

Meta-Analysis Using HLM

by

Ralph B. Taylor

breck@rbtaylor.net

All materials copyright (c) 1998-2002 by Ralph B. Taylor

This outline reviews meta-analytic procedures that can be followed using HLM.

Introduction

What is meta-analysis? Meta-analysis represents a family of techniques that can summarize results across studies that bear on a particular question.¹

What are the minimum requirements for a meta-analysis? You need a sizable number of studies. Although meta-analytic procedures are available for dealing with extremely small numbers of studies, such as 3-6, most meta-analytic studies base research on at least a dozen or more typically several dozen studies.

In addition, you need indicators for the same outcome concept across different studies. Although the actual variables used to indicate a concept need not be similar across studies, the indicators used should focus on the same underlying construct.

Finally, you need some quantitative reporting of differences on the outcome between different groups of participants. Different types of quantitative measures can be used in different studies. What is required, however, is that the different measures used be placed on a common metric.

What is the most typically used measure of an independent variable's effect on an outcome variable? The most common measure used is one of effect size (ES).

For example, in one study on the effects of therapeutic treatment for juveniles, used the following measure of effect size:

$$ES = \frac{X_E - X_C}{SD_C}$$

Where:

X_E = mean score of the experimental group on the dependent variable of interest.

X_C = mean score of the control group on the dependent variable of interest.

SD_C = standard deviation of the control group on the dependent variable of interest.²

¹ For an introduction to these techniques see R. B. Taylor (1993) Research methods in criminal justice New York: McGraw Hill; ch.16.

To obtain an unbiased estimate of the population effect size you will want to use the denominator that relies on the pooled variance of both the experimental and control groups. This pooling is used in the HLM transformations discussed below.

Example calculation with hypothetical data. The following hypothetical data illustrate the construction of a measure of effect size for one study, using the formula used by Garrett.

Outcome = scores on an index measure of psychological adjustment. Scores can range from -3 (extremely poorly adjusted) to + 3 (extremely well adjusted)

Treatment group (n = 50) average score after treatment = + 0.50.

Randomly assigned control group (n = 50) average score at end of observation = - 0.50.

Standard deviation of control group scores on outcome scale = 1.2

$$ES = \frac{+0.50 - (-0.50)}{1.2} = +.83$$

Note that had the experimental group and the control group had the same average score on the outcome:

$$ES = 0$$

Had the experimental group scored worse than the control group on the outcome:

$$ES < 0.$$

Interpreting ES. ES, a measure of treatment impact, can be interpreted in percentiles because it is in effect a standardized score or a z score. It is a measure of how many standard deviation units difference exists between the average person in the experimental group and the average person in the control group, on the outcome variable. Because of the properties of the normal curve, z scores show the proportions of a population falling above or below a particular score.

Therefore, if you observe an effect size of +.83 in a study, this can be translated into areas under the normal curve, and can be interpreted as follows:

The 'average' person in the experimental group, with a 50th percentile score in that group on the outcome, had a score that was equivalent to a person scoring in the 79th percentile in the control group.

How does HLM Conceptualize a Meta-Analysis?³

In a way we already have become acquainted with meta-analysis. When we summarize results looking at mean scores and slope differences across a range of neighborhoods, that constitutes a meta-analysis. Each neighborhood, and its results, can be viewed as a single study.

³ Much of the material following here merely restates material you can find in Chapter 7 of B&R.

The questions HLM allows us to address in meta-analysis are fivefold (B&R p. 156).

First, what is the average measure of 'effect' across studies? We want to obtain a measure of the average of the parameter, not the average that has been observed. Why? Because the average observed effects are contaminated or influenced by differences in studies, and, of course, sampling errors.

You realize, of course, that we have already been doing this, in all of our analyses estimating B00s. So what we are doing here is no different than what we have been doing before, except that now our B00s mean something different. Whereas in prior HLMs B00s reflected estimated true mean scores on an outcome variable, here they represent estimated true scores of effects of an independent variable on a dependent variable.

Second, you want to measure the variance associated with this effect size parameter, and test whether the variance is significantly greater than zero. Note this is not the variance of the estimates, but of the parameter or estimated true scores. The idea here: does the impact of an independent variable on a dependent variable vary significantly across a number of studies, or is that impact relatively consistent across different studies?

Again, we already have been doing this with our estimate of T00, reflecting the variance of B00, and our chi square test to see if $T00 > 0$.

