904	.068	.017	.183
Approval	.993	.089	.660	.300	.029	.564
Constructive thinking						
Beliefs	1.000[a]	—	.648	.292	.043	.580
Self-Talk	1.056	.178	.433	1.022	.097	.812
Imagery	1.890	.331	.870	.242	.123	.242

Note. Unst., unstandardized; St., standardized. Standardized estimates for measurement errors are proportions of unexplained variance.

[a]Not tested for statistical significance. For all other unstandardized estimates, $p < .05$.

TABLE 10.4. Maximum Likelihood Estimates of Factor Variances and Covariances and Error Covariance for a Measurement Model of Thought Strategies and Job Satisfaction

Parameter	Unstandardized	*SE*	Standardized
	Factor variances and covariances		
Job Satisfaction	.618	.081	1.000
Subjective Well-Being	.142	.031	1.000
Dysfunctional Thinking	.235	.031	1.000
Constructive Thinking	.212	.049	1.000
Constructive ↔ Dysfunctional	−.028[a]	.017	−.124
Constructive ↔ Well-Being	.024[a]	.014	.140
Constructive ↔ Job Satisfaction	.060	.029	.165
Dysfunctional ↔ Well-Being	−.088	.017	−.480
Dysfunctional ↔ Job Satisfaction	−.131	.030	−.344
Well-Being ↔ Job Satisfaction	.138	.028	.466
	Error covariance		
Happy ↔ $Mood_2$	−.043	.018	−.243

[a]$p \geq .05$. For all other unstandardized estimates, $p < .05$.

is –.243. This correlation is not large, but its presence helps to "clean up" some local fit problem in parts of the standard four-factor measurement model without this parameter.

The analyses described next concern the second step of two-step modeling—the testing of SR models with the measurement model established in the first step but with alternative versions of the structural models. The first SR model analyzed is one with a just-identified structural component. Because this SR model and the CFA measurement model in Figure 10.2 have the same number of paths among the factors (6), they are equivalent models. This fact is verified by the observation of identical values of fit statistics for the two models just mentioned (see Table 10.2). Equivalence also implies that estimates of factor loadings and measurement error variances and covariance will be identical within rounding error for the two models. Accordingly, at this point we need to consider just the parameter estimates for the structural part of the SR model, which are unique to this model.

Estimates for the just-identified structural model are presented in Figure 10.3. The direct effects in the figure depicted with dashed lines were predicted by Houghton and Jinkerson (2007) to be zero. The unstandardized path coefficient for the direct effect of constructive thinking on dysfunctional thinking (–.131) is not statistically significant, and the corresponding standardized path coefficient (–.124) indicates a relatively small effect size. It is no surprise, then, that constructive thinking explains only about 1.5% of the variance in dysfunctional thinking (R^2_{smc} = .015). The other two unstandardized

