

Issues and Challenges in Coding Interventions for Meta-Analysis of Prevention Research

Elizabeth C. Devine

INTRODUCTION

Meta-analysis is the statistical analysis of a large collection of results from individual studies for the purpose of integrating findings (Glass 1976). In other words, it is a quantitative review of existing research in a substantive area, involving multiple tests of a common hypothesis. When applied to intervention research, meta-analysis can be helpful in determining whether multiple tests of an intervention yield effects on an outcome construct of interest that are similar in direction and magnitude. Although the concept and some of the statistics used in meta-analysis date from the 1930s (Hedges and Olkin 1985), tremendous strides in the acceptance and popularity of meta-analysis have occurred in the last 15 years (Chalmers 1991; Myers 1991).

Many of the challenges facing meta-analysts arise from the fact that they are restricted to investigating what has been studied previously in primary research. Unless meta-analysts obtain additional information from individual primary researchers, they are further limited by the information provided in the research reports. Meta-analysts lack the control that primary researchers have to specify the population to be studied, the interventions to be tested, and the outcome measures to be used. In addition, it is rare to find exact replications within a body of research. Even studies of the same basic hypothesis may have noteworthy differences in sampling and operationalization of the intervention and outcome constructs.

Faced with what can aptly be described as "lumpy data," the meta-analyst must make many decisions, and possibly revise those decisions, as the extent and limitations of the existing data are discovered. Like primary researchers, meta-analysts must make many judgments that help to determine the final product. These include what to study, the source of data to use, the final selection criteria for the review, what to measure and how to operationalize the constructs of interest, who should collect and code the data, what analyses to perform, and how to report results. In both meta-analysis and primary research, the research decisions to be made far outnumber the

calculations to be performed. The prevalence of choices, judgments, and compromises commonly made in meta-analysis (Nurius and Yeaton 1987), as well as their influence on the outcome of the meta-analysis (Wanous et al. 1989), have been discussed to a limited extent in the meta-analysis literature.

Purpose

The primary purpose of this chapter is to discuss the issues and challenges involved in one of the major judgment areas in meta-analysis, that of coding interventions. The focus is on interventions typically found in field research, such as the evaluation of drug abuse prevention programs and health care-related interventions.

BACKGROUND

Presby (1978), an early critic of meta-analysis, noted that combining overly broad categories of interventions can obscure important differences between treatments. Similar cautions can be raised for creating overly broad categories of subjects or outcomes. With this in mind, and without the control to insure that there are sufficient numbers of studies in all of the potential subcategories of interest, the meta-analyst must determine the selection criteria, the coding categories, and the grouping of studies for analysis. The decisions made should be based on the populations and constructs that are the target of intended generalization. There are no simple, "canned" programs for making these choices; many different decisions are possible. In fact, many different constellations of decisions may be justified and yield useful results, assuming that they are based on the current knowledge in the field and are consistent with the objective of the review. For example, reviews of the same general content area may look quite different depending on whether the primary purpose of the review is to inform professional practice, to test theory, or to influence policy.

In the early literature on meta-analysis, combining studies with multiple differences in a single analysis was often referred to as the "apples and oranges" problem (Glass et al. 1981). At one extreme, critics saw meta-analysis as hopeless mishmash. At the other extreme, proponents saw it as a way to learn about constructs that include both apples and oranges (e.g., fruit salad). In other words, meta-analysis may provide a way to identify whether certain phenomena are stubbornly replicable. That is, do they occur across many studies

despite minor differences in subjects, settings, measures, or interventions? However, as the meta-analyst strives for general conclusions about phenomena, there is a nagging question that must never be far from the meta-analyst's mind: How many differences (e.g., in subjects, outcomes, or interventions) can be tolerated before the analyses obscure meaning rather than inform?

