Chapter 1
Introduction to Meta-Analysis and Structural Equation Modeling

Abstract Meta-analysis is a prominent statistical tool in many research disciplines. It is a statistical method to combine the effect sizes of separate independent studies, in order to draw overall conclusions based on the pooled results. Structural equation modeling is a multivariate technique to fit path models, factor models, and combinations of these to data. By combining meta-analysis and structural equation modeling, information from multiple studies can be used to test a single model that explains the relationships between a set of variables or to compare several models that are supported by different studies or theories. This chapter provides a short introduction to meta-analysis and structural equation modeling.

Keywords Meta-analysis · Introduction · Structural equation modeling · Path model · Factor model · Model fit

1.1 What Is Meta-Analysis?

The term "meta-analysis" was introduced by Glass (1976), who differentiated between primary analysis, secondary analysis, and meta-analysis. However, the techniques on which meta-analysis is based were developed much earlier (see Chalmers et al. 2002; O'Rourke 2007). In the terminology of Glass, primary analysis involves analyzing the data of a study for the first time. Secondary analysis involves the analysis of data that have been analyzed before, for example to check the results of previous analyses or to test new hypotheses. Meta-analysis then involves integration of the findings from several independent studies, by statistically combining the results of the separate studies. One of the first meta-analyses in the social sciences was performed by Smith and Glass (1977), who integrated the findings of 375 studies that investigated whether psychotherapy was beneficial for patients, a topic that was much debated at the time. By using a quantitative approach to standardizing and averaging treatment/control differences across all

S. Jak, *Meta-Analytic Structural Equation Modelling*,
SpringerBriefs in Research Synthesis and Meta-Analysis,
DOI 10.1007/978-3-319-27174-3_1

the studies, it appeared that overall, psychotherapy was effective, and that there is little difference in effectiveness across the different types of therapy. Around the same time as Smith and Glass performed this meta-analysis, other researchers developed similar techniques to synthesize research findings (Rosenthal and Rubin 1978, 1982; Schmidt and Hunter 1977), which are now all referred to as meta-analysis techniques. Meta-analysis is used to integrate findings in many fields, such as psychology, economy, education, medicine, and criminology.

1.1.1 Issues in Meta-Analysis

Compared with primary analysis, meta-analysis has important advantages. Because more data is used in a meta-analysis, the precision and accuracy of estimates can be improved. Increased precision and accuracy also leads to greater statistical power to detect effects.

Despite the obvious positive contributions of meta-analysis, the technique is also criticized. Sharpe (1997) identified the three main validity threats to meta-analysis: mixing of dissimilar studies, publication bias, and inclusion of poor quality studies. The mixing of dissimilar studies, also referred to as "mixing apples and oranges" problem, entails the issue that average effect sizes are not meaningful if they are aggregated over a very diverse range of studies. Card (2012) counters this critique by stating that it depends on the inference goal whether it is appropriate to include a broad range of studies in the meta-analysis (e.g. if one is interested in fruit, it is appropriate to include studies about apples, oranges, strawberries, banana's etc.). Moreover, a meta-analysis does not only entail aggregation across the total pool of studies, but can also be used to compare different subsets of studies using moderator analysis. The second threat, publication bias, is also referred to as the "file drawer" problem, and points to the problem that some studies that have been conducted may not be published, and are therefore not included in the meta-analysis. Publication bias is a real source of bias, because the non-published studies are probably those that found non-significant or unexpected results. Several methods exist that aim at avoiding, detecting and/or correcting for publication bias (see Rothstein et al. 2005; van Assen et al. 2014) but there is no consensus on the best ways to deal with the problem. The third issue, the inclusion of poor quality studies in the meta-analysis is also denoted as the "garbage in, garbage out" problem. Although it may seem logical to leave studies of poor quality out of the meta-analysis a priori, it is recommended to code the relevant features of the included primary studies that are required for high quality (e.g. randomization in an experiment), so that later on one can investigate whether these quality-conditions are related to the relevant effect sizes (Valentine 2009).

Cooper and Hedges (2009) distinguish six phases of research synthesis: Problem formulation, literature search, data evaluation, data analysis, interpretation of the results and presentation of the results. In this book we focus on the data analysis phase, referred to as meta-analysis. The other parts of research synthesis are discussed in for example Borenstein et al. (2009), Card (2012), Cooper et al. (2009), and Lipsey and Wilson (2001).