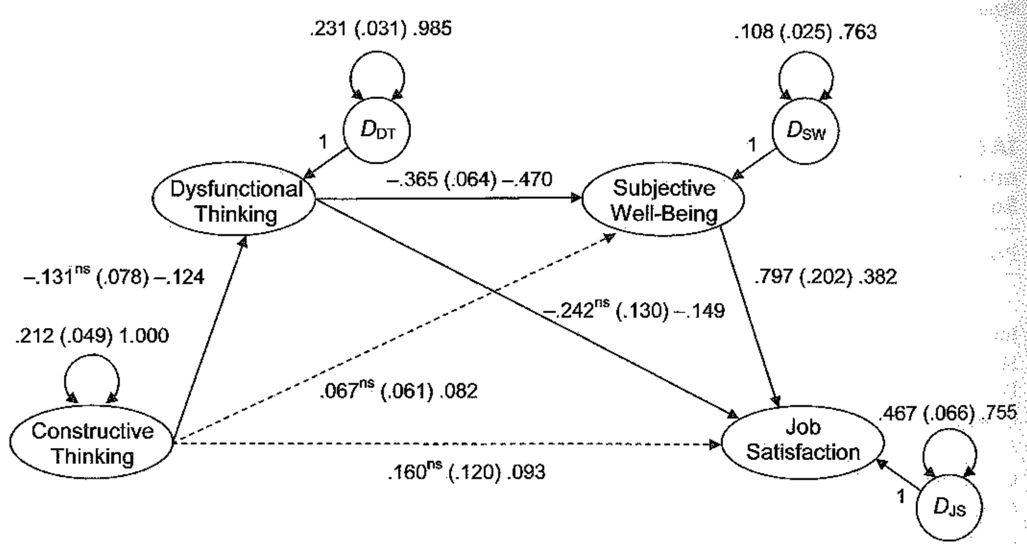

FIGURE 10.3. Structural model for a structural regression model of thought strategies and job satisfaction. Estimates are reported as unstandardized (standard error) standardized. Standardized estimates for disturbances are proportions of unexplained variance. The unstandardized estimates are all statistically significant at the .01 level except for those designated "ns," which means not significant.

path coefficients for constructive thinking, .067 and .160 for, respectively, direct effects on subjective well-being and job satisfaction, are also not statistically significant. This is consistent with predictions (Figure 10.3). Direct effects of dysfunctional thinking on subjective well-being and of subjective well-being on job satisfaction are both statistically significant and appreciable in standardized magnitude (respectively, –.470, .382). These results support the hypothesis that the effects of dysfunctional thinking strategies on job satisfaction are largely mediated through subjective well-being. Overall, about 25% of the variance in both the subjective well-being factor and job satisfaction factors is explained (the R^2_{smc} values are, respectively, .237 and .245; see Figure 10.3).

The final SR model retained by Houghton and Jinkerson (2007) had the four paths in the structural model represented with the solid lines in Figure 10.3. Values of selected fit statistics for this restricted SR model are reported in Table 10.2. The exact-fit hypothesis is not rejected for the restricted SR model (χ^2_M (49) = 60.010, p = .135), and its overall fit is not statistically worse than that of the unrestricted SR model with six direct effects (χ^2_D (2) = 3.348, p = .188). However, inspection of the correlation residuals (calculated in EQS) for the restricted SR model indicated some localized fit problems. For example, the correlation residual for the association between the "Work$_2$" indicator of job satisfaction and the "Imagery" indicator of constructive thinking is .142. Other absolute correlation residuals > .10 involved the "Beliefs" indicator of constructive thinking and both positive mood indicators of subjective well-being. Thus, dropping the two paths listed next:

Constructive Thinking → Job Satisfaction
Constructive Thinking → Subjective Well-Being

from the just-identified structural model in Figure 10.3 results in poor explanations of the observed correlations between the pairs of indicators just mentioned. This is an example of how dropping paths that are not statistically significant—here, from the structural model—can deteriorate the fit of some other parts of the model. Based on these results, I would retain the SR model with the just-identified structural model with six direct effects. You can download from this book's website (see p. 3) the Mplus syntax, data, and output files for the final four-factor measurement model (Figure 10.2) and the final SR model with six paths (Figure 10.3). Also available on the site are computer files for the same analyses in EQS and LISREL.

I used the Power Analysis procedure of STATISTICA 9 Advanced to estimate power for the final SR model with six paths in its structural model. Given N = 263, df_M = 47, and assuming α = .05 and ε_1 = .08, the power for the test of the close-fit hypothesis (H_0: $\varepsilon_0 \leq$ .05) is .869. Now assuming ε_1 = .01, the power for the test of the not-close-fit hypothesis (H_0: $\varepsilon_0 \geq$ .05) is .767. Thus, the probability of either rejecting a false model or detecting a correct model is quite good in this analysis, despite a sample size that is not large. This happens here because the relatively high degrees of freedom (47) for this larger model offset the negative impact of a smaller sample size on power.

EQUIVALENT SR MODELS

It is often possible to generate equivalent versions of SR models. An equivalent version of an SR model with a just-identified structural model was mentioned earlier: the measurement part of an SR model respecified as a CFA model, which assumes no causal effects among the factors, only unanalyzed associations (e.g., Table 10.2). Regardless of whether or not the structural model is just-identified, it may be possible to generate equivalent versions of it using the Lee–Hershberger replacing rules for path models (Chapter 8). For example, *any* rearrangement of the direct effects in the just-identified structural model in Figure 10.3 that respects these rules while holding the measurement model constant will result in alternative SR models that will fit the same data equally well. With the structural model held constant, it may also be possible to generate equivalent versions of the measurement model using Hershberger's reversed indicator rule, which involves reversing the direction of the causal effect between a factor and one of its indicators. That is, one indicator is specified as a cause indicator rather than as an effect indicator (Chapter 9). Given no change in the structural model, alternative SR models with equivalent measurement models would also fit the same data equally well. See Hershberger (1994) for more information and examples.