Over the years there have been dramatic changes in the way differences in outcomes have been treated in meta-analyses. In the very early years of meta-analysis, it was not uncommon to see all the effect size values from all dependent variables combined in a single unweighted mean. The outcome would be termed something global like "well-being." Meta-analysis has come a long way from those beginnings. For example, in recent years there is general consensus that only one effect size value should represent a study in any analysis. This helps to ensure the statistical independence of the data. It is also much more widely accepted that only effect size values from measures of the same construct should be aggregated (Hedges and Olkin 1985).

There have not been such dramatic changes in the way meta-analysts treat differences among subject characteristics. However, for several reasons, aggregating across studies of subjects with different characteristics often presents somewhat fewer conceptual or practical problems. First, researchers are accustomed to subjects with different characteristics being included in a single study. Second, subject characteristics often are better reported and thus are easier to code than treatment characteristics. Third, if the studies in the meta-analysis include relevant information on the subgroups of interest to the meta-analyst, it is fairly easy to disaggregate studies according to subject characteristics and determine if the pattern of results is consistent across relevant subpopulations of interest.

There has been less discussion of, and there is less consensus about, coding and aggregating interventions. In any topical area there are many potential consumers of a meta-analysis and many different purposes for doing a review of existing research. Even among high-quality meta-analyses, it should not be surprising if meta-analyses of similar topics vary widely in the coding and aggregating of interventions. Some of the issues meta-analysts must deal with as they decide about coding interventions are discussed below.

CODING INTERVENTIONS

There are at least four major areas to be considered related to coding interventions. To facilitate discussion, these areas are presented in a linear fashion. However, in the meta-analyst's reality, they present themselves like a bowl of jelly that jiggles all over no matter where it's touched.

The first decisions relate to what should be coded about experimental interventions. These are followed closely by decisions about when and by whom experimental interventions should be coded. Third are the decisions about how to minimize bias in the coding of experimental interventions. And finally, since the essence of each experimental intervention is defined by the ways in which it is different from the control condition, decisions must be made about what to code about the control condition.

What To Code About the Experimental Intervention

With regard to the experimental intervention itself, the meta-analyst must decide what and how much information about the experimental intervention is useful to code, and how to categorize and aggregate experimental interventions.

There are no simple answers to these questions. In addition to considering the purpose of the meta-analysis, it is essential to consider the analyses that are planned, the size of the research base of studies meeting the selection criteria, and the variability among interventions that have been tested in the literature under review.

What To Code. When one is summarizing and analyzing the results obtained from a group of studies that are construct replications (Lykken 1968) rather than exact replications, the actual content of interventions will most likely have been operationalized in many different ways. This may be particularly true of interventions such as drug abuse prevention, counseling, or patient education that are proposed to ameliorate specific (and often complex) social, interpersonal, or health-related problems. The variability of content typically included in drug abuse prevention programs is illustrated by the classifications of curriculum content developed and used by Hansen (1992) in a review of 45 drug abuse prevention curriculums. Hansen identified 12 content areas called domains of content. These areas are: information about drugs, resistance skills training, decisionmaking skills, pledges or personal commitments not to use

drugs, values clarification, norm setting, creating alternatives, stress management, self-esteem building, life-skills training, goal setting, and peer problemsolving skills. Each of those 12 domains of content can be operationalized in many different ways. Such differences may be clinically or theoretically important.

In addition to coding the actual content of the intervention, it is often useful to code information about the manner of experimental treatment delivery and the context in which the intervention is tested. This includes information such as:

- Who delivered the content (e.g., a teacher, a counselor, a peer);
- What was the format of the program (e.g., lecture, discussion, role play);
- How long was the program;
- What substances were the focus of the program;
- What was the goal of the program (e.g., delay onset of drug use, abstinence, decreased high-risk use); and
- Who is the audience (all regular students, volunteers for a drug education program, residents in a juvenile detention center).

