This is chapter 1: The growth of meta-analysis and implications for methodological controversies (pp. 3-7) from

Schulze, R. (2004). *Meta-analysis: A comparison of approaches*. Hogrefe & Huber.

1

The Growth of Meta-Analysis and Implications for Methodological Controversies

The research literature in most fields of science is steadily growing at a seemingly ever increasing rate. Nowadays, it appears to be virtually impossible for a researcher even in a relatively restricted field of study to keep track of all relevant published articles. Hence, there is a strong need for summaries of recent theoretical and empirical results in all scientific areas. Traditionally, there are reviews published in periodicals like the Annual Reviews, for example, where experts of the field are invited to present the current state of a field of study. Besides the function to inform interested researchers about the recent developments and findings such reviews are also relevant for an evaluation of the state of knowledge of a scientific area and even to guide decisions of policymakers to find scientifically well-founded solutions for everyday problems. However, clear-cut summaries of a research field are only easily established with a fairly consistent empirical basis, which is rather an exception than the rule, at least in the social sciences.

As Hunter and Schmidt (1996) have described for the field of psychology, making sense of heterogeneous results can be rather frustrating not only for researchers but also for policymakers. This may have the adverse effect of a negative appraisal of a whole scientific area potentially leading to cuts in funds and bad reputation. This kind of situation characterized the state of affairs in psychology in the early 1970s in the United States, with the negative consequences just described. It was in this climate when researchers became more occupied with the way summaries and reviews were actually carried out. Although the problem of summarizing the state of knowledge was not an entirely new one, the scientific examination of the review process itself was immensely intensified from this time on and ideas on the methods for a syn-

thesis of research began to appear in publications (e.g., Light & Smith, 1971). Yet it was not until Gene Glass coined the term *meta-analysis* (Glass, 1976) that the ways to conduct literature reviews and the synthesis of empirical evidence in a field of study became a research area of its own. From this point of time on increasing research activity was devoted to the development of guidelines and techniques for the conduct of systematic reviews now having its own name meta-analysis. However, meta-analysis was not associated with the invention of a new research problem, as Olkin (1990) has highlighted (see also Hunt, 1997), but with calls for more procedural and statistical rigor in the preparation of literature reviews. It is this rigor that still most prototypically marks the difference between traditional reviews and meta-analysis.

However, this was not the only attribute which appealed to members of the scientific community. The introduction of meta-analysis to the statistical toolbox was not totally detached from substantive problems. The motivation for its development was sparked by the interest to find answers to two very important problems in psychology, namely the comparison of the effectiveness of psychotherapies on the one hand, and the situational specificity of predictive validities of personnel selection procedures in occupational settings on the other. The former problem motivated Glass and co-workers to develop their methods of meta-analysis (see Glass, McGaw, & Smith, 1981). They subsequently published the first meta-analysis in clinical psychology (Smith & Glass, 1977) which provoked great interest[1] as well as harsh criticism of the method (Eysenck, 1978). The latter problem was addressed — coincidentally at the same time — by Schmidt and Hunter, and resulted in the development o6f their methods (Schmidt & Hunter, 1977), followed by applications in the area of personnel selection (for a recent overview, see Schmidt & Hunter, 1998). Thus, meta-analysis forcefully caught the attention by the early 1980s via two routes, methodological rigor and the potential to provide an elegant solution to substantive research problems.

After the inauguration of the term, presentation of procedural details, and publication of the first applications, meta-analysis was quickly adopted in the scientific field, and psychology in particular. This growth of meta-analysis in the past 30 years can be illustrated, for example, by the frequencies of published articles related to meta-analyses.

Figure 1.1 depicts the number of publications up to 2003 that matched the query "meta-analy* or metaanaly* or 'integrative review' " in two of the main databases of psychological research literature: PsycINFO (mainly English literature) and Psyndex (mainly German literature). The "hits" in this literature search represent articles concerning the development and evaluation of the statistical methods as well as applications of meta-analysis in psychology. It is clearly evident that the new field of research is still growing and tends to produce itself an enormous amount of research articles. As a caveat, however, it must be added in this context that the number of articles per year includ-

[1] At the time of writing, this article reached a citation count of 749 in the ISI Web of Knowledge.

