JOURNAL OF RESEARCH IN SCIENCE TEACHING VOL. 20, NO. 5, PP. 379-385 (1983)

SCIENCE EDUCATION: A META-ANALYSIS OF MAJOR QUESTIONS*

RONALD D. ANDERSON, STUART R. KAHL,
GENE V GLASS, AND MARY LEE SMITH

University of Colorado, Boulder, Colorado 80309

Abstract

A multi-institutional endeavor was initiated to integrate the findings of extant research studies directed toward the major science education research questions. The research questions were selected by a largely empirical process of identifiying the most frequently researched questions in the literature. These questions were assigned to various researchers who developed coding sheets and procedures with many features in common. This article describes the overall operation of the project, the research questions identified, and some rudiments of meta-analysis. The results of the several meta-analysis are reported in the other articles of this issue of the *Journal.* The final article in this issue deals with research topics for which data are drawn from one or more of the separate meta-analyses.

While meta-analysis (Glass, 1976) has been on the educational research scene only a few years, it has become established as an important technique. It is proving useful in translating the results of numerous studies on a particular topic into a concise form that is reflective of the multiplicity of data found in the many studies, and understandable to the educational practitioner who may be in a position to apply the results. The characteristics of this methodology and guidelines for employing it are well documented (Glass, McGaw, & Smith 1981). While this approach already has been utilized for several science education questions, it has additional potential value if applied to the wide sweep of major science education research questions in a systematic manner. Such an approach requires focusing on the major research questions in the field, giving attention to various subquestions subsumed under each major question and examining common themes that cut across the major questions.

A project of this design was conducted with support from the National Science Foundation. Within the conceptual framework described above, a large number of research studies were integrated with the results providing a basis for interpretive and integrative statements about the major questions addressed in the science education research literature.

A Multi-Institutional Endeavor

Although primarily conducted at the University of Colorado, major portions of the project work were done under a multi-institutional arrangement involving researchers from six other in-

This material is based upon work supported by the National Science Foundation under Grant No. SED 80-12310. Any opinions, findings, and conclusions or recommendations expressed in the publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Published by John Wiley & Sons, Inc. CCC 0022-4308/83/050379-07$01.70

stitutions. A leading researcher from each of these institutions constituted an advisory committee to aid in identifying the research questions pursued and assisted in designing an endeavor encompassing the work of one or more researchers from their home institutions in this project. The members of this advisory committee were as follows:

J. Myron Atkin, Stanford University
Robert Howe, Ohio State University
James Okey, University of Georgia
Lee Shulman, Michigan State University
James Shymansky, University of Iowa
Wayne Welch, University of Minnesota

The actual coding and analysis work was conducted by researchers located at the indicated six research centers and the University of Colorado. At each location an individual or a team of up to three researchers conducted this work (the individuals involved are identified through the references given at the end of this article).

Prior to beginning this coding and analysis work, all of the researchers attended a week-long session for training and coordination of work. During this time each individual or team developed the initial version of the coding forms with a large percentage of the categories and format in common. This process resulted in a database which can be examined across research questions.

This multi-institutional approach had both advantages and disadvantages. It was possible to involve a large research group which was not already extant at one institution. It had the further advantage of stimulating meta-analysis work in a variety of locations where in many cases it was not already underway. One of the disadvantages was the inability to readily shift manpower among questions as their scope became more clearly identified during the actual coding process. As a result there is variation in the thoroughness with which the literature has been sampled for each of the research questions. Though this variation is identified here as a disadvantage, it is not a serious problem as indicated in another article (Anderson, 1982).

Identifying the Research Questions

The first step in the project was to identify the major science education questions to pursue. It was accomplished by a combination of: (a) empirical analysis of the extant research, and (b) expert judgment as to the importance of particular questions. Major attention was given to the empirical analysis rather than the expert judgment, however, in that the basic approach was to include whatever empirical analysis showed to be the subject of a substantial number of research investigations.