Depending on the purpose of the review, it may be desirable to code treatments so that various operationalizations of each domain of content can be examined for differences in treatment effectiveness. However, depending on the sample of studies included in the meta-analysis, analyzing individual treatment components may or may not be possible. Hansen (1992) noted two problems that may limit the ability of meta-analysts to determine the effects of some specific components of drug abuse prevention interventions. First, most of the drug abuse prevention interventions contained elements of content from multiple domains of content. This may be advantageous from a clinical perspective when the goal is to determine if certain programs are effective in field settings. However, from theory or policy perspectives it is often desirable to identify the maximally efficacious components of a prevention program to understand causal mechanisms or to refine and streamline an intervention. This will require the testing of specific components of the intervention, often

using factorial-type designs or including within-study comparisons of alternate treatments. There must be multiple tests of individual treatment components in the extant research before meta-analysis will be useful in summarizing the effects of those treatment components. Nonetheless, if only multidimensional intervention programs have been tested, it is important that their effects be summarized as well.

The second problem is related to reporting weaknesses that are often found in primary research (Orwin and Cordray 1985). Hansen's (1992) analyses were limited by serious deficiencies in the documentation of intervention characteristics. For example, time on task within a multidimensional treatment was not reported consistently. Tobler (this volume) noted that the descriptions of interventions were variable, with the content of drug abuse prevention programs being better reported than the manner of treatment delivery. Sometimes it is possible to obtain missing information from primary researchers. Tobler (1994) reported contacting primary authors when information about the intervention program was missing or ambiguous. Although contacting primary researchers is time consuming and not frequently done, it should be considered if vital information is missing from the research report.

In addition to coding the characteristics of the experimental intervention that are absolutely essential for the planned analyses, it can also be helpful to abstract or code detailed treatment characteristics. Having this data readily available allows the reviewer to provide a thick description of the interventions included in the review. It also enables the reviewer to determine whether interventions grouped together are, in fact, quite variable in substantial ways. Such differences could be used to explain results if outcomes are not homogeneous.

There is a downside, of course. Detailed coding of the experimental intervention requires additional time and resources that may be hard to justify if the resulting data play only a minor role in the review. It also may turn out to be an inefficient use of resources if weaknesses are so prevalent that specific characteristics (e.g., time on task within a multidimensional treatment) are reported in too few studies to allow them to be used as meaningful descriptors or potential moderator variables.

How To Categorize and Aggregate Interventions. In addition to coding specific characteristics of the experimental intervention, it is usually necessary to group experimental interventions into meaningful

categories to enhance interpretability and facilitate data analysis. A review of several meta-analyses on drug abuse prevention programs reveals different approaches to establishing categories, different reporting about the processes involved and the information used to develop intervention categories, and rather different numbers of categories of interventions developed.

Some approaches have been theory driven, like the four-factor classification of prevention program orientation used by Rundall and Bruvold (1988; Bruvold 1993). Various empirical approaches to developing categories of interventions have been used. Tobler (1986) reported analyzing major themes reported by researchers and proposed five functional-content categories. In later work, Tobler (this volume) grouped programs empirically based on both content matter and the manner of program delivery. Bangert-Drowns (1988) also categorized programs functionally. However, that categorization scheme included only three types of interventions: knowledge only, affective education only, and mixed. Hansen (1992), on the other hand, identified and coded each program according to "building block theoretical concepts." The pattern of occurrence of these concepts within programs was used to develop a provisional conceptual framework containing six distinct types of program content. In other substantive areas, such as psychology, education, and health care, theoretical (Shadish 1992), empirical (Devine and Cook 1983, 1986; Fernandez and Turk 1989), and multidimensional scaling (Smith and Glass 1977) approaches have been used to classify interventions. All of these approaches can be useful for addressing certain research questions. Just as one study does not answer all research questions on a topic, neither can a single meta-analysis. It should, however, be up to the experts in drug abuse prevention to judge the usefulness and relevance of the categories developed by the various meta-analysts of the drug abuse prevention research.

