

3 Setting Up a Meta-Analysis

Traditional, narrative reviews of research literature are selective. Critics of such studies rightly ask, "Why were some studies selected but not others? What criteria were used? What kind of bias did the selection process introduce into the review?"

A meta-analysis should be comprehensive and replicable. It should not only examine as much of the research as possible, it should also describe how you found the research so that other researchers can evaluate *your* work.

This chapter describes the tasks required to organize and conduct a meta-analysis.

Focusing Your Aim	35
Conducting the Literature Search	36
Handling Questions of Study Quality	37
Dealing with Publication Bias	38
Documenting Your Search	38
Creating the Meta-Analysis in <i>Meta-Stat</i>	38
Defining Variables	39
Classes of Variables	39
Variables for a Meta-Analysis of the Effects of Psychotherapy on Asthma	39
Variables for a Meta-Analysis of the Effects of Coaching on SAT Scores	40
Deciding Which Variables to Use for Your Meta-Analysis	41
Setting Up Variables in <i>Meta-Stat</i>	42

continued

Coding the Results of Each Study	43
Dealing with Data Errors	43
Entering Study Data into <i>Meta-Stat</i>	44
Analyzing and Displaying Results	45
First steps	45
A comment on Schmidt and Hunter meta-analysis	47
Summary of Steps for a Meta-Analysis	49

Focusing Your Aim

Before setting out to collect studies, you must first decide the range of topics your review will encompass. Meta-analyses typically cover broad topics, sometimes loosely defined topics. They examine questions such as, Does psychotherapy work? Does computer-assisted instruction lead to more learning than traditional instruction? Is mastery learning better than traditional learning?

These questions, by themselves, are too broad to be meaningful. Does "psychotherapy," for example, mean reading a self-help book, or spending a few sessions with a college counselor, or completing a multi-year psychoanalysis?

To avoid vague generalities, you must make the focus of your meta-analysis much more explicit by establishing criteria for including or excluding studies. However, the criteria cannot be too restricting, or you might not find a sufficient number of studies; and even if you do find enough studies, you might not find anything interesting or illuminating.

To focus your aim, read a good sample of the literature and develop a thorough understanding of the concepts and methods that you want to analyze. Determine the effect, or outcome, you want to study, as well as the predictors, or independent variables, that you want to measure.

Conducting the Literature Search

A thorough literature search is critical to the validity of your meta-analysis:

How one searches determines what one finds; and what one finds is the basis of the conclusions of one's integration of studies. (Glass, 1976)

Finding research studies is difficult and time-consuming. The studies can be located in a variety of places, and often you must look beyond the titles. For example, if you are conducting a meta-analysis of sex differences, you will find that some studies show such differences even when that is not the principal variable of interest in the study and even when the study title makes no mention of sex differences.

To find studies, check the following general categories described by Rosenthal (1984):

! Books

- Authored books
- Edited books
- Chapters in edited books

! Journals

- Professional journals
- Published newsletters
- Magazines
- Newspapers

! Theses

- Doctoral theses
- Master's theses
- Bachelor's theses

! Unpublished work

- Technical reports
- Grant reports
- Convention papers not published in proceedings
- ERIC reports
- Films and cassette recordings
- Other unpublished work

Also consider soon-to-be-published works. Given the time lags between data collection, write-up, journal acceptance, and publication, the latest word on a topic may well be in the pipeline to publication. Your literature search should produce a list of names you can contact with inquiries about recent research. These researchers will also know about others who are doing research on similar topics.

Last, and perhaps most important, consult secondary sources such as review periodicals and the myriad abstract archives that are available in most fields of study. These sources will help you identify the primary research that will be the subject of your meta-analysis.

Handling Questions of Study Quality

One of the most controversial questions related to meta-analysis is the question of whether to include studies that are of doubtful or poor quality. Some critics invoke the Garbage-In, Garbage-Out principle by arguing that any meta-analysis that summarizes studies of widely differing quality is likely to be uninformative or flawed. These critics argue that studies with methodological flaws should be eliminated from consideration in the meta-analysis.

