

Evaluating Educational Research

- 1 Serious doubt has been expressed as to whether the majority of published educational research articles merit publication on the basis of their quality. For example, a recent editor of the *Review of Educational Research* has suggested that: "Probably, on the average, only 10 per cent of published papers in educational journals are worthy of being reported in the *Review*,"¹ and a professional logician has stated that: "By minimum acceptable research standards, 95 per cent of the work in the field ... that is concerned with causal analysis is, by either theoretical or practical standards, invalid or trivial."² Until recently, however, no attempt has been made to apply the methods of educational research to determine the quality of educational research articles. In 1962 the American Educational Research Association established a Committee on Evaluation of Research whose objectives included (1) obtaining evaluative judgments concerning the quality of educational research studies published in journals, and (2) analyzing the major weaknesses of such research. The portions of the committees' report³ which deal with these two objectives are summarized in the following section.

Summary of Proportions of "An Evaluation of Educational Research Published in Journals"

Edwin Wandt,
Georgia W. Adams, Dorothy M. Collett,
William B. Michael, Ryans, David G. Ryans,
Carleton B. Shay

Problem

- 2 The following questions were asked with respect to the quality of educational research articles published in 1962:
 - 3 a. What percent of the educational research articles, published in journals are considered by experts in educational research to be (1) of sufficiently high quality to merit publication without change, (2) in need of some revision to make them acceptable for publication, or (3) so low in quality that they should not have been published?

1. William B. Michael, "Teacher Personnel: A Brief Evaluation of the Research Reviewed," *Review of Educational Research*, 33:443, October, 1963.
2. Michael Scriven, "The Philosophy of Science in Educational Research," *Review of Educational Research*, 30:426, December, 1960.
3. "An Evaluation of Educational Research Published in Journals," (Unpublished) Report of the Committee on Evaluation of Research, American Educational Research Association (Edwin Wandt, Chairman, Georgia S. Adams, Dorothy M. Collett, William B. Michael, David C. Ryans, and Carleton B. Shay), 1965.

- 4 b. How do the experts rate representative educational research articles on specific characteristics related to the quality of research and of research reporting?
- 5 c. Is there a significant difference in quality, both overall and on the specific characteristics, between (1) educational research articles published in education journals and (2) those published in journals of related professions?
- 6 d. Is there a significant difference in quality, both overall and on specific characteristics, between (1) research articles published in education journals which primarily publish research and (2) those in education journals which primarily publish nonresearch articles?
- 7 e. When experts evaluate published research articles as ones which should be rejected or returned for revision, what specific shortcomings do they cite to substantiate their judgments?

Procedures

- 8 *The population of research articles.* The population of articles was defined by (1) identifying a group of journals which publish educational research articles and (2) identifying all articles published in these journals in 1962 to determine which met the criterion of research used by the committee (i.e., statement of problem, data, analysis of data, and conclusions). After the committee eliminated 16 journals which published fewer than three research articles in 1962 a population of 827 educational research articles in 41 journals was available for study.¹
- 9 *The sample of research articles.* A stratified (by journal) random sample of 125 articles was drawn from the population of 827 articles. Table 1 shows the number of articles included in the sample from each journal.²
- 10 *The judges.* Each of the 125 articles was evaluated by a different expert in educational research. In addition, 41 of the articles were evaluated by a second expert so that paired ratings on these articles could be used in a reliability study.
- 11 All the judges were members of the American Educational Research Association, 98 percent had earned doctor's degrees, 63 percent had taught a course in educational research methods, and 75 percent had supervised doctoral dissertations. The median number of years of college teaching was 11, the median number of years of full-time research experience was 7, and the median number of educational research articles published was 16.

1. The journals identified as "publishing educational research" met two criteria: (1) the journal had to be listed in the *Education Index* and (2) the journal had to be cited in the *Review of Educational Research* at least 10 times during the period 1959-1961. Fifty-seven journals met these criteria. Each article published in these journals in 1962 was examined and classified as "research" or "non-research." Of the 57 journals, eight published no research during 1962, and eight other journals published only one or two research articles during that period. These 16 journals were eliminated from the study. The remaining 41 journals published 827 research articles in 1962.