There are many sources of relevant information for the meta-analyst to use while developing categories of interventions. These include the research base itself; relevant theoretical, descriptive, and social policy literatures; and practice-derived knowledge. Among these sources, the studies under review provide the major limiting factor for the meta-analyst. The meta-analyst is working with an existing data set (the studies on the topic that have been completed and can be retrieved). If there is not a sufficient number of studies that included a specific version of the experimental intervention, then the effect of that specific type of treatment can not be examined. However, other approaches are possible; even if there is not a sufficient number of

studies in each of several types of norm-setting interventions, it may be appropriate and useful to group together somewhat different versions of the intervention into a more general category (e.g., of norm-setting interventions).

Two aspects of coding interventions are essential for the meta-analyst to keep in mind. First, whatever the process and the source of data used to develop categories of interventions, readers of the review should be well informed about how the categorization scheme was developed and the characteristics of intervention programs that were aggregated into specific categories. Without this type of information, the consumers cannot judge the validity of the conclusions drawn about the interventions. Second, it is essential to remember that between-study contrasts of the relative effectiveness of different types of experimental interventions (also called review-generated evidence by Cooper (1989)) have some inherent weaknesses. For example, caution must be used when one interprets differences in average effect size values from different subsets of studies. The subsets of studies may differ in many ways other than the fact that intervention "x" was the experimental treatment in one group of studies and intervention "y" was the experimental treatment in another group of studies. Differences in historic period, sample, setting, and/or operationalization of the outcome measure could be the reason for any observed differences in treatment effectiveness; therefore, between-study contrasts should be viewed as descriptive, and not as a basis for causal inference. However, between-study contrasts can provide a good basis for designing future research. Within-study contrasts of intervention "x" and intervention "y" from well-designed and executed studies can be used to examine causal relationships about the relative effectiveness of different types of experimental interventions. Within-study contrasts involve two or more experimental interventions being compared directly in the same study, or multiple experimental interventions being compared with a control group from the same study. By their very nature, within-study contrasts hold constant most, if not all, of the source of differences outlined above that plague between-study contrasts.

Timing and Personnel Involved in Coding Interventions

In addition to deciding what to code about interventions and how to group interventions for analysis, the meta-analyst must decide who will code this data and when it will be coded. It is fairly typical for meta-analyses to involve the use of an established coding form with specific directions. Nonetheless, the need for knowledgeable, well-

trained coders with adequate background in the substantive area is essential (Wachter 1988). In developing a coding form for studies, it is often possible to adapt certain sections of the coding form from other meta-analyses. For example, from a practical point of view, there are a fairly limited number of ways to record such things as publication source, publication date, or the number of subjects in the treatment group. In spite of the ambiguities often present in research reports, it is fairly easy to train individuals who have graduate research experience in the content area under review to code many characteristics of the study, the subjects, and the outcomes, with good interrater reliability.

Developing coding categories for interventions, on the other hand, requires much more specific knowledge about the substantive area under review. If coding categories for interventions are to be useful, they must be fine tuned to the purpose of the meta-analysis, the content area under review, and the typical reporting practices in the studies being reviewed.

Coding interventions requires extensive substantive knowledge about the phenomenon of interest. Given the typical length of most published research reports, it is probably not surprising that the details provided about interventions are often sketchy. When the description of the experimental intervention is very brief or vaguely worded, the coders are required to make many judgments about the appropriate categorization of various treatment characteristics. Coders who are well grounded in the substantive area will be better able to identify when coding categories need to be modified to capture the essence of the intervention, or to determine when reporting weaknesses or the use of outdated terms or operationalization of treatments is a factor in differences between interventions. Examples of the foregoing are provided by Hansen (1992) in a discussion of the substantial changes in how the constructs norm setting and alternatives training have been used in the drug abuse prevention intervention literature over the last two decades.

Minimizing Bias

Given the many judgments coders must make to categorize interventions, minimizing coder bias is a major concern (Cordray 1990a, 1990b). Experimenter expectancies (Rosenthal 1966), or in this case coder expectancies, are the main threat to the accuracy of coding interventions. In this context, coder expectancies refer to knowledge or beliefs on the part of the coder that adversely affect the

integrity of coding decisions. The issue here is not fraud or malevolent misrepresentation of treatment characteristics. Those can be problems, of course, but they require different remedies that are not addressed in this chapter. Inaccuracies arising from coder expectancies are much more subtle. If one wants to minimize this bias, its sources need to be recognized and steps taken to reduce its effect.

