Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Copyright is retained by the first or sole author, who grants right of first publication to *Practical Assessment, Research & Evaluation*. Permission is granted to distribute this article for nonprofit, educational purposes if it is copied in its entirety and the journal is credited. PARE has the right to authorize third party reproduction of this article in print, electronic and database forms.

Volume 2, Number 8, November, 1991 ISSN=1531-7714

Meta-Analysis in Educational Research.

Robert L Bangert-Drowns,
SUNY

Lawrence M Rudner
ERIC Clearinghouse on Assessment and Evaluation

"I had hoped to find research to support or to conclusively oppose my belief that quality integrated education is the most promising approach. For every study that contains a recommendation, there is another, equally well documented study, challenging the conclusions of the first...No one seems to agree with anyone else's approach. But more distressing: no one seems to know what works."

Senator Fritz Mondale's quote illustrates a common plight. Educational research often produces contradictory results. Differences among studies in treatments, settings, measurement instruments, and research methods make research findings difficult to compare. Even frequent replications can prove inconclusive. Literature on a topic may be so extensive as to obscure trends with an overwhelming amount of information.

Meta-analysis is a collection of systematic techniques for resolving apparent contradictions in research findings. Meta-analysts translate results from different studies to a common metric and statistically explore relations between study characteristics and findings.

This article first describes meta-analysis as a research method. The need and general approach are discussed. We then identify some common approaches toward conducting meta-analysis in education and outline their advantages and disadvantages.

META-ANALYSIS AS A RESEARCH METHOD

Gene Glass first used the term "meta-analysis" in 1976 to refer to a philosophy, not a statistical technique. Glass argued that literature review should be as systematic as primary research and should interpret the results of individual studies in the context of distributions of findings, partially determined by study characteristics and partially random. Since that time, meta-analysis has become a widely accepted research tool, encompassing a family of procedures used in a variety of disciplines. A recent search of the ERIC database identified over 800 articles written after 1980 that use or discuss meta-analysis.

Meta-analysis responds to several problems in educational research. First, important issues are studied by numerous investigators. The amount of information on a given topic therefore is often overwhelming and not amenable to summary. Even when there are relatively few studies on a given topic, it is difficult to determine if outcome differences are attributable to chance, to methodological inadequacies, or to systematic differences in study characteristics. Informal methods of narrative review permit biases to remain easily undetected. Reviewers' biases can influence decisions about study inclusion, relative weights given to different findings, and analysis of relations between study features and outcomes. These biases can have clandestine effects when reviewers do not systematically seek to reduce them or provide sufficient information for readers to evaluate their extent.

Meta-analysis typically follows the same steps as primary research. The meta-analyst first defines the review's purpose. Organizing frameworks can be practical or theoretical questions of varying scope, but they must be clear enough to guide study selection and data collection. Second, sample selection consists of applying specified procedures for locating studies that meet specified criteria for inclusion. Typically, meta-analyses are comprehensive reviews of the full population of relevant studies. Third, data are collected from studies in two ways. Study features are coded according to the objectives of the review and as checks on threats to validity. Study outcomes are transformed to a common metric so that they can be compared. A typical metric in educational research is the effect size, the standardized difference between treatment and control group means. Finally, statistical procedures are used to investigate relations among study characteristics and findings.

Criticisms of meta-analysis tend to fall into two categories. Some complain that meta-analysis obscures important qualitative information by "averaging" simple numerical representations across studies. Other critics argue that research is best reviewed by a reflective expert who can sift kernels of insight from the confusing argumentation of a field.

META-ANALYTIC APPROACHES

VOTE-COUNTING -- Some reviews categorize findings as significantly positive (favoring the treatment group), significantly negative, or nonsignificant. The category with the most entries is considered the best representation of research in this area. This as an inexact approach to integrating research. Vote-counting confuses treatment effect and sample size because statistical significance is a function of both. Given the modest power of typical educational research to detect true effects as statistically significant, conclusions from vote-counting can be very misleading.