2. Since the journals which published only three research articles in 1962 did not contain a sufficient number of articles to justify including an article from each of them, the articles from these journals were treated as a single group for purposes of sampling. Only three articles were selected at random from this pool of articles, with the restriction that no journal could be represented by more than one article. As a result of this sampling procedure, two journals which published only three research articles in 1962 are not represented in the sample of articles studied. Hence, the number of journals studied was reduced from 41 to 39.

- 12 *The evaluations.* Each expert evaluated his assigned article in three ways. First, the article was evaluated on 25 characteristics which reflected the quality of the research and the adequacy of reporting. Each characteristic was rated on a 5-point scale as follows:

Level of Quality	Description
5	Excellent — A model of good practices
4	Good — A few minor defects
3	Mediocre — Not good, not bad
2	Poor — Some serious defects
1	Completely incompetent — a “horrible example”

- 13 If the characteristic did not apply to the article, the judge was instructed to rate it as “does not apply.” The 25 characteristics are listed in Table 2. Second, the expert was asked to project himself into an editor’s role and make an overall judgment, deciding whether he would (1) accept the article and publish it without revision, (2) ask for revisions and publish the article if the revisions were made, or (3) reject the article. These ratings are referred to as ARR (accept-revise-reject) ratings. Third, each expert was asked to specify the shortcomings of his assigned article if he had recommended either asking for revisions or rejecting the article.

Results

- 14 *Overall quality.* For the entire sample of 125 articles, the experts rated 19 percent as “acceptable” (acceptable for publication without revision), 41 percent as “revise” (acceptable for publication after revision), and 40 percent as “reject” (not acceptable for publication).
- 15 These overall ARR ratings were summarized separately for articles published in “education” journals and those published in “related-profession” journals (psychology, sociology, and medical science). The percents are shown in Table 1 for each of these groups. A statistical test (chi-square) showed that there was a highly significant difference between the ratings assigned to the two groups, with the related-profession journals having the better ratings. The percent of articles rated “accept” was 41 percent for the related-profession journals, as compared with only 7 percent for the education journals. The percent of articles rated “reject” was 52 percent for education journals and 18 percent for related-profession journals.
- 16 The education journals were divided into two groups: (1) those that published a majority of research articles and (2) those that published a majority of nonresearch articles. Comparison of the ARR ratings for articles in each of these groups showed no significant differences. The percents for these groups are also shown in Table 1. Despite the relatively favorable record of related-profession journals, one cannot assume that publication in any journal represents a guarantee of quality. Each journal which was represented by two or more articles in the sample of 125 articles had at least one article which was judged to be either in need of revision or to be so low in quality that it should have been rejected.
- 17 *Ratings on the 25 characteristics.* The 25 characteristics on which the articles were rated are listed in Table 2, with the mean ratings assigned to each of the characteristics for (a) the entire sample of 125 articles, (b) the 81 articles in the education journals, and (c) the 44 articles in the related-profession journals.
- 18 The median of the mean ratings of the 25 characteristics was 3.27 (approximately one-fourth point above mediocre) for the sample of 125 articles. For the education journals the median was 3.02 (mediocre), and for the related profession journals the median was 3.93 (good).

Table 1. Accept-revise-reject ratings of 125 educational research articles published in 1962

Source of the articles	Number of			
	Articles	Accept	Revise	Reject
All educational journals	81	7%	41%	52%
Education journals which primarily publish research articles	37	5%	49%	46%
California Journal of Educational Research	(3)			
Journal of Educational Research	(13)			
Journal of Experimental Education	(5)			
Personnel and Guidance Journal	(9)			
Science Education	(7)			
Education journals which primarily publish non-research articles	44	9%	34%	57%
Adult Education	(2)			
Arithmetic Teacher	(3)			
Audio-Visual Communications Review	(3)			
Bulletin of the National Association of Secondary School Principals	(4)			
Clearing House	(2)			
Education	(1)			
Educational Record	(1)			
Elementary English	(2)			
Elementary School Journal	(3)			
Exceptional Children	(2)			
High School Journal	(1)			
Journal of Higher Education	(1)			
Journal of Negro Education	(2)			
Journal of Research in Music Education	(1)			
Journal of Teacher Education	(2)			
Junior College Journal	(2)			
NEA Research Bulletin	(1)			
Phi Delta Kappan	(1)			
Quarterly Journal of Speech	(1)			
Religious Education	(1)			
School and Society	(1)			
School Science and Mathematics	(1)			
School Review	(2)			
Science Teacher	(1)			
Vocational Guidance Quarterly	(2)			
Volta Review	(1)			
Related profession journals	44	41%	41%	18%
American Journal of Mental Deficiency	(4)			
Child Development	(11)			
Educational and Psychological Measurement	(6)			
Journal of Counseling Psychology	(5)			
Journal, of Educational Psychology	(8)			
Journal of Educational Sociology	(2)			
Journal of Personality	(5)			
Journal Speech and Hearing Disorders	(3)			
All journals	125	19%	41%	40%