There are two main sources of experimenter expectancies. The first arises from close affiliation, on the part of the principal investigator or the coders, with certain studies or types of treatments under review. Although the author has stressed how important it is for the meta-analyst and the coders to be very familiar with relevant literatures, theories, and practice, such knowledge can be a source of bias. If the individuals making coding decisions are too closely affiliated with (and biased toward) particular studies or types of treatments, or if the data coders presume the principal investigator wants a certain outcome no matter what, their ability to code studies accurately may be restricted. Studies coinciding with a reviewer's beliefs may be evaluated much more favorably than those that do not, an effect that Mahoney (1977) called the "confirmatory bias."

The second potential source of bias comes from the research reports themselves (Cooper 1989). Among almost any large group of studies there will be variability in the prestige of the author(s), in the source of funding for the study, and in the prestige of source of the research report (e.g., a major journal in a field or an unpublished thesis). There will also be variability in the writing ability of the author(s), the direction and magnitude of treatment effects, and whether statistically significant results were obtained. Differences like these can create halo or shadow effects that may influence the decisions coders make about the experimental intervention (Peters and Ceci 1982).

There is no perfect solution to this problem. However, at a minimum the following checks and balances are recommended to improve the accuracy of coding studies.

- Prior to coding studies the research team should thoroughly review the coding categories and the coding directions for clarity, relevance, and comprehensiveness;
- Intercoder agreement should be examined, and the training of coders and refining of directions should continue until established;

- Coders should take frequent breaks to minimize errors arising from fatigue and boredom;
- Outside readers (or research team members), ideally with different theoretical or professional perspectives, can serve as supplemental coders to consider particularly ambiguous research reports or to help resolve conflicts when intercoder agreement is not achieved; and
- Coders should be made aware of the potential sources of bias so that they can be critical of their own decisions.

Another approach to minimizing coder bias that is gaining in popularity is to blind the coders to certain nonpertinent sections of the research report while coding other sections of the study (Chalmers et al. 1988; Devine 1992; Devine and Cook 1983; Sacks et al. 1987). For example, to minimize the effect of information such as the direction or magnitude of treatment effect or the title of the journal when coding interventions, steps could be taken to black out or cut out all irrelevant and potentially biasing information. Ideally, one would want the coders to have available only the sections of the research report related to the type of information being coded. This information can be photocopied, with irrelevant information blacked out, and then labeled with only an identification number. Depending on purpose of the review and the variables being examined, when coders are determining the content and nature of the experimental and control interventions, it might be helpful for them to have a photocopy of the introduction to the study, the review of literature, relevant parts of the methods sections, and any related papers by the same author for reference. These various sources of information might be critical for coding the experimental and control interventions. For example, coders may be better able to determine the implications of certain program descriptors if they are aware of the theoretical underpinnings upon which the study is based. It is also important for coders to review related papers by the same author, if they include greater detail about the intervention than the primary published report of the study.

There is another advantage in abstracting detailed information about the experimental intervention. One must be concerned with coder expectancies at the time of initial coding of studies and also during data analysis. If the effect size values obtained from a group of studies are heterogeneous (i.e., statistical testing suggests there is an

interaction between type of intervention and magnitude of treatment effect), then often studies are partitioned to test for homogeneity within more refined subsets of interventions (Hedges and Becker 1986). It is desirable to avoid knowledge about the direction and magnitude of effect size values when making decisions about how to partition studies. If detailed descriptions of the interventions are available in a form clearly separated from effect size information, then it is easier for the meta-analyst to make less biased decisions about how to reaggregate studies into subgroups for analysis.