1.1.2 Statistical Analysis

Usually, the units of analysis in a meta-analysis are not the raw data, but summary statistics (effect size statistics) that are reported in the individual studies. The type of effect size statistic that is investigated depends on the nature of the variables involved. For example, if the interest is in differences between a treatment and control group on some continuous outcome variable, the meta-analysis may focus on the standardized mean difference (like Cohen's d or Hedges' g). If the hypothesis is about the association between two continuous variables, the (z-transformed) product moment correlation coefficient may be the focus of the analysis. If the interest is in association between two dichotomous variables, the (logged) odds ratio is often an appropriate effect size statistic. Once the effect size statistics of interest are gathered or reconstructed from the included studies, the statistical analysis can start, using fixed effects or random effects analysis.

The fixed effects approach is useful for conditional inference, which means that the conclusions cannot be generalized beyond the studies included in the analysis (Hedges and Vevea 1998). In the most common fixed effects model, it is assumed that the effect size statistics gathered from the studies are estimates of one population effect size, and differences between studies are solely the result of sampling error. The analysis focuses on obtaining a weighted mean effect size across studies. The weights are based on the sampling variance in the studies, so that studies with larger sampling variance (and smaller sample size) contribute less to the weighted mean effect size (which is the estimate of the population effect size).

The random effects approach facilitates inferences to studies beyond the ones included in the particular meta-analysis (unconditional inference). The random effects approach assumes that the population effect sizes vary from study to study, and that the studies in the meta-analysis are a random sample of studies that could have been included in the analysis. Differences in effect sizes between studies are hypothesized to be due to sampling error and other causes, such as differences in characteristics of the respondents or operationalization of the variables in the different studies. The random effects analysis leads to an estimate of the mean and variance of the distribution of effect sizes in the population.

Apart from the average effect size, it is often also of interest if and why studies differ systematically in their effect size statistics. Therefore, researchers often code study characteristics (e.g. average age of respondents, measurement instruments used, country in which the study was conducted), and investigate whether the effect sizes are associated with these study-level variables. This is called moderator analysis, and is used to investigate whether the association between the variables of interest is moderated by study characteristics. These moderator variables may explain variability in the effect sizes. If all variability is explained, a fixed effects model may hold, implying that conditional on the moderator variables, all remaining variability is sampling variability. If effect sizes are regressed on study level variables in a random effects approach, reflecting that the moderator variables do not explain all variability across the studies, this is called mixed effects meta-analysis.

To be consistent with recent terminology, I use the term “fixed effects model” for all models that do not estimate between-studies variance. This terminology is common in meta-analysis, but not in line with the statistical literature, where the fixed effects model denotes the model in which heterogeneity is explained by study-level variables. The model that assumes homogeneity of effect sizes, without study-level variables, is also called the “equal effects model” (Laird and Mosteller 1990). I use the term “fixed effects model” for both these models, and will explicitly state when study-level variables are included in the model.

1.2 What Is SEM?

Structural equation modeling (SEM) has roots in two very different techniques developed in two very different fields. Path analysis with its graphical representations of effects and effect decomposition comes from genetics research, where Wright (1920) proposed a method to predict heritability of the piebald pattern of guinea-pigs. Factor analysis is even older, with an early paper by Spearman (1904), and was developed in research on intelligence, to explain correlations between various ability tests (Spearman 1928). Jöreskog (1973) coined the name LISREL (LInear Structural RELations) for the framework that integrates the techniques of path analysis and factor analysis, as well as for the computer program that made the technique available to researchers.

1.2.1 Path Analysis

SEM is a confirmatory technique, which means that a model is formulated based on theory, and it is judged whether this model should be rejected by fitting the model to data. If multivariate normality of the data holds, the variance covariance matrix of the variables of interest and the sample size are sufficient to fit models to the data. This is a very convenient aspect of SEM, because it means that as long as authors report correlations and standard deviations of their research variables in their articles, other researchers are able to replicate the analyses, and to test different hypotheses on these data. In order to test hypotheses, these hypotheses have to be translated in a statistical model. The statistical model can be formulated in different ways, for example using a graphical display. The graphical displays that are used for structural equation models use squares to represent observed variables, ellipses to represent latent variables, one-headed arrows to represent regression

