A critique of contemporary methods of research synthesis

Harsh Suri

The University of Melbourne

Introduction

A single study can rarely provide a generalisable and definitive answer to a research question focused within the social sciences, (Cooper, 1989; Hunter, Schmidt & Jackson, 1982; McGaw, 1997). Results of a single study are frequently influenced by sampling characteristics such as the sample population, study setting, and timing. The research environment is often difficult to control and human behaviour complex to explain (Wolf, 1986). In many areas, particularly Education, economic constraints may restrict the scale of any single study (Draper et al., 1992). As a consequence, the comprehensive investigation of an area, such as numeracy, may require the combination of results from several individual studies.

More causal factors of a particular effect are likely to be detected by a research synthesis than by a single study (Cook et al., 1992). Often different individual studies provide conflicting results which can have confusing implications (Wolf, 1986). Knowledge in the social sciences, therefore, should progress by recognising the generalisable trends and underlying principles across a large body of empirical studies (Niemi, 1986). Synthesis of primary research is also important to transmit the accumulated knowledge to lay persons and to determine the direction of subsequent research, policies and practice (Cooper & Rosenthal, 1980; Sandelowski, Docherty & Emden, 1997).

Research review plays an important role in dissemination of knowledge and in shaping further research and practice. Therefore the methodology of research synthesis is crucial (Glass, McGaw & Smith, 1981; Dunkin, 1996). Traditional narrative reviews, meta-analyses and best-evidence syntheses are three frequently used methods of synthesising primary research in educational research review journals such as Review of Educational Research.

This paper highlights the relative strengths and weaknesses of the contemporary methods of research synthesis and proposes a multi-stage approach to research synthesis that draws on the strengths of each of these individual methods.

Traditional narrative reviews of research

Traditional literature reviews are often narrative reports of an intuitive aggregate of individual research findings (Johnson, 1989). These reviews are flexible in their methodology and can be undertaken effectively by an experienced research reviewer. But this flexibility can be associated with a high level of subjectivity that may explain inconsistencies in the conclusions of different reviews on the same issue. The criteria for the inclusion of particular studies in a narrative review have not always been made sufficiently clear which makes it difficult for the reader to fully appreciate the effect of the reviewer's theoretical position on the review's findings. Different empirical studies may use different methodologies and precision levels, which in turn are handled differently by different reviewers (McGaw, 1997).
Traditional narrative reviews are often inconclusive, especially when the review includes several individual findings supporting conflicting hypotheses. Therefore, when compared to statistical procedures, traditional narrative reviews are more inclined to have type II error (Cooper & Rosenthal, 1980). These reviews usually ignore unpublished research, which in turn, introduces a publication bias (Glass, McGaw & Smith, 1981). Thus, at times different traditional reviews may even consistently misrepresent the literature and be biased in one direction (Wolf, 1986).

At times, these reviews use a "voting method" to determine if an effect exists. In a voting method, all the findings are divided into three categories: those with statistically significant results in one direction; those with statistically significant results in the opposite direction; and those with statistically insignificant results. This method tends to give equal weight to studies with different sample sizes and effect sizes at varying significance levels, resulting in misleading conclusions. No matter what conclusion is reached, a major problem remains to determine the size of the effect (Abrami, Chambers, Poulsen, De Simone, d'Apollonia & Howden, 1995; Hunter, Schmidt & Jackson, 1982). Further, these methods often fail to identify the variables, or study characteristics, that could moderate the effect (McGaw, 1997).

Meta-analysis

Glass (1976) argued that variability and uncertainty of data in research synthesis are as evident as in the data analysis of primary research. Hence research synthesis requires the same statistical rigour as is demanded in the data analysis of an empirical study. With these views in mind, he proposed a statistical method of research integration which he called "meta-analysis". Meta-analysis is the quantitative integration and analysis of the findings from all the empirical studies relevant to an issue and amenable to quantitative aggregation. It not only quantifies the effect of a treatment, but also identifies potential moderator variables of the effect. In a meta-analysis, findings from different studies are expressed in terms of a common metric called the effect size. In general, the effect size is the difference between the means of the experimental and control conditions divided by the standard deviation (Glass, 1976; Wolf, 1986).

Meta-analysis has several advantages over traditional narrative review. It not only shows the direction of the effect of a treatment, but also quantifies the effect and identifies the moderator variables. It includes all the quantitative empirical studies relevant to the research question and should be free from the subjectivity introduced by selective sampling. The criteria used for selecting the findings included in the synthesis is explicitly stated to remove any unstated ambiguity (Hunter, Schmidt & Jackson, 1982). Meta-analysis can provide a general conclusive answer to a question (Glass, McGaw & Smith, 1981). It is sufficiently robust to deal with a large number of empirical studies (McGaw, 1997).

However, meta-analyses are not free from criticisms. They are prone to overgeneralise, include results from poorly designed studies, be biased in favour of published research in comparison to unpublished research, give more weight to studies with multiple results and ignore studies for which the effect size cannot be computed. In particular, qualitative studies are inevitably excluded from such research syntheses (Slavin, 1986).