While blinding coders to certain information can reduce bias, one must make sure that other problems are not created by artificially dividing studies into so many pieces for coding. Not all apparently credible studies provide credible results. For example, upon close inspection the reviewer may find that there was a poor fit between the program goals and the outcomes measured; a program could have been well designed but not faithfully implemented, or there could have been a fatal flaw arising from failed random assignment or considerable treatment diffusion across research design levels. Given the variability encountered in research reports, such information might appear in the methods section, in a footnote, or at the very end of the discussion section. Thus it is important for the research report to be read in its entirety by a research team member. This way, important information is less likely to be missed.

Coding Control Conditions

In order to fully interpret the content, manner, and context of the experimental intervention, it is essential to know something about the control condition. Control group data can be derived from many sources. Meta-analysis selection criteria usually specify that only studies with certain types of control conditions will be included in the review. In addition to the variability of control groups that arises from the manner of assignment to treatment condition, there is variability in control groups arising from the extent to which there is overlap between the experimental and control interventions. Control group interventions can include "no treatment," the "usual" treatment for someone in their situation, a placebo treatment, or an alternate treatment. There also can be noteworthy variation within each of these categories of control treatment. For example, among no-treatment control groups there can be varying degrees of treatment delivered by others in the environment.

In the drug abuse prevention area, while many forces in an adolescent's life foster drug abuse, many other forces are at work to prevent drug abuse. In addition to the efforts of parents and teachers, there may be relevant programming in the mass media (including video arcade games) and community-based or church-related efforts. If such efforts decrease drug use in the target population, they decrease the base rate of drug use and make it much more difficult for the effectiveness of the intervention to be demonstrated. Because these less formal interventions are rarely documented and vary within communities over time, it is difficult for the reviewer to account for their influence.

Placebo and usual care treatments also can vary in the degree of overlap with experimental interventions. In the health care literature and probably in other literatures as well, the actual content of usual care is rarely documented and so the degree of overlap is difficult to assess. The actual content of placebo interventions is usually better documented. Variability among placebo interventions in the degree of overlap with the experimental intervention has been shown to be related to the magnitude of treatment effect (Devine and Cook 1983). In social situations it is difficult to create placebo treatments that are both credible and as inert as the prototypical sugar pill. With very brief interventions, it often is possible to create an attention-type control treatment that provides equivalent time with an interested researcher or professional but contains irrelevant content. Creating credible placebo treatments for longer experimental interventions is much more difficult. Placebo treatments containing content (e.g., conflict resolution skills training) that is likely to affect an outcome of interest (e.g., drug use) are closer to alternate treatments, and it is better to treat them as such. Alternate interventions as control treatments provide special challenges as well as advantages to the meta-analyst. While they may provide excellent theory-relevant tests of causal mechanisms or the relative effectiveness of different treatments, they address very different hypotheses than contrasts with no-treatment control groups. Although it is problematic to combine tests of different hypotheses into a single analysis, contrasts between alternate treatments should not be disregarded. They are particularly valuable as a source of within-study contrasts from which one can get appropriate data for testing the relative effectiveness of different types of treatments.

Reporting weaknesses about control conditions is a major problem. While primary researchers usually report factors that affect the equivalence between experimental and control groups (e.g., the

manner of assignment to treatment condition), the actual experiences of the control group are often not as well reported. This makes it particularly challenging to code the content of control treatments and to conduct fine-grained analyses that account for differences in type of control group.

CONCLUSIONS

Coding experimental and control interventions is a critical step in the meta-analysis of intervention research. This information helps to define the specific constructs involved in the independent variable of the hypothesis tested. Special challenges exist when the effects of multidimensional treatments are the focus of interest or when important information is not detailed in the written report. There are no simple solutions. However, two guiding principles can help the meta-analyst and the consumers of a meta-analysis. First, as with all forms of research, the procedures and protocols guiding decisions should be explicit enough to allow critique and replication. And second, treatments that are aggregated should be similar enough to make combining their outcomes meaningful to likely consumers of the review.