Equivalent versions of the just-identified structural model in Figure 10.3 for analysis of the Houghton and Jinkerson (2007) data include any other possible just-identified variation of this model. This includes structural models where the causal effects "flow" in the opposite direction, such as from job satisfaction to subjective well-being to dysfunctional thinking to construct thinking. Houghton and Jinkerson (2007) offered a detailed rationale of their original directionality specifications. But without such an argument, there is no way to prefer one just-identified structural model over an equivalent variation.

SINGLE INDICATORS IN PARTIALLY LATENT SR MODELS

At times a researcher has only one measure of some construct. Scores from a single indicator are unlikely to be both perfectly reliable and valid. There is an alternative to representing a single indicator in the structural part of an SR model as one would in path analysis (i.e., without a measurement error term). This alternative requires an a priori estimate of the proportion of variance in a single indicator that is due to measurement error (10%, 20%, etc.). This estimate may be based on the researcher's experience or on results of previous studies. Recall that (1) one minus a reliability coefficient, $1 - r_{XX}$, estimates the proportion of observed variance due to random error, which is only one source of measurement error (Chapter 3). (2) Specific types of reliability coefficients estimate only one kind of random error. Thus, the quantity $1 - r_{XX}$ may *underestimate* the proportion of total variance due to measurement error.

Suppose that X_1 is the only indicator of an exogenous factor A and that the researcher estimates that the 20% of X_1's variance is due to measurement error. Given this estimate, it is possible to specify an SR model like the one presented in Figure 10.4(a). Note that X_1 in the figure is specified as a single indicator *and* has an error term. The unstandardized variance of the latter is fixed to equal .20 times the observed variance, or $.2\,s^2_{X_1}$. For example, if the observed variance of X_1 is 30.00, then 20% of this variance, or .2 (30.00) = 6.00, is the estimated error variance. Because factor A must be scaled, the unstandardized loading of X_1 on A is fixed to equal 1.0. With the specification of a residual term