- 19 Study of Table 2 shows that many important characteristics were assigned low ratings, particularly for the group of education journals. Note, for example, that the mean rating of the articles from the education journals on characteristic number 21 was only 2.63. This characteristic (conclusions are substantiated by the evidence presented) is of the utmost importance to consumers of research since the conclusions are worthless unless they are substantiated by the evidence presented.
- 20 The means for the 25 characteristics were higher, in every case, for the articles in the related-profession journals than for those in the education journals. For 23 of the 25 characteristics, these differences were statistically significant (.05).

Table 2. Mean ratings of 25 characteristics of 125 educational research articles published in 1962

Characteristic	Mean Rating		
	All Articles	81 articles in educ. journals	44 articles in related prof. jnls.
1 Problem is clearly stated	3.62	3.41	4.00
2 Hypotheses are clearly stated	3.24	3.04	3.53
3 Problem is significant	3.59	3.31	4.09
4 Assumptions are clearly stated	2.73	2.40	3.30
5 Limitations of the study are stated	2.74	2.41	3.33
6 Important terms are defined	3.16	2.84	3.67
7 Relationship of the problem to previous research made clear	3.13	2.60	4.02
8 Research design is described fully	3.35	3.03	3.93
9 Research design is appropriate to the solution of the problem	3.03	2.65	3.67
10 Research design is free of specific weaknesses	2.81	2.42	3.50
11 Population and sample are described	3.35	3.18	3.67
12 Method of sampling is appropriate	2.97	2.85	3.23
13 Data-gathering methods or procedures are described	3.50	3.14	4.14
14 Data-gathering methods or procedures are appropriate to the solution of the problem	3.27	2.99	3.77
15 Data-gathering methods or procedures are utilized correctly	3.38	3.01	4.02
16 Validity and reliability of the evidence gathered are established	2.74	2.49	3.18
17 Appropriate methods are selected to analyze the data	3.24	2.63	3.98
18 Methods utilized in analyzing the data are applied correctly	3.49	3.11	4.12
19 Results of the analysis are presented clearly	3.44	3.11	4.02
20 Conclusions are clearly stated	3.44	3.06	4.11
21 Conclusions are substantiated by the evidence presented	3.11	2.63	3.95
22 Generalizations are confined to the population from which the sample was drawn	3.26	3.07	3.67
23 Report is clearly written	3.46	3.21	3.91
24 Report is logically organized	3.71	3.46	4.16
25 Tone of the report displays an unbiased, impartial scientific attitude	3.69	3.42	4.19

- 21 Only three of the characteristics were rated significantly differently for the two groups of education journals, i.e., (1) those which publish primarily research, as compared with (2) those which publish primarily non-research articles. In each of the three cases the articles in the journals which published primarily research had the higher mean rating.
- 22 *Shortcomings cited by the judges.* In response to the request to specify shortcomings of articles which were rated “revise” or “reject” the judges cited 379 specific shortcomings in 93 of the articles. The most frequently cited areas of shortcoming are shown in Table 3.