The first step was initiated by collecting and examining a representative sample of science education research studies. Literature was sampled across time and type of publication and included studies from *The Journal of Research in Science Teaching, Science Education, Dissertation Abstracts,* and the most recent abstracts of presentations for the *National Association for Research in Science Teaching* annual convention. About 300 such research articles were sampled, and the major (as well as subsidiary) questions addressed recorded.

The questions collected were then classified into some broad, general categories. Five persons classified separate portions of the questions into categories. These categories, developed independently by each of the five persons, had much in common. The entire group of five then examined the questions and organized them into a simple classification system. It resulted in 13 general areas encompassing all but a small percentage of studies which neither fit within these 13 categories nor constituted a meaningful grouping themselves.

The researchers then went back to the literature (including the Curtis digests of *Research in Science Education* of several decades ago) to see if additional research questions fit within the framework that had been empirically derived. This cross-validation indicated the categories were appropriate.

The next step was to develop a full description of each of these 13 areas. They were identified by a generic question for each area along with sample subquestions. These sample subquestions were examples of a larger set of such subquestions; they were a representative and not exhaustive set. In addition, definitions of terms, descriptions of some variables, and a limited rationale for considering the questions were provided.

A form was then developed on which responses could be obtained from other science education researchers concerning these categories. Twenty people were mailed a full description of the 13 areas, a response form, and a cover letter requesting that they be prepared to discuss the material by phone. All 20 people responded to a telephone request for their judgments on the relative importance of these questions and the adequacy of the literature for doing a meta-analysis. While these judgments of the relative importance of the questions were of value, the judgments of the adequacy of the literature were largely subordinated by an empirical search of the total science education research.

Literature searches were conducted on a sampling basis to obtain an estimate of the size of the literature and determine if sufficient studies existed for a meta-analysis of each question. Abbreviated computer searches were conducted using databases such as ERIC, *Dissertation Abstracts,* and *Social Science Research.* The citations obtained then were screened to eliminate those items which were not research publications. Subsequent investigation indicated some problems with the manner in which the computer searches had been conducted, so additional searches were done "by hand" as a check. They were done on a sampling basis using selected annual reviews of science education research and *Science Education—A Dissertation Bibliography,* a listing of all doctoral dissertations pertaining to science education conducted between 1950 and 1977. These procedures provided a rough estimate of the size of the literature pertaining to each of the 13 questions.

At this point a two-day conference of the advisory committee was convened to confer with the project staff and produce a final classification of research questions for meta-analysis, as well as identify important variables to include when integrating the research for each question.

One of the original questions ("What are the goals and priorities of science education?") was eliminated due to an insufficient number of empirical studies, even though it was ranked high in importance. The other 12 questions were recombined into a broader set of questions as follows:

I. What are the effects of different curricular programs in science?
II. What are the effects of different instructional systems used in science teaching (e.g., programmed instruction, mastery learning, departmentalized instruction)?
III. What are the effects of different teaching techniques (e.g., questioning behaviors, wait-time, advanced organizers, testing practices)?
IV. What are the effects of different preservice and in-service teacher education programs and techniques?
V. What are the relationships between science teacher characteristics and teacher behaviors or student outcomes?
VI. What are the relationships between student characteristics and student outcomes in science?

While these six questions as stated were pursued initially, some of them were delimited further when subsequent search activities made it clear that they were too broad to complete within the resources of the project.

The Literature Search Process

Identifying and collecting the research studies to be part of a meta-analysis is a major step in the total endeavor. This aspect of the project will be described in terms of the: (a) limitations placed on the studies to be included, (b) search strategies employed, and (c) variations in the literature covered among the major questions within the total project.

Restrictions on Scope of the Questions

Because of the need to keep the meta-analysis to a manageable size and to maintain some degree of commonality among the studies included under a particular question, the following restrictions were placed on the studies to be included.

1. The studies were limited to those conducted in the context of grades K-12.
2. The studies included were limited to those conducted within the United States.
3. For questions I-IV, only those with a control group were included.
4. The studies were limited to those published in 1950 or later.