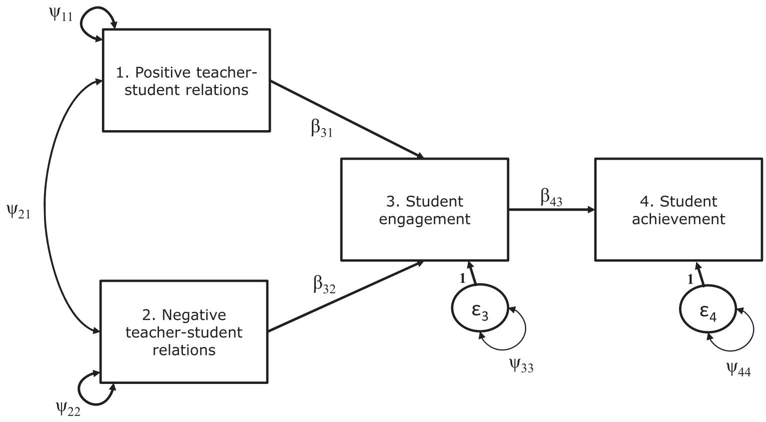

Fig. 1.1 Hypothesized path model in which the effects of Positive and Negative relations on achievement is fully mediated by engagement

coefficients, and two-headed arrows to represent covariances. Consider the path model in Fig. 1.1, in which the effect of negative and positive relations with teachers is hypothesized to affect student achievement through student engagement.

The four observed variables are depicted in squares. Student engagement is regressed on Positive and Negative relations, and Student Achievement is regressed on Student Engagement. There are no direct effects of Positive and Negative relations on Student Achievement, reflecting the hypothesis that these effects are fully mediated by Student Engagement. In this model, Engagement and Achievement are called endogenous variables, reflecting that other variables are hypothesized to have an effect on them. Variables that are not regressed on other variables are called exogenous variables. Positive and Negative relations are exogenous variables in this model. The two exogenous variables are assumed to covary, indicated by the two-headed arrow between them. There are also two-headed arrows pointing from the variable to itself, reflecting the variance of the variable (a covariance with itself is equal to a variance). The endogenous variables have a latent variable with variance pointing to it. This latent variable is called a residual factor, and could be viewed as a container variable representing all other variables that also explain variance in the endogenous variable, but that are not included in the model. The regression coefficient of the variable on the residual factor is not estimated but fixed at 1 for identification of the model. The variance of the residual factor represents the unexplained variance of the endogenous variable. So, part of the variance in Student Engagement is explained by Positive and Negative relations, and the remaining variance is residual variance (or, unexplained variance). Similarly, part of the variance in Student Achievement is explained by Student Engagement, and the remaining variance is residual variance. For the exogenous variables, actually, all variance is unexplained. So it seems logical to depict two more residual factors with variance pointing to Negative and Positive relations, instead of the double headed arrow pointing to the variables themselves. Indeed, this would be correct, but to keep the graphs simple they are often not depicted. Actually, the residual factor pointing to an endogenous variables is also often not fully depicted, but represented by a small one-sided arrow.

Attached to the arrows in the graphical display, the Greek symbols represent the model parameters. In a path model, the direct effects are often denoted by β and variances and covariances by ψ. For example, β_{43} represents the regression coefficient of Variable 4 on Variable 3, ψ_{44} represents the residual variance of Variable 4, and ψ_{21} represents the covariance between Variable 1 and Variable 2. The model parameters are collected in matrices. A path model on observed variables can be formulated using two matrices with parameters, matrix **B** and matrix **Ψ**, and an identity matrix, **I**. For the example, these matrices look as follows, with rows 1–4 and columns 1–4 corresponding to the variables Positive relations, Negative relations, Student Achievement, and Student Engagement, respectively:

$$\mathbf{B} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \beta_{31} & \beta_{32} & 0 & 0 \\ 0 & 0 & \beta_{43} & 0 \end{bmatrix}, \quad \boldsymbol{\Psi} = \begin{bmatrix} \psi_{11} & & & \\ \psi_{21} & \psi_{22} & & \\ 0 & 0 & \psi_{33} & \\ 0 & 0 & 0 & \psi_{44} \end{bmatrix} \text{ and } \quad \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Matrix **Ψ** is a symmetrical matrix, so the covariance between Variables 1 and 2 is equal to the covariance between Variable 2 and 1. Using these parameters, a model implied covariance matrix ($\boldsymbol{\Sigma}_{\text{model}}$) can be formulated. The model implied covariance matrix is a function of the matrices with parameters:

$$\boldsymbol{\Sigma}_{\text{model}} = (\mathbf{I} - \mathbf{B})^{-1} \boldsymbol{\Psi} (\mathbf{I} - \mathbf{B})^{-1\mathrm{T}}. \tag{1.1}$$

The resulting model implied covariance matrix ($\boldsymbol{\Sigma}_{\text{model}}$) for the current example can be found in Appendix A. The basic hypothesis that is tested by fitting a structural equation model to data is:

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\text{model}}. \tag{1.2}$$

Note however, that the population covariance matrix, **Σ**, is generally unavailable to the researcher, who only observed a covariance matrix based on a sample, denoted **S**. Suppose that observed covariance matrix of the four variables based on 104 respondents is as given in Table 1.1.