The necessity for evaluating all research studies

- 23 The results of the study just summarized leave little doubt that a large percentage of educational research articles published in 1962 contained serious flaws. We lack any evidence that the situation is different for articles published before or after that time; hence the prudent assumption to make, as a consumer of research, is that educational research articles, regardless of year of publication, must be carefully evaluated.
- 24 The reader who turns to the research literature for assistance in solving a problem needs to know whether the conclusions reached in the articles he reads are probably correct. The only sound basis for determining the correctness of the conclusions lies in careful evaluation of the articles. Although there is a difference among journals in the quality of the articles they publish, the reputation of a journal cannot be used in lieu of the evaluation of a specific article. According to the study summarized previously, every journal which had two or more articles in the sample had at least one of those articles rated as “revise” (in need of revision) or “reject” (unacceptable for publication).
- 25 Caution must also be exercised in judging an article by the reputation of the author. Authors with good reputations have been known to publish articles which could not stand up under evaluation. There is no substitute for evaluation of published research. Readers who do not feel capable of evaluating a specific article in which they are interested would be well advised to seek assistance from someone with greater competency in evaluation of research.

Table 3. Shortcomings most frequently cited by experts in evaluating 125 educational research articles published in 1962

<u>Area of shortcoming</u>	<u>Frequency of citation</u>
Results of analysis not clearly presented	33
Incorrect methods used to analyze data	32
Inappropriate or defective design	31
Validity and reliability of the evidence not established	29
Conclusions not substantiated by the evidence	29

Suggestions for evaluating research studies

- 26 Evaluations of research studies are usually either (1) assigned as learning experiences, designed to increase the student’s competence in evaluating and/or conducting research, or (2) undertaken by educators who want to improve some educational practice in their own situation on the basis of relevant research findings.

Evaluation of research articles by students as a learning experience

- 27 Almost all textbooks in educational research methods include a list of criteria to be used in evaluating educational research. In some textbooks, differentiated lists of criteria are presented for evaluating different types of research studies. The 25 characteristics listed in Table 2 can be used in evaluating most research studies. The following questions are suggested as a basis for making evaluative reports of research articles for class discussion:
- (1) What are the specific strengths of the article?
 - (2) What are the specific weaknesses of the article?

- (3) If it were desired to repeat the study, using the same general approach, what changes should be made to improve the quality of the research?
- (4) What approaches, other than that used by the author, could be used to study the same problem?

28 These questions can be used as the sole basis for analysis or in conjunction with questions about such specific aspects of the research as are included in the list of 25 characteristics.¹

Evaluation of research practitioners seeking to improve educational practice

- 29 The educator who wishes to improve some educational practice in his own situation may benefit in several ways from his evaluations of research studies: (1) he may identify particular methods of inquiry which he can use in conducting a local study, (2) the studies may suggest hypotheses which can be tested in a local study, or (3) the studies may provide a solution to his problem which can be used directly, without the necessity of conducting a local study.
- 30 The primary question to be answered in evaluating a research study for the purpose of improving educational practices is: "What conclusions can be justified by the evidence presented in the study?" The second question to be answered is: "Which of these conclusions, if any, apply to the particular situation in which, the reader is interested?" If none of the conclusions can be applied directly, the next question might be: "Which of the conclusions suggest hypotheses which can be tested by a local study?"
- 31 Conclusions reached in educational research studies may be divided into two categories: (1) those which are descriptive and (2) those which state cause-and-effect relationships. Descriptive conclusions by themselves provide no basis for recommendations to improve education. Some cause-and-effect relationship must be established before recommendations for improvement, are justified. To illustrate, suppose that a carefully conducted study concludes that there is a high positive correlation between students' posture ratings and their reading achievement, i.e., the students who have better posture are the better readers, and vice versa. This conclusion described the relationship between two variables, but does not justify any action by the teacher since no causal relationship has been established. Should teachers who desire to upgrade their students' reading achievement devote some effort to improving the posture of their students? Such a course of action assumes that gain in posture is a cause of gain in reading achievement. Should teachers who wish to improve their students' postures concentrate on improving their reading? This course of action assumes that gain in reading achievement is a cause of gain in posture. Neither of these courses of action is justified by the observed correlation since no cause-and-effect relationship was demonstrated. Only by systematically attempting to improve reading by improving posture, or attempting to improve posture by improving reading, could the existence of a cause-and-effect relationship be tested.
- 32 Evaluating a research study with the objective of improving some educational practice cannot be reduced to a mechanical formula, since each research study is unique. However, all research studies do contain some common elements (e.g., problem, data, conclusions); hence, it is possible to suggest a framework within which evaluation can take place.