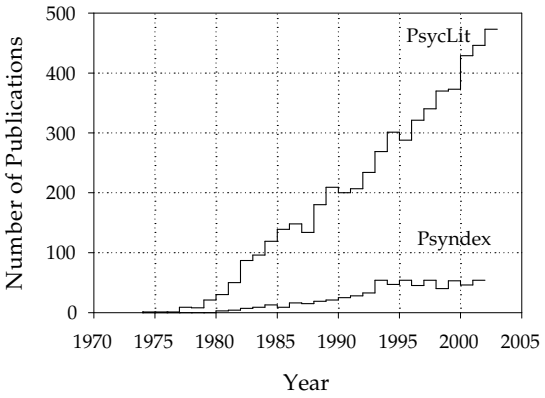

Figure 1.1 Number of publications in the research databases PsycINFO and Psyndex from 1974 to 2003.

ing these search terms may partly reflect expanded journal coverage of the databases. Nevertheless, along with this rising interest in the development and applications of meta-analysis the technique also seems to have been adopted in the canon of research tools in psychology. This is evidenced, for example, by the fact that general introductions to meta-analysis have found their way into general methodological handbooks (Cooper & Lindsay, 1998) as well as treatments of methods in more specific areas like social and personality psychology (e.g., Johnson & Eagly, 2000), organizational psychology (e.g., Holling & Schulze, in press), and clinical psychology (e.g., Durlak, 2003).

In some areas of research in psychology, there is now even a need to summarize applications of meta-analyses to keep track of the main empirical results in a field of study. There are, for example, mainly narrative reviews of meta-analyses for entire subdisciplines of psychology (e.g., Hunter & Hirsh, 1987; Tett, Meyer, & Roese, 1994) as well as more focused and even quantitative reviews (i.e., "meta-analyses of meta-analyses"), for instance on the relationship between personality measures and performance (Barrick & Mount, 2003) and personnel selection (Hermelin & Robertson, 2001). Moreover, the integration of meta-analytical findings can also be used to assess methodological effects in scientific research (e.g., Wilson & Lipsey, 2001). There are even reviews of meta-analyses for psychology as a whole discipline (Lipsey & Wilson, 1993) which are generally favorable in results as far as the effectiveness of psychological treatments is concerned. This fact may also have contributed to the popularity of meta-analysis as a new research tool, because it was associated with the promise of revealing "true" effects of psychological treatments which are otherwise buried in an enormous morass of contradictory study findings.

Interestingly, with respect to the seemingly inconsistent and highly variable results in psychology and related fields already mentioned, the application of meta-analysis also lead to the conclusion that the results in psychology are actually no more variable than results in some quarters in the physical sciences, which often are taken as the standard of so-called hard sciences (Hedges, 1987;

but see also Sohn, 1997, for a contrarian view). This possibly added to the evolving positive attitude towards meta-analysis, at least for those social science researchers somewhat envying their colleagues in the natural sciences for their hard facts.

Moreover, supplementing the expectations of unravelled research controversies by applying meta-analytic methods, meta-analysis was also proposed to even be a useful tool for theory development and testing (Miller & Pollock, 1995; Viswesvaran & Ones, 1995). Hence, in addition to serving an expedient function for the synthesis of extant empirical evidence, meta-analysis may also have the promise to generate new knowledge in a field of study and help in developing and testing new theories. Furthermore, its results may also be used to focus new research efforts and designs on interesting effects emerging from its application (see Czienskowski, 2003).