Table 1.1 Variances (on diagonal) and covariances of four research variables, N = 104

Variable	1	2	3	4
1. Positive relations	0.81			
2. Negative relations	−0.36	1.21		
3. Engagement	0.63	−0.60	1.69	
4. Achievement	0.14	−0.33	0.50	1.44

The model parameters that make up $\boldsymbol{\Sigma}_{\text{model}}$ can be estimated by minimizing a discrepancy function. This means that parameters are estimated in order to minimize the difference between the model implied covariance matrix ($\boldsymbol{\Sigma}_{\text{model}}$), and the observed covariance matrix (**S**). The more parameters a model has, the easier it is to make the $\boldsymbol{\Sigma}_{\text{model}}$ close to **S**. The maximum number of parameters that a model can have in order to be identified is equal to the number of observed variances and covariances in **S**. In our example with four variables, the number of variances and covariances is ten. The number of parameters in the $\boldsymbol{\Sigma}_{\text{model}}$ equals eight (three regression coefficients, one covariance, four variances). The degrees of freedom (df) of a model are equal to the difference between these two. This model has 2 degrees of freedom. The larger the degrees of freedom of a model is, the more the model is a simplification of reality. Simple models are generally preferred over complicated models. But, the larger the degrees of freedom, the larger the difference between $\boldsymbol{\Sigma}_{\text{model}}$ and **S** will be, meaning that the absolute fit of a model will be worse.

Having less parameters than observed variances and covariances is not the only requirement for identification of the model. For a model to be identified, all parameters in the model need to be identified. See Bollen (1989) for an overview of methods to assess the identification of model parameters. If a model is identified, the parameters can be estimated. The most used estimation method is maximum likelihood (ML) estimation. The discrepancy function F_{ML} that is minimized with ML estimation is:

$$\text{F}_{\text{ML}} = \log|\boldsymbol{\Sigma}_{\text{model}}| - \log|\mathbf{S}| + \text{trace}\left(\mathbf{S}\boldsymbol{\Sigma}_{\text{model}}{}^{-1}\right) - \text{p}, \qquad (1.3)$$

where p is the number of variables in the model. If the model fits the data perfectly, the model implied covariance matrix will be equal to **S**, and F_{ML} will be zero. If the model does not fit perfectly, F_{ML} will be larger than zero. See Bollen (1989) for a description of ML and other estimation methods and their assumptions.

1.2.2 Model Fit

An important property of the ML estimator is that it provides a test of overall model fit for models with positive degrees of freedom. Under the null hypothesis ($\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\text{model}}$), the minimum F_{ML} multiplied by the sample size minus one ($n - 1$) asymptotically follows a chi-square distribution, with degrees of freedom equal to the number of non-redundant elements in **S** minus the number of model parameters. If the chi-square value of a model is considered significant, the null hypothesis is rejected. The chi-square of a model may become significant because the discrepancy between **S** and the estimated $\boldsymbol{\Sigma}_{\text{model}}$ is large, or because the sample is large. With a very large sample, small differences between **S** and the estimated $\boldsymbol{\Sigma}_{\text{model}}$ may lead to a significant chi-square, and thus rejection of the model. Other fit measures are available in SEM, which do not test exact fit of the model, but are based on the idea that models are simplifications of reality and will never exactly hold in the population. The Root Mean Squared Error of Approximation (RMSEA, Steiger and Lind 1980) is the most prominent fit measure next to the chi-square. The RMSEA is interpreted using suggested cut-off values that should be regarded as rules of thumb. RMSEA values smaller than 0.05 are considered to indicate close fit, values smaller than 0.08 are considered satisfactory and values over 0.10 are considered indicative of bad fit (Browne and Cudeck 1992). Another prominent fit measure is the Comparative Fit Index (CFI, Bentler 1990) that is based on a comparison of the hypothesized model with the "independence model", which is a model in which all variables are unrelated. CFI values over 0.95 indicate reasonably good fit. For an overview of these and other fit indices see Schermelleh-Engel et al. (2003).