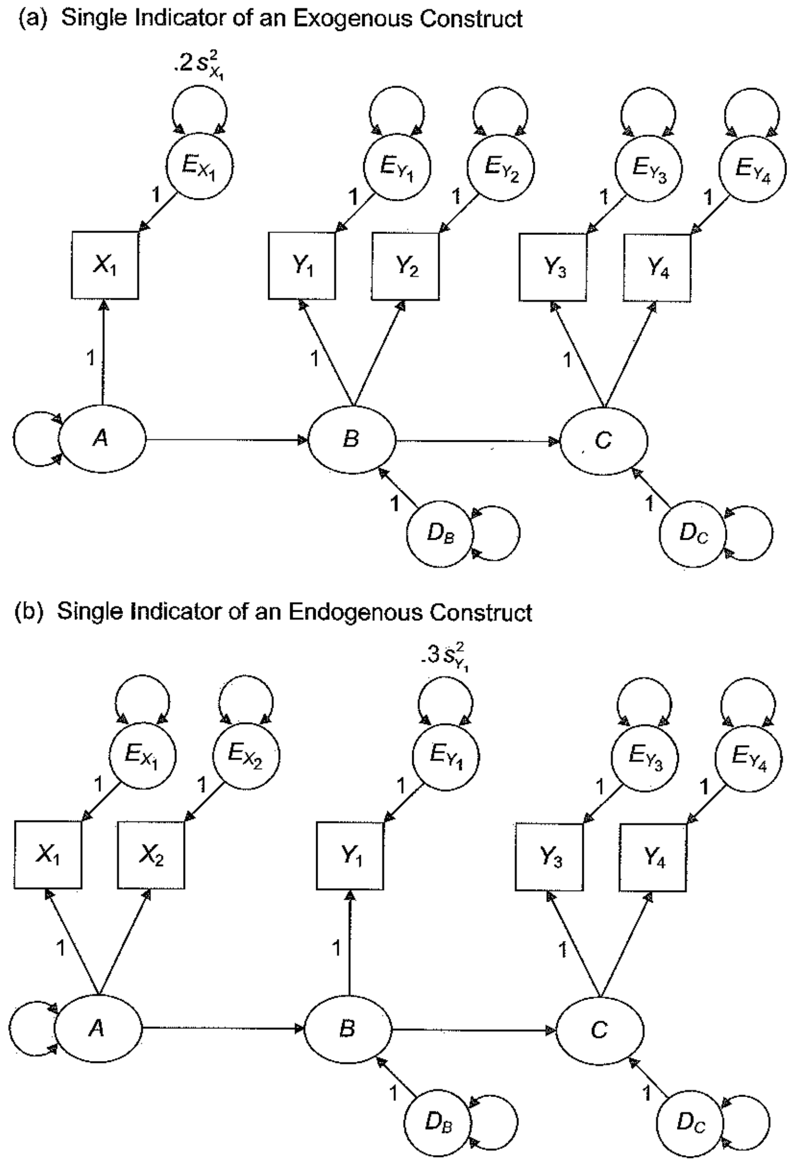

FIGURE 10.4. Two structural regression models with single indicators that correct for measurement error. It is assumed that the proportion of error variance for X_1 is .20 and for Y_1 it is .30.

for X_1, the direct effect of factor A is estimated controlling for measurement error in its single indicator.

Now look at the SR model of Figure 10.4(b), in which Y_1 is specified as the sole indicator of the endogenous factor B. The proportion of measurement error in Y_1 is estimated to be .30. Given this estimate, the variance of the error term for Y_1 is fixed to equal .30 times the observed variance of Y_1. Because Y_1 has an error term, both the disturbance variance for B and the direct effect of this factor will be estimated while controlling for measurement error in its single indicator. Three points should be noted about this method for single indicators:

1. A common question is, why not just specify the error variance for a single indicator as a free parameter and let the computer estimate it? Such a specification may result in an identification problem (see Bollen, 1989, pp. 172–175). A safer tactic with a single indicator is to fix the value of its measurement error variance based on a prior estimate.
2. A related question is, what if the researcher is uncertain about his or her estimate of the error variance for a single indicator? The model can be analyzed with a range of estimates, which allows the researcher to evaluate the impact of different assumptions about measurement error on the solution (i.e., conduct a sensitivity analysis).
3. It is theoretically possible to specify a path model where every observed variable is represented as the single indicator of an underlying factor and every indicator has a measurement error term. This tactic would be akin to fitting a path model to a covariance matrix based on correlations disattenuated for unreliability (Equation 3.7). See Bedeian, Day, and Kelloway (1997) for more information.

The models in Figure 10.4 illustrate that SR models with single indicators that are identified may nevertheless fail the two-step rule for identification (Rule 6.9): when either model in the figure is respecified as a CFA measurement model, one factor (A or B) will have only one indicator, which is one less than the minimum for a standard multifactor model (Rule 6.5). Fixing the error variance of X_1 in the model of Figure 10.4(a) or Y_1 in the model of Figure 10.4(b) to a constant, however, identifies the model.

Shen and Takeuchi (2001) administered within a stratified random sample of 983 native-born Chinese Americans and immigrants of Chinese descent measures of the degree of acculturation, socioeconomic status (SES), stress, and depression. Descriptive statistics for these variables are summarized in Table 10.5. Note in the table that there is just a single indicator of depression. This data matrix is ill scaled because the ratio of the largest variance (11.834) over the smallest variance (.058) exceeds 200. Therefore, I multiplied the original variables by the constants listed in Table 10.5 in order to make their variances more homogeneous.