Others counter by noting that it is often difficult to assess methodological quality and researchers often disagree on quality. Despite a researcher's best attempts to provide an objective measure of quality, decisions to include or exclude studies introduce bias into the meta-analysis.

Others note that the quality of a study may not have an effect on the study's outcome. When in doubt, include the study in the meta-analysis and use an independent variable to code the quality of a study. Then examine empirically whether the outcome does in fact vary with study quality. You can do this with *Meta-Stat* by examining the relationship between study quality and effect size.

Dealing with Publication Bias

Well-recognized publication biases can produce bias in your meta-analysis.

First, publication policy is biased toward statistically significant findings. In other words, it is easier to get published if you have something statistically significant to report. A number of meta-analyses have indeed found that effect sizes reported in journals differ widely from unpublished work. The effect size is about 33 percent larger in published research.

Second, publication biases can skew the direction of effect. For example, Smith (1980) reported that, in the published literature, counselors and therapists tend to view women more negatively than men. In the unpublished literature, the opposite is true. As Strube (1985) theorized:

Such biases probably reflect the fact that some findings better "fit" the prevailing scientific atmosphere (Zeitgeist) and are scrutinized less closely than are novel or counterintuitive results.

To overcome publication bias, gather everything you can find, both published and unpublished. You can then use empirical methods to examine the question of publication bias. For example, with **Meta-Stat** you can compare the effect sizes for published research and unpublished work.

Documenting Your Search

A traditional, narrative review of literature often devotes only one or two paragraphs to a discussion of the methods that were used to conduct the literature search. A meta-analysis requires a much more rigorous approach. Only through a comprehensive description can other researchers evaluate the validity of your work and, perhaps, gather clues about overlooked sources.

Creating the Meta-Analysis in *Meta-Stat*

Before you can define your variables in **Meta-Stat**, as described in the following sections, you must create the meta-analysis in **Meta-Stat**. Chapter 4 of this manual describes how to create and modify the set of data files that, together, make up a meta-analysis.

See Chapter 4 for complete information.

Defining Variables

After you collect the literature for your meta-analysis, you must determine which study features your meta-analysis will examine. These features become the variables in the meta-analysis.

Classes of Variables

There are three classes of variable:

! Variables that identify characteristics of the study

You will need some variables to identify when and where the research was published, whether a control group was used, the type of effect that was measured, and so forth.

These variables will help you show the relationships, if any, between study methodology and results.

! Variables that identify characteristics of the sample

Use these variables to identify the subject population—age, educational level, socioeconomic status, and so on.

! Variables that identify characteristics of the intervention

These variables could include the type of treatment or intervention, its duration, the type of effect that was measured, and so on.

Variables for a Meta-Analysis of the Effects of Psychotherapy on Asthma

Schlesinger et al. (1978) conducted a meta-analysis that evaluated the effects of psychotherapy on asthma. This meta-analysis found that psychotherapy does indeed significantly lessen both the medical and nonmedical effects of asthma.

In the meta-analysis, the following variables were used:

This variable ...	Was used to identify ...
Therapy type	Type of psychotherapy received by treatment subjects For example, some subjects received hypnotherapy while others received group therapy.
Age	Average age of subjects
Hours of Therapy	Number of hours that subjects received therapy
Control Group	Type of therapy received by the control group Some control subjects received no treatment, others received relaxation therapy, and others received medical treatment.
Follow-Up Time	Amount of time between the end of treatment and the measurement of the outcome variable
Dependent Variable	The type of outcome or effect that was measured For example, some studies measured the effect of psychotherapy on the use of drugs to treat asthma; others measured the effect on asthma attacks and hospitalization.
Effect size	The size of the effect of treatment

Variables for a Meta-Analysis of the Effects of Coaching on SAT Scores

Becker (1990) conducted a meta-analysis that tested the effects of coaching on Scholastic Aptitude Test scores. Among other results, this meta-analysis found that the effectiveness of coaching varied widely across the studies, with much of the variation resulting from studies without comparison groups. The magnitude of the coaching effect is related to study design and to the duration of coaching intervention.