Fitting the model from Fig. 1.1 to the observed covariance matrix in Table 1.1 gives the following fit indices: $\chi^2 = 2.54$, df $= 2$, $p = 0.28$, RMSEA $= 0.05$ and CFI $= 0.99$. So, exact fit of the model is not rejected, and the model also fitted the data according to the rules of thumb for the RMSEA and CFI. If the model fits the data, the parameter estimates can be interpreted. If a model does not fit the data, the parameter estimates should not be interpreted because they will be wrong. Table 1.2 gives an overview of the unstandardized parameter estimates, the 95 % confidence intervals and the standardized parameter estimates of the model. See Appendix B for an example of an OpenMx-script to fit the current model.

All parameters in this model differ significantly from zero, as judged by the 95 % confidence intervals. For interpretation, it is useful to look at the standardized parameter estimates. For example, the standardized β_{31}, means that 1 standard deviation increase in Positive relationships is associated with 0.45 standard deviations increase in Engagement, controlled for the effect of Negative relationships. The standardized residual variance is interpreted as the proportion of residual

Table 1.2 Unstandardized parameter estimates, 95 % confidence intervals and standardized parameter estimates of the path model from Fig. 1.1

Parameter	Unstandardized estimate	95 % confidence interval		Standardized estimate
		Lower bound	Upper bound	
β_{31}	0.64	0.40	0.89	0.45
β_{32}	−0.30	−0.50	−0.10	−0.26
β_{43}	0.30	0.13	0.47	0.32
ψ_{21}	−0.36	−0.60	−0.36	−0.36
ψ_{11}	0.81	0.62	1.08	1.00
ψ_{22}	1.21	0.93	1.61	1.00
ψ_{33}	1.10	0.85	1.47	0.65
ψ_{44}	1.29	0.99	1.72	0.90
$\beta_{31} \times \beta_{43}$	0.19	0.08	0.34	0.14
$\beta_{32} \times \beta_{43}$	−0.09	−0.19	−0.03	−0.08

variance. This means that in the standardized solution, the proportion of explained variance in Student achievement is calculated as $1 - \psi_{44}, = 0.10$. The proportion of explained variance in Engagement is 0.35. Indirect effects are calculated as the product of the two direct effects that constitute the indirect effect. With OpenMx, one can estimate confidence intervals for indirect effects as well. The indirect effects of Positive and Negative relationships on Student Achievement are both small but significant (see the last two rows in Table 1.2). This shows that as expected, there is significant mediation. Whether there is full or partial mediation can be investigated by testing the significance of the direct effects of Positive and Negative relationships on Student Achievement. This is shown in Chap. 5.

1.2.3 Factor Analysis

Factor analysis can also be seen as a special case of structural equation modeling. Factor models involve latent variables that explain the covariances between the observed variables. Consider the two-factor model on five scales measuring children's problem behavior depicted in Fig. 1.2.

In a factor model, each indicator is affected by a common factor that explains the covariances between the indicators. The regression coefficients linking the factor to an indicator are called factor loadings. The larger a factor loading is, the more variance the factor explains in the indicator. Not all indicator variance may be common variance, which is reflected by the residual factors that affect each indicator. The variance of these residual factors is called residual variance (denoted by θ) and is assumed to consist of random error variance and structural

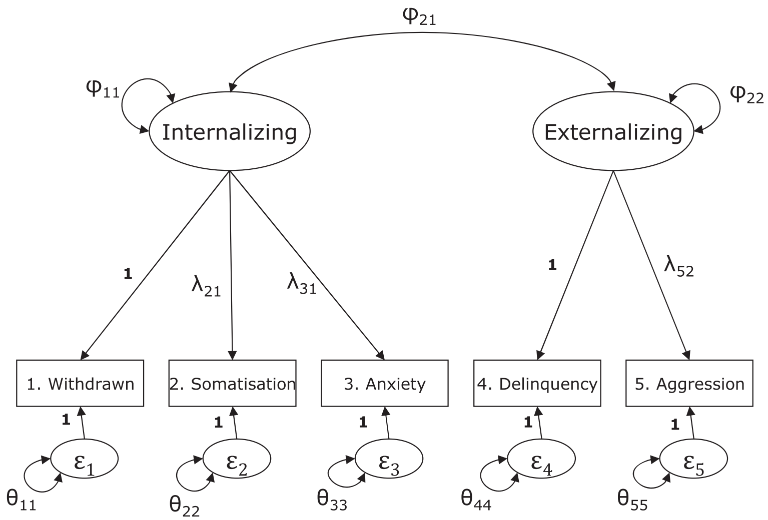

Fig. 1.2 A two-factor model on the five problem behavior variables

variance. For example, there may be a structural component in Somatisation that is not correlated with Anxiety or Withdrawn behavior.