In addition to these results now rather indicating the successful application of psychological interventions, meta-analysis has also been connected with more far-reaching implications within the realm of epistemological questions of scientific research in the social sciences. The related discussion in methodological quarters of the social sciences centered around the notion of science as an endeavor of the accumulation of knowledge and the way current empirical practices may have to be changed with the methods of meta-analysis at hand, especially the use of significance tests in the social sciences. Whether science in its entirety, and social science in particular, is cumulative in nature is a controversial issue that has its supporters (e.g., Hunter & Schmidt, 1990, 1996; Schmidt, 1992; Schmidt & Hunter, 1995) as well as critics (e.g., Meehl, 1978, 1990; Sohn, 1997). Meta-analysis as a research tool touches upon this issue for it is applied to synthesize current knowledge and its results are supposed to "reveal" or even prove the cumulation of knowledge by "cleaning up and making sense of research literature" (Schmidt, 1992, p. 1179). This conveys the notion of psychology as a research discipline that produces reliable and useful results, a highly welcomed point of view for scientists and the research consumer with a positive attitude towards the social sciences. Furthermore, it has been claimed that meta-analysis is a valid tool to fundamentally change current research practices by *replacing* significance tests, which have been identified as retarding cumulation of scientific knowledge (Rossi, 1997; Schmidt, 1996; Schmidt & Hunter, 1997). Along with this position comes a devaluation of the impact of individual studies and the view to regard them only as data points for a subsequent meta-analysis. Again, there are also critics of meta-analysis that raise serious doubts about the notion of accumulation of knowledge in (social) sciences and support significance testing as a tool in a theory-corroborating scientific approach (e.g., Chow, 1988; Mulaik, Raju, & Harshman, 1997). Although the notion of devaluing individual studies as merely providing data points for a meta-analysis has also been heavily criticized (Harris, 1997; Landy, 2003; Sohn, 1995, 1997). A discussion of these issues can be found, for example, in the volume edited by Harlow, Mulaik, and Steiger (1997) devoted to the significance test controversy in psychology.

However, critics did not only address philosophy of science issues. Applications of meta-analysis were also criticized for various other reasons mostly on substantive grounds (e.g., Eysenck, 1978). Objections were raised, for example, under such headings as "mixing of apples and oranges" to point out potential problems in meta-analyses combining results from studies in which very different characteristics were measured (see also Cortina, 2003), experimental manipulations were different, and so forth. Moreover, the so-called "garbage-in, garbage-out" objection addresses the problem of pooling studies of very different quality, an issue that may, however, be dealt with within the framework of meta-analysis (Wortman, 1994).

In sum, meta-analysis is not only regarded as a new data-analytical tool, but it is also associated with more far-reaching consequences, though the role of meta-analysis in the ongoing significance test controversy in the methodological literature is not yet entirely fixed (Andersson, 1999; Chow, 1996), and the replacement of significance testing by meta-analysis has not taken place to date (Hubbard, Parsa, & Luthy, 1997). This replacement will presumably also not happen in the future since the excessive promises associated with the method are still opposed by challenges of the usefulness of meta-analysis as a method to synthesize the research literature (e.g., Bobko & Stone-Romero, 1998; Chow, 1988). General reviews of meta-analysis as a method also discussing various problems that may be associated with it can be found in detail elsewhere (Beelmann & Bliesener, 1994; Bailar, 1995; Sharpe, 1997).

Finally, it is interesting to note that meta-analysis is now also widely recognized in other sciences like medicine (Dickersin & Berlin, 1992; Normand, 1999; Sutton, Abrams, Jones, Sheldon, & Song, 2000) with positive appraisal in majority (Lau, Ioannidis, & Schmid, 1998; but see also Bailar, 1995; Feinstein, 1995). Controversies like the one described above do not seem to have taken place but other methodological issues are more intensively debated. For example, the value of meta-analysis is challenged on the grounds that it is more of an observational study type and may therefore not lead to reliable causal claims (see e.g., Sauerbrei & Blettner, 2003). This controversy, in turn, is not an issue in psychology where meta-analyses are classified as quasi-experimental (e.g., Farley, Lehmann, & Ryan, 1981). Furthermore, even when considering the same issues researchers in different fields seem to come to different conclusions. For example, whereas in psychology single studies are not given very high value in deciding upon controversial research issues (for reasons, see Gadenne, 1984), single (large) clinical trials are taken as a standard of comparison for the results of meta-analyses. They are even used to judge the validity of claims made on the basis of meta-analytical results. Hence, different research traditions lead researchers to focus on different issues and potential problems of meta-analysis and may fruitfully complement each other in advancing the development of the techniques (for an overview of recent developments in medical and social sciences, see Schulze, Holling, & Böhning, 2003).