Special actions may be needed to help the meta-analyst overcome the problems associated with working with an existing and limited set of data. In order to protect the integrity of coding, steps should be taken to reduce the likelihood of experimenter expectancies. It may be necessary for the meta-analyst to contact the primary researchers to obtain needed information. Special caution is always needed to appropriately interpret between-study contrasts.

Prospective meta-analysts and consumers of meta-analysis are cautioned to be modest in their expectations. Meta-analysis is limited by the extent, quality, reporting detail, and specific operationalizations tested in the existing research on the topic of interest. All forms of research review have these same limitations. The two major functions of most research reviews are to summarize what is known and to foster the further development of knowledge in an area through recommendations about future research topics and practices. To the extent that meta-analyses are more explicit, comprehensive, and critical than other forms of research reviews, they make unique contributions to the building of knowledge.

REFERENCES

- Bangert-Drowns, R.L. The effects of school-based substance abuse education—a meta-analysis. *J Drug Educ* 18(3):243-264, 1988.
- Bruvold, W.H. A meta-analysis of adolescent smoking prevention programs. *Am J Public Health* 83(6):872-880, 1993.
- Chalmers, T.C. Problems induced by meta-analyses. *Stat Med* 10:971-980, 1991.
- Chalmers, T.C.; Berrier, J.; Sacks, H.S.; Levin, H.; Reitman, D.; Nagalingam, R.; and Sacks, H.S. Meta-analysis of randomized control trials as a method of estimating rare complications of non-steroidal anti-inflammatory drug therapy. *Aliment Pharmacol Ther* 2(1):9-26, 1988.
- Cooper, H.M. *A Guide for Literature Reviews*. 2nd ed. Newbury Park, CA: Sage Publications, 1989.
- Cordray, D.S. Meta-analysis: An assessment from the policy perspective. In: Wachter, K., and Straf, M., eds. *The Future of Meta-Analysis*. New York: Russell Sage Foundation, 1990a. pp. 99-119.
- Cordray, D.S. Strengthening causal interpretations of nonexperimental data: The role of meta-analysis. In: Sechrest, L.; Perrin, E.; and Bunder, J., eds. *Research Methodology: Strengthening Causal Interpretations of Nonexperimental Data*. Washington, DC: U.S. Department of Health and Human Services Public Health Service, Agency for Health Care Policy and Research, 1990b. pp. 151-172.
- Devine, E.C. Effects of psychoeducational interventions: A meta-analytic analysis of studies with surgical patients. *Diss Abstr Int* 44:3356B, 1984.
- Devine, E.C. Effects of psychoeducational care with adult surgical patients: A theory-probing meta-analysis of intervention studies. In: Cook, T.D.; Cooper, H.; Cordray, D.S.; Hartmann, H.; Hedges, L.V.; Light, R.J.; Louis, T.A.; and Mosteller, F., eds. *Meta-Analysis for Explanation: A Casebook*. New York: Russell Sage Foundation, 1992. pp. 35-82.
- Devine, E.C., and Cook, T.D. A meta-analytic analysis of effects of psychoeducational interventions on length of post-surgical hospital stay. *Nurs Res* 32:267-274, 1983.
- Devine, E.C., and Cook, T.D. Clinical and cost-saving effects of psycho-educational interventions with surgical patients: A meta-analysis. *Res Nurs Health* 9:89-105, 1986.