With factor analysis, Σ_{model} is a function of factor loadings, depicted by λ's, factor variances and covariances, depicted by φ's, and residual variances, depicted by θ's. Note that one factor loading for each factor is fixed at 1. This is needed to identify the model. As factors are unobserved variables, the scale of the variables is not known, and a metric has to be given to the factors by fixing one factor loading per factor. Alternatively, one can fix the factor variances φ_{11} and φ_{22} at some value (e.g. 1) and estimate all factor loadings. In advanced models (e.g. multigroup and longitudinal models) one method of scaling may be preferred over the other, but in this example it is arbitrary how the factors are given a metric. The unstandardized parameters will differ based on the scaling method, but the model fit and the standardized parameter estimates will not. The factor model can be represented by three matrices with parameters, a full matrix $\mathbf{\Lambda}$ with factor loadings, a symmetrical matrix $\mathbf{\Phi}$ with factor variances and covariances, and a symmetrical matrix $\mathbf{\Theta}$ with residual variances and covariances. For the current model, the three matrices look as follows.

$$\mathbf{\Lambda} = \begin{bmatrix} 1 & 0 \\ \lambda_{21} & 0 \\ \lambda_{31} & 0 \\ 0 & 1 \\ 0 & \lambda_{52} \end{bmatrix}, \quad \mathbf{\Phi} = \begin{bmatrix} \varphi_{11} & \\ \varphi_{21} & \varphi_{22} \end{bmatrix} \text{and} \quad \mathbf{\Theta} = \begin{vmatrix} \theta_{11} & & & & \\ 0 & \theta_{22} & & & \\ 0 & 0 & \theta_{33} & & \\ 0 & 0 & 0 & \theta_{44} & \\ 0 & 0 & 0 & 0 & \theta_{55} \end{vmatrix}.$$

The rows of $\mathbf{\Lambda}$ are associated with variables 1 through 5 from Fig. 1.2, as well as the rows and columns of $\mathbf{\Theta}$. The columns of $\mathbf{\Lambda}$ and the rows and columns of $\mathbf{\Phi}$ are associated with the Internalizing and Externalizing factors respectively.

The factor model is specified using these matrices as:

$$\mathbf{\Sigma}_{\text{model}} = \mathbf{\Lambda\Phi\Lambda}^{\mathrm{T}} + \mathbf{\Theta}, \tag{1.4}$$

leading to the model implied covariance matrix given in Appendix C.

Suppose that we observed the covariance matrix of the five variables from a sample of 155 parents with children suffering from epilepsy that is given in Table 1.3.

Fitting the model from Fig. 1.2 to these data leads to good fit with the following fit measures: $\chi^2 = 4.08$, df $= 4$, $p = 0.40$, RMSEA $= 0.01$ and CFI $= 1.00$. The unstandardized parameter estimates, 95 % confidence intervals and standardized

Table 1.3 Variances (on diagonal) and covariances of five research variables

Variable	1	2	3	4	5
1. Withdrawn	12.55				
2. Somatization	6.31	10.06			
3. Anxiety	11.15	9.64	26.02		
4. Delinquency	2.85	2.09	4.84	3.72	
5. Aggression	12.44	9.68	22.20	9.96	51.02

Table 1.4 Unstandardized parameter estimates, 95 % confidence intervals and standardized parameter estimates of the factor model from Fig. 1.2

Parameter	Unstandardized estimate	95 % confidence interval		Standardized estimate
		Lower bound	Upper bound	
λ_{11}	1	–	–	0.74
λ_{21}	0.85	0.11	8.03	0.70
λ_{31}	1.78	0.18	9.25	0.86
λ_{42}	1	–	–	0.77
λ_{52}	4.54	0.50	9.04	0.94
φ_{11}	6.78	1.37	4.96	1
φ_{22}	2.18	0.43	5.11	1
φ_{21}	2.78	0.54	5.16	0.72
θ_{11}	5.69	0.84	6.81	0.46
θ_{22}	5.12	0.707	7.252	0.51
θ_{33}	6.78	1.571	4.314	0.26
θ_{44}	1.52	0.255	5.936	0.41
θ_{55}	5.72	3.945	1.451	0.11

parameter estimates are given in Table 1.4. All standardized factor loadings are larger than 0.70, meaning that they are substantially indicative of the common factor on which they load. The correlation between the common factors internalizing and externalizing is significant and quite large, 0.72. The proportion of explained variance is largest in indicator 5 ($1 - 0.11 = 0.89$) and smallest in indicator 2 ($1 - 0.51 = 0.49$). See Appendix D for an annotated OpenMx-script from this example.