The following table describes some of the variables that were used in the meta-analysis:

This variable ...	Was used to identify ...
Year of publication	Year the study was published

This variable ...	Was used to identify ...
Type of publication	Whether the article was published in an academic journal
ETS Authorship	Whether any author was affiliated with Educational Testing Service, which develops and administers the SAT
Control Group	Type of control group
Use of randomization	Whether subjects were randomly assigned to control and coaching groups
Selectivity	The type of students in the study (e.g. low achievers, public-school students, college-prep students)
Voluntariness	Whether participation in coaching was voluntary or compulsory
Duration	Length of coaching program
Presence of Test Practice	Whether students practiced taking complete sample tests
Homework	Whether students were given coaching-related exercises for completion at home
Use of Computerized Instruction	Whether students were given computerized practice

Deciding Which Variables to Use for Your Meta-Analysis

In some cases, simple analysis will help you decide which variables are required for a meta-analysis. But it's not always so simple. Definitions of study features may vary from study to study, and different researchers often use different terminology.

The safest procedure is to have two researchers independently code all study features, then negotiate an agreement in places where they disagree.

Setting Up Variables in *Meta-Stat*

Chapter 5 describes how to create and modify variables in *Meta-Stat*. The program comes with many variables already predefined for you. There are variables to measure:

- ! Study characteristics, including the year of publication and the type of publication in which the research appeared
- ! Sample characteristics, including the sizes of the treatment and control groups
- ! Intervention characteristics, including the effect size and the unbiased effect size

You can also create your own variables to handle other study features. You can create different types of variables to handle different types of data.

See Chapter 4 for complete information about *Meta-Stat* variables.

Coding the Results of Each Study

After defining the variables you will use to measure studies, you must obtain the studies and code data for each variable.

Unfortunately, this task is not as straightforward as it might seem. Meta-analysts sometimes find that the required data are missing and must be estimated. Sometimes, researchers find errors in the original research. Errors in the publication are relatively easy to correct. But that is not the case for errors that were made when the data were originally recorded. These errors will contaminate your meta-analysis, but it is no easy matter to determine whether that contamination is severe.

Dealing with Data Errors

In social science research, errors commonly occur in favor of the anticipated outcome. In one study, researchers gave genetically identical white rats to students. The researchers told the students that some of the rats were fast runners and some were slow. When the students recorded running times in a maze, the "fast" rats had faster times than the "slow" ones.

Another study collected the original data from 27 published studies. The median error rate was about 1 percent, but rates of 3 percent and 4 percent were not uncommon. Two-thirds of these errors supported the researcher's hypothesis. Of course, if the researchers were unbiased, half of the errors should have supported the hypothesis and half should have countered it.

Dealing with Data Reporting Deficiencies

The scientific method requires researchers to report their studies in enough detail that anyone who wants to replicate a study can do so. Unfortunately, published research seldom exhibits the required clarity and detail. Further, there is no standardization among different journals—or even within a journal, as editors change—about what all studies must contain. About 40 percent of all studies even lack the means and standard deviations that are important to most meta-analyses. Fortunately, this deficiency does not make a meta-analysis impossible.

Deficient reporting can affect your meta-analysis by forcing you to exclude variables that,

although potentially important, are not described in the study results. Deficient reporting can also force you to abandon precise variables in favor of rougher groupings of data. This problem causes less variance in your variables and attenuates important relations in your data.

In your meta-analysis, it may be useful to add a variable that you can use to rate data quality. If you rate the quality of each study's data on a three-point scale, you can then examine whether the quality of reporting systematically affects the outcomes of studies.

Entering Study Data into *Meta-Stat*

Chapter 6 describes how to enter each study into your meta-analysis. For each study, you must identify the author and title. You must then enter data from the study for each variable that you have set up.

See Chapter 6 for full information.

Analyzing and Displaying Results

Chapters 7 and 8 describe how to analyze and display the results of your meta-analysis.