- Fernandez, E., and Turk, D.C. The utility of cognitive coping strategies for altering pain perception: A meta-analysis. *Pain* 38:123-135, 1989.
- Glass, G.V. Primary, secondary, and meta-analysis of research. *Educ Res* 5:3-8, 1976.
- Glass, G.V.; McGaw, B.; and Smith, M.L. *Meta-analysis in Social Research*. Beverly Hills, CA: Sage Publications, 1981.
- Hansen, W.B. School-based substance abuse prevention: A review of the state of the art in curriculum, 1980-1990. *Health Educ Res* 7(3):403-430, 1992.
- Hedges, L.V., and Becker, B.J. Statistical methods in the meta-analysis of research on gender differences. In: Hyde, J.S., and Linn, M.C., eds. *The Psychology of Gender*. Baltimore: Johns Hopkins University Press, 1986. pp. 14-50.
- Hedges, L.V., and Olkin, I. *Statistical Methods for Meta-Analysis*. Orlando, FL: Academic Press, 1985.
- Lykken, D.T. Statistical significance in psychological research. *Psychol Bull* 70:151-159, 1968.
- Mahoney, M. Publication prejudices: An experimental study of confirmatory bias in the peer review system. *Cogn Ther Res* 1:161-175, 1977.
- Myers, D.G. Union is strength: A consumer's view of meta-analysis. *Pers Soc Psychol Bull* 17(3):265-266, 1991.
- Nurius, P.S., and Yeaton, W.H. Research synthesis reviews: An illustrated critique of "hidden" judgments, choices, and compromises. *Clin Psychol Rev* 7(6):695-714, 1987.
- Orwin, R.G., and Cordray, D.S. Effect of deficient reporting on meta-analysis: A conceptual framework and reanalysis. *Psychol Bull* 97:134-147, 1985.
- Peters, D., and Ceci, S. Peer-review practices of psychological journals: The fate of published articles, submitted again. *Behav Brain Sci* 5:187-255, 1982.
- Presby, S. Overly broad categories obscure important differences between therapies. *Am Psychol* 33(5):514-515, 1978.
- Rosenthal, R. *Experimenter Effects in Behavioral Research*. New York: Appleton-Century-Crofts, 1966.
- Rundall, T.G., and Bruvold, W.H. A meta-analysis of school-based smoking and alcohol use prevention programs. *Health Educ Q* 15(3):317-334, 1988.
- Sacks, H.S.; Berrier, J.; Reitman, D.; Ancona-Berk, V.A.; and Chalmers, T.C. Meta-analysis of randomized control trials. *N Engl J Med* 316:450-455, 1987.
- Shadish, W.R., Jr. Do family and marital psychotherapies change what people do?: A meta-analysis of behavioral outcomes. In:

- Cook, T.D.; Cooper, H.; Cordray, D.S.; Hartmann, H.; Hedges, L.V.; Light, R.J.; Louis, T.A.; and Mosteller, F., eds. *Meta-Analysis for Explanation: A Casebook*. New York: Russell Sage Foundation, 1992. pp. 129-208.
- Smith, M.L., and Glass G.V Meta-analysis of psychotherapy outcome studies. *Am Psychol* 32: 752-760, 1977.
- Tobler, N.S. Meta-analysis of 143 adolescent drug prevention programs: Quantitative outcome results of program participants compared to a control or comparison group. *J Drug Issues* 16(4):537-567, 1986.
- Tobler, N.S. Drug prevention programs can work: Research findings. *J Addict Dis* 11(3):1-27, 1992.
- Tobler, N.S. Meta-analytic issues for prevention intervention research. In: Collins, L.M., and Seitz, L.A., eds. *Advances in Data Analysis for Prevention Intervention Research*. National Institute on Drug Abuse Research Monograph 142. NIH Pub. No. 94-3639. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1994.
- Wachter, K.W. Disturbed by meta-analysis. *Science* 1407-1408, 1988.
- Wanous, J.P.; Sullivan, S.E.; and Malinak, J. The role of judgment calls in meta-analysis. *J Appl Psychol* 74(2):259-264, 1989.

ACKNOWLEDGMENTS

The preparation of this manuscript was supported in part by a grant from the National Institutes of Health, National Institute of Nursing Research R01 NR1539-05.

AUTHOR

Elizabeth C. Devine, Ph.D., FAAN
Associate Dean for Research
Associate Professor
School of Nursing
University of Wisconsin-Milwaukee
P.O. Box 414
Milwaukee, WI 53217

[Click here to go to page 147](#)