In the two examples given in this chapter the input matrix was a covariance matrix. Maximum likelihood estimation assumes analysis of the covariance matrix, and not of the correlation matrix. However, sometimes only the correlation matrix is available. Treating the correlation matrix as a covariance matrix leads to incorrect results when estimating confidence intervals or when testing specific hypotheses (Cudeck 1989). To obtain correct results, a so-called estimation constraint can be added. This constraint enforces the diagonal of the model implied correlation matrix to always consist of 1's during the estimation.

The factor model and path model are the two basic models within the structural equation modeling framework. Once a factor model has been established, the analysis often goes some steps further, for example by including predictor variables like age to investigate age differences in the latent variables Internalizing and Externalizing problems. Another extension is multigroup modeling, in which a model is fitted to covariance matrices from different groups of respondents simultaneously, giving the opportunity to test the equality of parameters across groups. For example, in the path model from Fig. 1.1, it may be hypothesized that the effect of Positive and Negative relations on Engagement may be stronger for children in elementary school than for children in secondary school.

Some cautions about SEM have to be considered. If a model fits the data well, and is accepted by the researcher as the final model, it does not mean that the model is the correct model in the population. If the model is not rejected, this could be due to lack of statistical power to reject the model. Moreover, there may be other models that fit the data just as well as the hypothesized model. Therefore, it is important to consider equivalent models (MacCallum et al. 1993). If a model is rejected however, the conclusion is that the model does not hold in the population. This chapter is far too short to discuss all relevant issues in SEM. Several books have been written that can be used to learn about SEM, see for example Bollen (1989), Byrne (e.g. 1998), Geiser (2012), Loehlin (1998), and Kline (2011).

1.3 Why Should You Combine SEM and MA?

Most research questions are about relations (or differences) between a set of variables. The hypothetical model in Fig. 1.1 for example, states that positive and negative relations lead to achievement through engagement. Current practice in meta-analysis is to meta-analyze each effect in this model separately. The questions these analyses answer are: What is the pooled effect of positive relations on engagement? And: What is the pooled effect of engagement on achievement? However, what the researcher also may want to know is: Is this model a good representation of the data? Are the effects of positive and negative relations on achievement fully mediated by engagement? Which effects are lacking in this model?

Using MASEM, information from multiple studies is used to test a single model that explains the relationships between a set of variables or to compare several models that are supported by different studies or theories (Becker 1992; Viswesvaran and Ones 1995). MASEM provides the researcher measures of overall fit of a model, as well as parameter estimates with confidence intervals and standard errors. By combining meta-analysis and SEM, some of the difficulties in the separate fields may be overcome.

Structural equation modelling requires large sample sizes. Small samples lead to low statistical power, and non-rejection of models. If several (small) studies investigate the same phenomenon, they may end up with very different final models, leading to a wide array of models describing the same phenomena. By combining the information from several (possibly underpowered) primary studies, general conclusions can be reached about which model is most appropriate. Norton et al. (2013) for example, used MASEM to investigate the factor structure of an anxiety and depression scale, by comparing ten different models that were proposed based on different primary studies. Furthermore, MASEM can be used to answer research questions that are not addressed in any of the primary studies. Even about models that include a set of variables that none of the primary studies included all in their study. For example, Study 1 may report correlations between variable A and variable B. Study 2 may report correlations between variables B and C, and Study 3 between variable A and C. Although none of the studies included all variables, one model can be fit on these three variables using MASEM (Viswesvaran and Ones 1995).

I use the term MASEM for the process of fitting a structural equation model on the combined data from several studies. SEM can also be used to perform ordinary meta-analysis (SEM-based meta-analysis). SEM-based meta-analysis is outside the scope of this book, but see Cheung (2008, 2015) for an explanation.