Chapter 7 describes the statistical techniques you can use to explore your data. Chapter 8 describes the graphs you can use to analyze and display results.

First steps

The meta-analyst is well advised to perform a number of very basic analyses before launching into more complex analyses in which some of the more basic features of the data base are obscured. Specifically, after entering the data, it is wise to perform simple frequency distributions and scatter plots to see if data have been entered correctly or whether some studies contain data that are clearly aberrant in the context of all studies in the analysis. These aberrations can occur in a variety of ways. Most simply, perhaps, are clerical errors committed in the entry of the data into Meta-Stat. Bad data points can enter the process earlier: typos in printed reports of the studies; data analysis errors in the primary data analysis that yield an F-ratio in error by a factor of two or three, for example; or conditions invalidating the transformation of a reported statistic into an effect size or correlation coefficient. However they arise, these errors produce odd values of variables that can distort the analysis of the study outcomes and their relationships with study characteristics. The errors need to be detected and corrected or removed.

Two analyses are particularly useful to this end. Under CHARTS can be found the options of BAR GRAPH and EFFECT SIZE BY STUDY. An early analysis should be the graphing of EFFECT SIZE (or CORRELATION in the case of correlations as the study outcome) as a BAR CHART. The results are then inspected for a small number of entries that are clearly separated from the bulk of the distribution of effects. How far removed from the mean or the next closest entries is "clearly separated"? No definitive answer can be given. Values more than five standard deviations from the mean in samples of 100 or fewer cases would certainly raise suspicions about errors in reporting or calculation. But it is difficult to be more precise than this here. (The reader interested in pursuing this point to a more precise answer should consult Dixon & Massey, 1969, or Tukey, 1977.)

The graph of EFFECT SIZE BY STUDY can add information to the BAR GRAPH. All effect sizes are arrayed in either ascending or descending order and each is bracketed by a confidence interval that reflects on the size of the samples in the study yielding the effect size. If the largest or smallest of these effects are very distant from their neighbors and the confidence intervals are likewise quite distinct and not overlapping, then it might be well to

return to the studies that produced these effects and see if some error is apparent. Not every aberration will have a simple explanation; when they lack reasons, the analyst is faced with the difficult question of eliminating the odd data point without good reasons or leaving it in and having it exercise undue influence on the later statistical analyses.

The problem of detecting outliers can be more complex than indicated above. Consider the follow data, which for the sake of example can be thought of as effect size from nine experiments in which a drug is tested against a placebo: .40, .45, .50, .55, .60, .75, .80, .85, .90. There appear to be no outliers or aberrant data points here. But suppose that the first four and the ninth effect sizes are from studies that are double blind and the fifth through the eighth effects are from single blind experiments: when the effects are grouped and inspected, the value of .90 appears aberrant and sends the analyst looking for errors or explanations:

Single Blind Effects: .60 .75 .80 .85

Double Blind Effects: .40 .45 .50 .55 .90

Consequently, it is advisable not only to inspect the entire data set for outliers but to break down the whole data set into various groups to see if aberrations appear when effects are compared with effects observed under similar conditions. Perhaps the best way to accomplish this analysis is to select EFFECT SIZE BY STUDY under the CHARTS menu and then use SELECT IF to construct selection criteria to view the effects for various subgroups of discrete study characteristics.

For continuous study characteristics, the counterpart to the above analysis is a scatter plot relating the study characteristic to the study outcome (e.g., EFFECT SIZE). When SCATTER PLOT is selected from the CHARTS menu and EFFECT SIZE plotted against a study characteristic, a quick visual fitting of the relationship will reveal points that fall far away from the general relationship between the two variables. The studies contributing these outliers can be identified with the POINT SHOOT feature, and reread or checked for obvious errors.

An alternative means of detecting bad data points is to perform REGRESSION analyses under the ANALYSIS menu and inspect the residuals graph that the analysis produces. Very large and aberrant residuals identify data points that have peculiarities unrelated to the study characteristics used in the regression.