Third, if we establish that effect size parameters do differ significantly across studies, we will try and predict why the effects are larger in some studies than in others.

Again, we already have been doing this in our HLMs using L-2 variables to understand why B00 is higher in some L-1 units and lower in others. Here we use characteristics of the study, or of its population, to predict the estimated true score of the effect size.

Fourth, once we have predicted the effect size parameter, we can test to learn if the remaining residual variance in the effect size parameters (i.e., T00 residualised) is significant, or if the L-2 (study-level) features we have used to predict have explained most of T00.

Fifth, we can recover from each study estimates of its effect size.

How is Meta-Analysis with HLM Different From Other HLMs?

Meta-analysis via HLM is different from what we already have been doing with HLM in the following two respects.

...In meta-analysis the raw data from each study are rarely available. Instead, only summary statistics published in research reports are accessible to the meta-analyst. Second, different studies typically use different outcome measures, even though these are viewed as measures of the same construct ... meta-analysts have employed a variety of standardized measures of effect ... Using standardized measures of effect translates each study's results to a common scale so that they may be compared. (B&R p. 157).

In short: we only have summary data with which to work, and must put results from different studies onto a common ruler.

There is also a third way meta-analysis is different, stemming from sampling theory. (p. 157) If we use a standardized measure of effect size, and if the studies in question include more than about 30 participants in each study, the statistic itself -- the measure of effect size -- will be normally distributed "with known variance at Level 1."

So meta-analysis represents a special case of V-known applications.

Steps in Meta-analysis

Getting your standardized measure

B&R recommend computing a standardized measure. The measures you use probably will need to be transformed to measures that better approximate a normal distribution using a monotonic transformation. (See Table 7.4, p. 169.)

If you start with:	You transform into:	Using the ___ transform:
Mean differences	effect size	
Correlations	z scores	Fisher r to z
Proportions	logits	logit
Standard Deviations	logs	log (natural)

Note: since you can work with standard deviations as your starting point, you can focus on degrees of variation across studies, as well as mean differences (p. 169).

What you start with is

d_j : a sample estimate of the difference between two groups of respondents in a study, for example: mean fear difference between elderly and non-elderly; mean outcome score difference on cognitive adjustment scale for control and experimental juveniles

Make your transformation of difference, estimate variance

You transform this into two different elements:

δ_j [delta-j]: the parameter estimate (i.e., estimated true score) of the effect or the difference

V_j : "the sampling variance of D_j as an estimate of δ_j " (p. 159). In other words: how much does D_j vary around δ_j ? If most of the variance in D_j is due to true variation in delta-j, i.e., T_{00} , then V_j will be small.

$$\text{Var}(d_j) = T + V_j \quad [\text{Eq. 7.11}]$$

You accomplish this using the transformations listed in Table 7.4, p. 169.

Level-1 Model

Your unit of analysis is individual level studies.

Your L-1 model decomposes the variation in your observed effect measures into true scores and error (B&R p. 159, Eq. 7.5):

$$d_j = \delta_j + e_j$$

where

d_j = measure observed effect size

δ_j = estimated true score (parameter) for effect size

e_j = "sampling error associated with d_j as an estimate of δ_j " with a mean of 0 and a variance of V_j (p. 159).

Level-2 Model

Now you will go on to try and predict, using study features as L-2 predictors (Ws), what study features account for the mean difference or variations measured in the studies in L-1.

ANOVA. At the simplest level:

$$\delta_j = G_0 + U_{0j}$$

where

G_0 = your precision-weighted, grand-mean estimated true score of effect

U_{0j} = measure of cross-study variation in estimated true scores of effect, with a mean of 0 and a variance of T_0

Conditional model. And, of course, you can add in predictors so that:

$$\delta_j = G_0 + G_1W_{1j} + G_2W_{2j} \dots + U_{0j}$$

where W = study-level features, and U_{0j} = residualized cross-study error in measures of effect size.

Setting Up the Data

This is explained in the manual, pp. 66-68

Col. 1 (Character) ID variable

Col 2 ... m Q statistics representing observed effect size measures.

IF you have more than one Q these measures must be followed by:

Col 4 ... n Variance-covariance matrix for the Q statistics, entered in lower triangle matrix order, starting with the variance of Q1.

IF you have only one Q statistic it is followed by:

Col. 3 Your variance measure for Q. See Table 7.4 in B&R.

Col. o ... p Scores on potential L-2 predictors.