MASEM is a fairly young field of research, and it seems to be growing in popularity, both in substantive and methodological research. At this moment, a special issue about MASEM is being edited for the journal *Synthesis Research Methods*.

References

Becker, B. J. (1992). Using results from replicated studies to estimate linear models. *Journal of Educational Statistics, 17*, 341–362.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238–246.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: Wiley.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. (2009). *Introduction to meta-analysis*. Hoboken, NJ: Wiley.

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research, 21*, 230–258.

Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Card, N. A. (2012). *Applied meta-analysis for social science research*. New York, NY: Guilford Press.

Chalmers, I., Hedges, L. V., & Cooper, H. (2002). A brief history of research synthesis. *Evaluation and the Health Professions, 25*(1), 12–37.

Cheung, M. W.-L. (2008). A model for integrating fixed-, random-, and mixed-effects meta-analyses into structural equation modeling. *Psychological Methods, 13*(3), 182–202.

Cheung, M. W.-L. (2015). *Meta-analysis: A structural equation modeling approach*. Chichester, UK: Wiley.

Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). (2009). *The handbook of research synthesis and meta-analysis*. New York, NY: Russell Sage.

Cooper, H. M., & Hedges, L. V. (2009). Research synthesis as a scientific enterprise. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 3–16). New York, NY: Russell Sage.

Cudeck, R. (1989). Analysis of correlation matrices using covariance structure models. *Psychological Bulletin, 105*(2), 317–327.

Geiser, C. (2012). *Data analysis with Mplus*. New York, NY: Guilford.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *The Educational Researcher, 10*, 3–8.

Hedges, L. V., & Vevea, J. L. (1998). Fixed-and random-effects models in meta-analysis. *Psychological Methods, 3*, 486.

Jöreskog, K. G. (1973). Analysis of covariance structures. *Multivariate analysis, 3*, 263–285.

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York, NY: Guilford.

Laird, N. M., & Mosteller, F. (1990). Some statistical methods for combining experimental results. *International Journal of Technology Assessment in Health Care, 6*, 5–30.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications.

Loehlin, J. C. (1998). *Latent variable models: An introduction to factor, path, and structural analysis*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin, 114*(1), 185–199.

Norton, S., Cosco, T., Doyle, F., Done, J., & Sacker, A. (2013). The Hospital Anxiety and Depression Scale: A meta confirmatory factor analysis. *Journal of Psychosomatic Research, 74*, 74–81.

O'Rourke, K. (2007). An historical perspective on meta-analysis: Dealing quantitatively with varying study results. *Journal of the Royal Society of Medicine, 100*(12), 579–582.

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (Eds.). (2005). *Publication bias in meta-analysis: Prevention, assessment and adjustments*. Hoboken, NJ: Wiley.

Rosenthal, R., & Rubin, D. B. (1978). Interpersonal expectancy effects: The first 345 studies. *Behavioral and Brain Sciences, 1*(03), 377–386.

Rosenthal, R., & Rubin, D. B. (1982). Comparing effect sizes of independent studies. *Psychological Bulletin, 92*(2), 500–504.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of psychological research online, 8*(2), 23–74.

Schmidt, F. L., & Hunter, J. E. (1977). Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology, 62*(5), 529–540.

Sharpe, D. (1997). Of apples and oranges, file drawers and garbage: Why validity issues in meta-analysis will not go away. *Clinical Psychology Review, 17*(8), 881–901.

Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist, 32*, 752–760.

Spearman, C. (1904). "General Intelligence," Objectively Determined and Measured. *The American Journal of Psychology, 15*(2), 201–292.

Spearman, C. (1928). The sub-structure of the mind. *British Journal of Psychology. General Section, 18*(3), 249–261.

Steiger, J. H., & Lind, J. C. (1980). *Statistically based tests for the number of common factors*. Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA (Vol. 758).

Valentine, J. C. (2009). Judging the quality of primary research. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 129–146). New York, NY: Russell Sage.

van Assen, M. A. L. M., van Aert, R. C. M., & Wicherts, J. M. (2014). Meta-analysis using effect size distributions of only statistically significant studies. *Psychological methods.* Advance online publication. http://dx.doi.org/10.1037/met0000025

Viswesvaran, C., & Ones, D. (1995). Theory testing: Combining psychometric meta-analysis and structural equations modeling. *Personnel Psychology, 48*, 865–885.

Wright, S. (1920). The relative importance of heredity and environment in determining the piebald pattern of guinea–pigs. *Proceedings of the National Academy of Sciences, 6*, 320–332.