The Search Process

In a departure form many past meta-analysis, it was decided that the search process would begin with dissertations because of the thoroughness with which data are typically reported therein, and because such a large percentage of research studies are conducted within that context. This process of searching dissertations was greatly facilitated by the existence of the previously mentioned bibliography which lists all doctoral dissertations pertaining to science education conducted between 1950 and 1977. This document lists approximately 3200 science doctoral dissertations; the entire document was systematically examined to identify each potential dissertation which, by title and categorization within the bibliography, appeared to be a potential for the meta-analysis. These approximately 1000 dissertations were obtained on microfilm from the Science and Mathematics ERIC center at Ohio State University. Each dissertation was read to determine if it actually pertained to the topic at hand and, if so, it was utilized in the meta-analysis.

Another facet of the search process was screening the bibliographies in each coded publication to identify additional studies to be included in the meta-analysis. In addition to identifying journal articles through this standard bibliographic search method, ERIC searches and simple screening of the entire collection of issues for the relevant years of selected journals were conducted. Among the various research sites, the procedures for identifying journal reports to be included varied considerably. Whatever mechanism was used, a high percentage of the articles located were reports of studies already coded from dissertations. Finally, some studies utilized in this meta-analysis were reported in other sources such as books or unpublished reports.

Variations in Literature Covered

While there was considerable variation in the amount of literature covered among the several research sites, there was consistency in removing many studies from consideration without coding them once they had been read and their exact character ascertained. While 769 studies were coded, nearly 2000 studies were read in the process. Among the reasons for excluding studies were the following:

a. The most common reason for eliminating a study was inadequate reporting, i.e., not enough information was provided to make it possible to calculate an effect size.
b. The study did not utilize a control group.
c. The study was not within the K-12 limit; most studies eliminated were college level.
d. The study was conducted outside the United States.

Even given this limiting of the studies included, many of the researchers were faced with a body of literature larger than was possible for them to code and analyze completely within their time limitations. The means of limiting the number of studies varied from one site to another, but generally were one of the following three approaches: (a) Some sites found it possible to code and analyze essentially the entire body of literature located through the search procedure described above and contained within the boundaries cited earlier. (b) Some sites chose to limit the scope of their original question to one or more key subquestions. (c) Some maintained the scope of their coverage but selected only a portion of the studies for analysis.

Coding the Studies

Meta-analysis endeavors are very labor-intensive; the most time consuming part is reading each study and recording on the coding sheets each relevant piece of information. Of the dozens of items of information potentially available for a given study, the major one is an effect size that provides a quantitative comparison of the effects of the experimental and control group (or in the case of a correlational study, the correlation between two variables). For an experimental study, an effect size is calculated which provides a normalized measure of the difference in performance of the two groups with respect to a specified dependent variable such as achievement, attitude toward science, or any other outcome variable. Symbolized by the Greek letter Δ and abbreviated E.S., effect size is defined as the difference between the given variable for the experimental group and control group divided by the standard deviation of the control group.

$$\Delta = \frac{\bar{X}_+ - \bar{X}_c}{S_c}$$

where

$\bar{X}_+$ = mean of experimental group,
$\bar{X}_c$ = mean of control group, and
S_c = standard deviation of the control group.

The calculations involved in determining the effect size vary considerably depending upon the particular form of the data reported in a given study. The numerous procedures required in the various situations including those for computing average correlation coefficients, are well developed and described in Glass, McGaw, and Smith (1981).

Integrating the Results

Once the coding (recording information on all demographic, independent, and dependent variables available in the report) for all of the studies in the meta-analysis has been completed,

attention is turned to integrating this information. This step involves calculating an *average effect size* (a simple arithmetic average) from all those obtained on a given outcome variable such as achievement (and/or some particular category of achievement), attitude toward science laboratory skills, or whatever outcome variable has been examined within some subset of the studies involved. Furthermore, an average effect size can be calculated for a particular *outcome* variable from all studies with a particular *independent* variable and this average effect size then can be compared to the average effect size on the same outcome variable for those studies having a different independent variable. For example, in the meta-analysis of studies of instructional systems in science (at K-12 levels), the average effect size on cognitive achievement for five studies of audio-tutorial systems was 0.09 standard deviations higher than the control groups, while the average effect size on cognitive achievement for seven studies of "Keller Plan" systems was 0.49 standard deviations higher than their control groups. This same type of comparison can also be made for other outcome variables. For example, one of the audio-tutorial studies had an affective measure; it was an effect size of 0.33 in favor of the experimental group. Two "Keller Plan" studies had an affective measure with an average effect size of 0.52. Similar statements can be made about these two instructional systems with respect to any other outcome measures included in some of the studies, and similar comparisons can be made with other instructional systems with respect to any outcome measures included in studies of these systems.