Best-evidence synthesis

To overcome the limitations of the methods of traditional narrative review and meta-analysis Slavin (1986) proposed the method of "best-evidence synthesis" which, in theory, draws on
the strengths of the methods of traditional narrative review as well as meta-analysis. According to Slavin, best-evidence syntheses incorporate the statistical rigour of meta-analyses to synthesise quantitative findings together with the flexibility of traditional narrative reviews. The method is freed from unacknowledged subjectivity by including well-justified and well-described inclusion criteria for empirical studies (Slavin, 1986).

The method of best-evidence synthesis does not prescribe a rigid set of criteria for selecting the empirical studies. Like traditional narrative reviews, best-evidence syntheses allow for the individual differences in priorities from review to review. Like meta-analyses, best-evidence syntheses explicitly state the criteria for including or excluding the individual research reports. Best-evidence syntheses do not exclude all the studies for which computation of the effect size is not possible. Unlike meta-analyses, best-evidence syntheses are not limited to statistical aggregation and analysis of only quantitative findings from individual studies. In this method, statistical analysis is supplemented with a rich literature review which explains any discrepancies observed and summarises the results which cannot be quantified (Slavin, 1986).

A closer inspection of best-evidence syntheses reveals some major differences in the meta-analytic aspect of Slavin's method and the contemporarily acceptable meta-analytic procedures. For instance, unlike meta-analyses, best-evidence syntheses take the median effect size rather than the appropriately weighted mean effect size as the pooled effect size (Slavin, 1986; Veenman, 1995). While Slavin's modifications are rarely referenced in the meta-analytic literature, contemporary meta-analytic procedures are successors of rigorous criticisms and modifications, as evident in the vast literature on different aspects of meta-analysis. Slavin's method also fails to provide guidelines for systematic and rigorous methods of synthesising qualitative research.

**Synthesis of qualitative research**

Qualitative researchers argue that synthesis of qualitative research should be interpretive rather than aggregative. While preserving the integrity and holism of individual studies, inductive and interpretive techniques should be used to sufficiently summarise the findings of individual studies into a product of practical value.

According to Jensen and Allen (1996), an interpretive synthesis is essentially a reciprocal translation of key metaphors of each study in terms of the key metaphors of other studies. Hence, they argue that an interpretive synthesis should include studies that use similar methodologies only. For instance, reports within an interpretational paradigm should not be synthesised with those within a narrative paradigm. However, reciprocal translational synthesis is only one of the possible types of interpretive synthesis. "Refutational" synthesis and "lines of argument" synthesis are two other forms of interpretive synthesis advocated by . The criteria for inclusion of individual studies should be based on conceptual considerations rather than only methodological considerations.

The purpose of an interpretive synthesis of qualitative research is not to generate predictive theories, but to facilitate a fuller understanding of the phenomenon, context or culture under consideration. Hence policy making should be informed not only by qualitative research findings, but also quantitative research findings.
Complementing the quantitative and qualitative approaches

Several criticisms of each of the above mentioned methods of synthesis are not specific to meta-analyses, best-evidence syntheses or interpretive syntheses *per se*, but can be generalised to every research synthesis method. Likewise, issues of rigour at various stages of synthesis are often comparable across different methods of research synthesis. Instead of arbitrarily excluding any body of literature because of its methodological paradigm, a good research synthesis should comprehensively include quantitative as well as qualitative findings. The quantitative and qualitative approaches should be complementary rather than adversarial.

In educational research, the researcher often does not have control over all the variables. Leinhardt and Leinhardt (1997), therefore, emphasise the importance of exploratory data analysis. They urge educational researchers to immerse themselves in their data and let the procedures for analysis be guided by the nature of the data, before performing inferential statistics.

This paper argues that a similar inductive approach is also required in the synthesis of educational research. The notion of inductive analysis is not exclusive to the synthesis of qualitative findings. The approach is equally applicable in the meta-analytic synthesis of quantitative findings where the selection of particular statistical techniques for data analysis should be determined by the nature of the data, rather than any rigidly prescribed rules. For example, the spread of individual effect sizes should be examined before deciding whether to use parametric or non-parametric tests for statistical analyses (McGaw, 1997). Consistent with the spirit of exploratory data analysis, the preliminary data analysis may also be enriched through the use of graphs and visual displays of data (Light & Pillemer, 1984).

The criteria for inclusion of individual studies should be conceptual rather than only methodological. All the individual studies relevant to the particular context, concept, culture, or strategy under examination should be included in the synthesis. Within each phase, studies using similar methodologies and research focus should be synthesised first using an aggregative or reciprocal translational approach. The findings thus obtained from each subset may further be synthesised at the next level. In the initial phase, an inductive and interpretive approach should be used to synthesise qualitative findings. This should be followed by meta-analytic synthesis of quantitative findings. Finally, the synthesist should compare the findings from the above two phases and conduct an inductive and interpretive synthesis at a higher level.

Summary

As research reviews play an important role in the dissemination of knowledge and in shaping future research and practice, the methodology of research synthesis is crucial. Traditional narrative reviews, meta-analyses, best-evidence syntheses, and qualitative research syntheses, have their own strengths and weaknesses. This paper argues that a comprehensive research synthesis should include quantitative as well as qualitative research findings. The process of synthesising research should be inductive and interpretive rather than a rigid set of procedures and techniques.
References:


Cambridge: Harvard University Press.


365-71.