CLASSIC OR GLASSIAN META-ANALYSIS -- Glass' early meta-analyses set the pattern for conventional meta-analysis: define questions to be examined, collect studies, code study features and outcomes, and analyze relations between study features and outcomes. These early meta-analyses, and later ones following this tradition, share three distinguishing features. First, "classic" meta-analysis applies liberal inclusion criteria. Glass argued that one should not disregard studies on the basis of study quality a priori; a meta-analysis itself can determine if study quality is related to variance in reported treatment effect. Second, the unit of analysis is the study finding. A single study can report many comparisons between groups and subgroups on different criteria. Effect sizes are calculated for each comparison. Third, meta-analysts using this approach may average effects from different dependent variables, even when these measure different constructs.

Glassian meta-analysis has proven quite robust when submitted to critical re-analysis. Its use of conventional statistical tests render the method and its results accessible to most educational researchers. However, using study findings as the units of analysis produces nonindependent data and gives greater weight to studies with many comparisons. Averaging across constructs and including studies with obvious methodological flaws can confuse the reliability of findings.

STUDY EFFECT META-ANALYSIS -- Study effect meta-analysis alters the Glassian form in two ways. First, inclusion rules are more selective. Studies with serious methodological flaws are excluded. Second, the study is the unit of analysis. One effect size is computed for each study. This preserves the independence of the data and gives equal weight to all included studies. Unfortunately, it also reduces the number of data points analyzed in the review. And, of course, a reviewer's biases may operate in decisions to exclude studies.

TESTS OF HOMOGENEITY -- Some reviewers argue that conventional statistical tests are inappropriate for meta-analysis. Homogeneity tests were developed to determine the likelihood that variance among effect sizes is due only to sampling error. If the homogeneity statistic is significant for a group of studies, a procedure analogous to analysis of variance can be used. Studies are repeatedly divided into subgroups according to study features until within-group variation is nonsignificant.

Numerous factors can cause variation in effect sizes: measurement unreliability, range restrictions, reporting errors, within-study statistical adjustments, unreported factors, etc. Homogeneity tests are very likely to indicate heterogeneity among effect sizes even when the variation is of no practical or theoretical importance. Successively dividing subgroups according to these tests can capitalize on chance and cause the incorrect identification of moderators. Kulik and Kulik defend conventional analysis of variance for meta-analysis and suggest that homogeneity tests may ignore an important nesting factor.

PSYCHOMETRIC META-ANALYSIS -- Hunter and Schmidt's approach to meta-analysis combines some of the best features of other approaches. All studies related to a given topic are gathered, regardless of quality. The distribution of effect sizes is corrected for sampling error, measurement error, range restriction, and other systematic artifacts. If the remaining variance is still large, effect sizes are grouped into subsets according to preselected study features, and each subset is meta-analyzed separately. Ideally, the meta-analysis should estimate true treatment effects under conditions typical of those represented in the studies and predict treatment effects under conditions determined by the reviewer. Unfortunately, this technique requires substantial information from individual studies for accurate correction of effect sizes. This information is not always available in research reports.

SUGGESTED READING

Bangert-Drowns, R.L. (1986). Review of developments in meta-analytic methods. *Psychological Bulletin,* 99, 388-399.

Glass, G.V, McGaw, B., & M.L. Smith (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.

Hunter, J.E., & F.L. Schmidt (1990). *Methods of meta-analysis*. Newbury Park, CA: Sage.

Kulik, J.A., & C.-L.C. Kulik (1989). Meta-analysis in education. *International Journal of Educational Research*, 13, 221-340.

Descriptors: Comparative Analysis; Data Collection; Definitions; *Educational Research; *Effect Size; Evaluation Methods; Literature Reviews; *Meta Analysis; Psychometrics; *Research Methodology; Sampling; Statistical Data

Citation: Bangert-Drowns, Robert L. & Rudner, Lawrence M. (1991). Meta-analysis in educational research. *Practical Assessment, Research & Evaluation*, 2(8). Available online: http://PAREonline.net/getvn.asp?v=2&n=8.