Although outliers can be critical when they occur in the study outcome variable, say EFFECT SIZE, they can occur in other places in a data set. Occasionally they cause problems in subsequent analyses. The methods suggested above for searching for outliers in study outcomes can also be applied to any other variable (with the exception of the

EFFECT SIZE BY STUDY graph that is only available for the EFFECT SIZE variable). In addition, the BREAKDOWN tables under ANALYSIS are useful in detecting bad data points in Grouping Variables that describe study characteristics.

A comment on Schmidt and Hunter meta-analysis

The effect-size meta-analysis approach in *Meta-Stat* tends to follow the concepts outlined by Glass and by Hedges. These techniques are very similar to those of Schmidt-Hunter, but they do differ in several critical ways. Schmidt and Hunter argue that

- 1) the effect size should be based on the pooled variance estimate rather than that of the control group,
- 2) the meta-analyst should correct for sampling error, and
- 3) the meta-analyst should correct for measurement error.

The argument for using the pooled variance estimate is that it is based on more observations and subject to less error than the control group variance estimate. The counter argument is that the control group variance is unaffected by the treatment. The effect-size calculator in *Meta-Stat* lets you choose your technique. In practice there will be very little difference.

Schmidt and Hunter recommend grouping, trimming, or selecting your data until the ratio of the error variance to the variance of the effect sizes is .75 or greater. They argue that if 75% of the variance is due to error, then the rest of the variance is probably also due to error. Therefore the population variance is zero and the model of a single effect size is consistent with the data. *Meta-Stat* provides you with their variance ratio under Analysis/Descriptive when the criterion variable is UNBIASED.

Their last key recommendation is to correct for study artifacts such as the lack of perfect reliability of the criterion measure and data dichotomization. For example, they recommend correcting for measurement error by dividing the effect-size estimate by the square root of the reliability of the criterion variable. They note that measurement error inflates the standard deviation and thus lowers the effect size. *Meta-Stat* can accommodate this correction. The program allows you to override values entered into an equation when one or more of the variables are missing. Thus, you can correct for measurement error and other artifacts by following these steps.

1. Use the effect size calculator to compute EFFECTSZ..
2. Manually divide the EFFECTSZ (or UNBIASED) value by the square root of the

dependent measure reliability (reliability for commercial tests can be found in the *Buros Mental Measurements Yearbook*).

3. Divide the above value by other correction formulas if desired and available.
4. Delete the value for EFFECTSZ.
5. Insert the corrected effect-size in UNBIASED.

Not all meta-analysis methodologists will agree with each of Hunter & Schmidt's recommendations. We diverge from their plan in at least three respects. It seems unwise to contaminate the variance estimate with the treatment effect. Furthermore the 75% rule-of-thumb test is quite arbitrary and not in the spirit of intelligent, contextualized data analysis. Third, corrections for measurement unreliability often make conceptual sense but are impossible to implement. Finding a test reliability (in a book or manual or from a computer printout) is one thing; but assessing a sensible and appropriate error variance is quite another. The error that enters the control group variance estimate and contributes to observed differences among persons is generally not the error assessed by ordinary convenient measures of test reliability like Cronbach's Alpha, Kuder-Richardson or various split-half coefficients. In addition, after reanalyzing several meta-analysis studies using the Glass and the Schmidt-Hunter techniques, Hough and Hall (1991) demonstrated that the corrections are trivial and that both approaches lead to the same conclusions.

Summary of Steps for a Meta-Analysis

To conduct a meta-analysis:

1. Focus your aim by deciding on an area of the literature that you want to explore and a specific topic that you want to analyze.
2. Conduct a thorough literature search.
3. Create the meta-analysis in *Meta-Stat*. See Chapter 4 of this manual.
4. Add variables that you will use to code study features. See Chapter 5.
5. Read the studies and code the variables for each study.
6. Enter the data for each study into *Meta-Stat*. See Chapter 6.
7. Explore and display the data with various statistical techniques. See Chapter 7 and 8.