Presented in Figure 10.5 is the SR model analyzed by Shen and Takeuchi (2001). This model reflects the hypothesis that stress is directly affected by the degree of acculturation and that depression is directly affected by both SES and stress. Values of selected

TABLE 10.5. Input Data (Correlations and Standard Deviations) for Analysis of a Structural Regression Model of Acculturation and Mental Health Status with a Single Indicator

Variable	1	2	3	4	5	6	7	8
Acculturation								
1. Acculturation Scale	1.00							
2. Generation Status	.44	1.00						
3. Percent Life in U.S.	.69	.54	1.00					
Socioeconomic status								
4. Education	.37	.08	.24	1.00				
5. Income	.23	.05	.26	.29	1.00			
Stress								
6. Interpersonal	.12	.08	.08	.08	−.03	1.00		
7. Job	.09	.06	.04	.01	−.02	.38	1.00	
Single indicator								
8. Depression	.03	.02	−.02	−.07	−.11	.37	.46	1.00
Original s^2	.608	.168	.058	10.693	11.834	.137	.203	.102
Constant	4.00	8.00	10.00	1.00	1.00	8.00	8.00	10.00
Rescaled s^2	9.728	10.752	5.800	10.693	11.834	8.768	12.992	10.200
Rescaled *SD*	3.119	3.279	2.408	3.270	3.440	2.961	3.604	3.194

Note: These data are from Shen and Takeuchi (2001); $N = 983$.

fit statistics calculated by LISREL 8.8 with ML estimation for the model in Figure 10.5 are as follows:

$$\chi^2_M(16) = 59.715, \quad p < .001$$
$$\text{RMSEA} = .053\ (0.39\text{–}.068); \quad p_{\text{close-fit } H_0} = .343$$
$$\text{GFI} = .985; \quad \text{CFI} = .977; \quad \text{SRMR} = .032$$

The exact-fit hypothesis is rejected, so there is a need to understand why this test was failed. I inspected the correlation residuals (derived in EQS), and none of their absolute values are > .10. Also, the parameter estimates for the model in Figure 10.5 seemed reasonable in the converged and admissible solution. In this case, the chi-square test may be failed due more to the relatively large sample size ($N = 983$) than to appreciable discrepancies between observed and predicted correlations or covariances. This outcome indicates the need to routinely examine the residuals in every analysis.

The disturbance for the single indicator of depression in Figure 10.5 reflects both measurement error and omitted causes, which is not ideal. Assuming a score reliability of $r_{XX} = .70$, Exercise 2 will ask you to respecify the model in Figure 10.5 such that measurement error in the depression scale is estimated separately from the effects of omitted causes. Next, use an SEM computer tool to fit this respecified model to the data