A variety of issues have been raised about the interpretation of such results as described above. For a discussion of the issues the reader is referred to the recent JRST article on the topic (Glass, 1982) or the recent book on the topic (Glass, McGaw, & Smith, 1981).

Project Results

The results of the meta-analysis in this project are reported in several research papers contained in this issue of the *Journal of Research in Science Teaching* (Druva & Anderson, 1983; Lott, 1983; Malone & Fleming, 1983; Shymansky, Kyle, & Alport, 1983; Sweitzer & Anderson, 1983; Willett, Yamashita, & Anderson, 1983; Wise & Okey, 1983). They include one research article associated with each of the previously identified questions (two articles in the case of question III) and an article (Anderson, 1983) dealing with research issues for which data are drawn from one or more of the separate meta-analyses. Brief descriptions of the data files acquired are provided in each of the individual research papers. The total database has been compiled on one master file at the University of Colorado and is available, along with a User's Manual (Kahl & Anderson, 1982), to other researchers who wish to use it. Copies of the coding sheets used and the complete bibliography of research studies coded are provided both in the User's Manual and the final report of the project submitted to the National Science Foundation (Anderson et al., 1983).

References

Anderson, R. D. A consolidation and appraisal of science education meta-analysis. *Journal of Research in Science Teaching,* 1983, **20**(5), 495-507.

Anderson, R. D., Kahl, S., Glass, G. V, Smith, M. L., Fleming, M. L., & Malone, M. R. *Science meta-analysis project: final report of N.S.F. project no. SED 80-12310.* Boulder, CO: Laboratory for Research in Science and Mathematics Education, University of Colorado, 1982, 797 pages.

Druva, C., & Anderson, R. D. Science teacher characteristics by teacher behaviors and student outcomes: a meta-analysis. *Journal of Research in Science Teaching,* 1983, **20**(5), 467-479.

Glass, G. V. Primary, secondary, and meta-analysis of research. *Educational Researcher,* ,976, 5(10), 3-8.

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

Kahl, S., & Anderson, R. D. *Science meta-analysis project: user's guide for the machine-readable raw data file.* Boulder, CO: Laboratory for Research in Science and Mathematics Education, University of Colorado, 1982, 137 pages.

Lott, G. W. The effect of inquiry teaching and advance organizers upon student outcomes in science education: a meta-analysis of selected research studies, *Journal of Research in Science Teaching,* 1983, **20**(5), 437-451.

Malone, M. R., & Fleming, L. The relationship of student characteristics and student performance in science as viewed by meta-analysis research. *Journal of Research in Science Teaching,* 1983, **20**(5), 481-494.

Shymansky, J. A., Kyle, W. C., Jr., Alport, J. M. The effects of new science curricula on student performance. *Journal of Research in Science Teaching,* 1983, **20**(5), 387-404.

Sweitzer, G. L., & Anderson, R. D. A meta-analysis of research on science teacher education practices designed to produce outcomes associated with inquiry strategy, *Journal of Research in Science Teaching,* 1983, **20**(5), 453-466.

University Microfilms International. *Science education—a dissertation bibliography.* Ann Arbor, MI: Dissertation Publishing, University Microfilms International, 1978.

Willett, J. B., Yamashita, J. J. M., & Anderson, R. D. A meta-analysis of instructional systems applied in science teaching. *Journal of Research in Science Teaching,* 1983, **20**(5), 405-417.

Wise, K. C., & Okey, J. R. A meta-analysis of the effects of various science teaching strategies on achievement. *Journal of Research in Science Teaching,* 1983, **20**(5), 419-435.

Manuscript accepted January 17, 1983