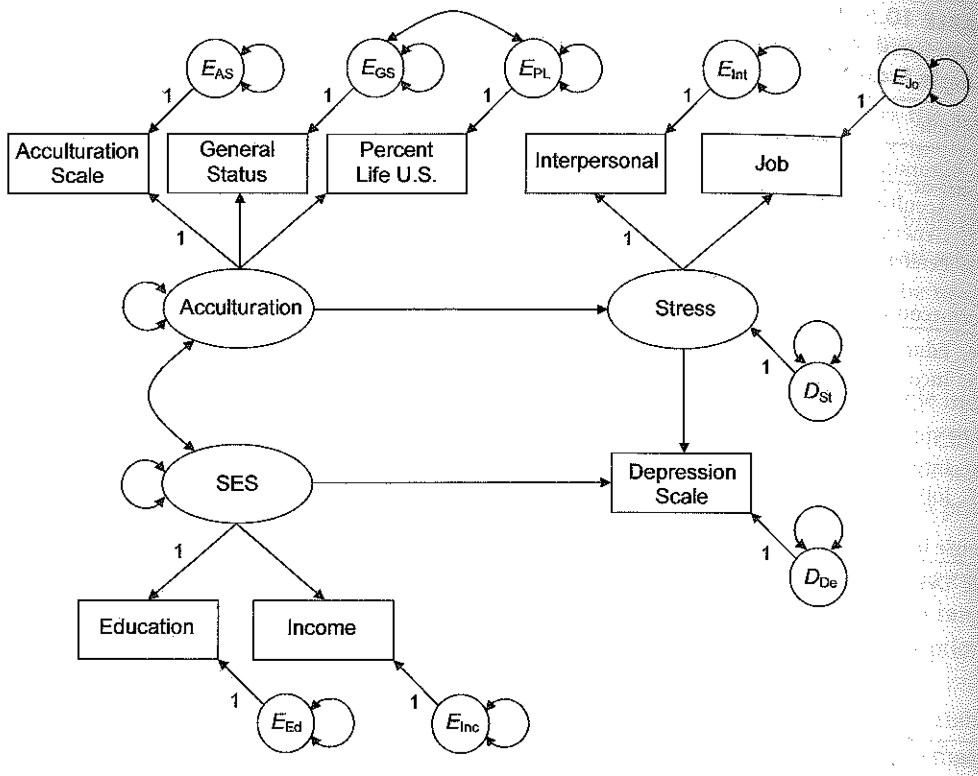

FIGURE 10.5. A structural regression model of acculturation and mental health status with a single indicator.

in Table 10.5. Look for a "surprise" among the parameter estimates. You can download the EQS and LISREL syntax and output files for the analysis just described from this book's website (p. 3).

CAUSE INDICATORS AND FORMATIVE MEASUREMENT

Observed variables in standard measurement models are represented as effect (reflective) indicators that are presumed to be caused by the underlying factors and their measurement errors. This directionality specification describes reflective measurement. This approach assumes (1) that equally reliable indicators are interchangeable, which implies that they can be substituted for one another without affecting construct definition. It also requires (2) positive intercorrelations among the indicators of the same factor. Finally, (3) factors are conceptualized in reflective measurement as unidimensional latent variables (Chapter 5).

The assumptions just listed are not suitable for some research problems, especially in areas where composites, or **index variables**, are analyzed. Recall the example from Chapter 5 of SES as a composite that is determined by measured variables such as income, education, and occupation, not the other way around. This view is consistent with a formative measurement model wherein manifest variables are specified as cause (formative) indicators (Chapter 5). The origins of formative measurement lie in the operational definition model (Diamantopoulos & Winklhofer, 2001). An older, strict form of operationalism views constructs as synonymous with the single indicator that corresponds to its definition. More contemporary forms of operationalism allow for both multiple indicators and disturbance terms for composites. The latter permits the representation of **latent composites** that are determined in part, but not entirely, by their cause indicators. Cause indicators are *not* generally interchangeable. This is because removal of a cause indicator is akin to removing a part of the underlying construct (Bollen & Lennox, 1991). Cause indicators may have *any* pattern of intercorrelations, including ones that are basically zero. This is because composites reflect the contribution of multiple dimensions, albeit with a single score for each case (i.e., composites are not unidimensional). There are many examples of the analysis of composites in economics and business research (Diamantopoulos, Riefler, & Roth, 2005).

Presented in Figure 10.6 are three "mini" measurement models that illustrate differences between reflective measurement and formative measurement. The model of Figure 10.6(a) depicts standard reflective measurement. Grace and Bollen (2008) use the term **L → M block** (latent to manifest) to describe the association between factors and their effect indicators in reflective measurement models. Measurement error in such models is represented at the indicator level by the error terms E_1–E_3 in Figure 10.6(a).[2]

A formative measurement model is represented in Figure 10.6(b). It depicts an **M → L block** (manifest to latent) because the latent composite in this model is presumed to be caused in part by its formative indicators, X_1–X_3. In Figure 10.6(b) I used a circle to represent the latent composite because, like error terms but unlike factors, a latent composite is not unidimensional. With no disturbance, the composite in Figure 10.6(b) would be just a linear combination across its cause indicators. To scale the latent composite, the unstandardized direct effect of one of its cause indicators, X_1, is fixed to 1.0. Cause indicators in formative measurement models are exogenous variables and have no error terms. This means that (1) cause indicators are free to vary and covary, which explains the presence of the symbols in Figure 10.6(b) that represent their variances and covariances (respectively, and). Also, (2) measurement error in a formative measurement model like the one in Figure 10.6(b) is manifested in the disturbance term, D_{LC}. That is, measurement error is represented at the construct level, not at the indicator level as in reflective measurement (e.g., Figure 10.6(a)). Note that the model in Figure 10.6(b) is not identified. In order to estimate its parameters, it would be necessary to embed it in a larger model. Identification requirements of formative measurement models are considered momentarily.

[2]A factor can also be endogenous in a reflective measurement model, but the term L → M block still applies.