1. See page 14 for a Bibliography of Guidelines for Evaluating Educational Research.

A framework for evaluating educational research for the purpose of improving some educational practice

- 33 A. Read the entire study to get an overview. Identify the portions of the report that deal with (a) the statement of the problem, (b) the procedures used to collect data, (c) analysis of the data, and (d) the conclusions. Determine whether the data gathered are related to the problem as stated, and whether the conclusions are related to the evidence presented.
- 34 B. Examine the sections dealing with the procedures used to collect data, and the analysis of the data, to determine whether sufficient information has been provided to enable a judgment to be made regarding the quality of the research. Questions to be answered include:
- has the population studied been adequately described?
 - if a sample was used, was the method of sampling specified?
 - are the methods and instruments used to gather data adequately described?
 - are the methods used to analyze the data described?
 - are provisions (if any) made to control variables adequately described?
 - are the results of the analyses presented clearly?
- 35 Unless all of the questions listed above can be answered “Yes,” there is a serious question as to whether any further evaluation of the study is possible. All of the above questions related to *what the investigator did* and *what he found*. Unless complete information is available in these two areas, it is difficult, if not impossible, to continue with the evaluation of the study. While it is desirable that the sections dealing with the statement of the problem and the conclusions be reported fully and clearly, it is possible to make sense out of a study which is deficient in either or both of these respects. The account of what the author did and what he found can be used as a basis for formulating the problem actually studied (which does not always correspond to the problem stated by the author), as well as for formulating valid conclusions on the basis of the evidence presented.
- 36 C. Examine each conclusion stated by the author to, determine whether it is substantiated by the evidence presented. Questions to be answered include:
- have the validity and reliability of the evidence gathered been established?
 - were appropriate methods selected to analyze the data?
 - were the methods utilized in analyzing the data applied correctly?
 - if the conclusion states a cause-and-effect relationship, has the author demonstrated that the stated cause is more probably the cause than any other alternative cause?
 - Determine if other conclusions, not stated by the author, are justified by the evidence presented.
 - Determine which conclusions, if any, can be generalized to the particular educational setting in which the improvement is to be made.

Difficulties involved in establishing cause-and-effect relationships

- 37 One of the most prevalent errors in educational research is to conclude that a cause-and-effect relationship has been proven when it has not. One illustration of such an erroneous conclusion, which has been widely disseminated to the public, is given in the following criticism of educational research:

From the logician's point of view, then, gross deficiencies of self-awareness in educational research exist, although techniques are available for handling most of these difficulties, As long as those in education allow their own institutions to put out written and cinematographic propaganda which seeks support for higher education by arguing that the average income of graduates is so much higher than that of non-graduates as to more than reimburse them for the cost of higher education within very few years (without adducing any grounds

whatsoever for supposing that this connection is in fact a causal connection and is not, for example, due to the higher income group of the families from which college students come) — so long will they fall short of achieving maturity for their own subject. This is an excellent example of an argument which is scientifically unsound and significantly immoral, since it encourages people to spend money on the basis of a belief which is not known to be well founded.¹

- 38 To establish a cause-and-effect relationship, it is necessary to show that the variable named as the cause (e.g., graduation from college) is a more probable cause of the effect (e.g., higher average income) than any other possible cause (e.g., higher income of parents). Unless this is done, any claimed cause-and-effect conclusion does not merit acceptance. Since there are always numerous variables in educational situations which might account for an effect, establishing cause-and-effect relationships is every difficult. Yet such relationships must be demonstrated if progress is to be made since: "One cannot apply anything one learns from descriptive research to the construction of theories or to the improvement of education without having some causal data with which to implement it."²
- 39 One of the most widely used types of educational research which attempts to establish cause-and-effect relationships consists of comparing two methods of teaching a subject to determine whether one of the methods produces a greater amount of achievement by students than does the other method. Assume that method N_1 is used with students S_1 taught by teacher T_1 in environment E_1 with results R_1 , and that method N_2 is used with students S_2 taught by teacher T_2 in environment E_2 with results R_2 . If at the end of the study the students exposed to method N_1 achieve, as a group, results R_1 which are significantly greater than R_2 , which are achieved by the students exposed to method N_2 , what can be concluded about the relative effectiveness of methods M_1 and M_2 ? The answer, obviously, is, "nothing can be concluded," since differences between the students, the teachers, the environments, or some combination of these might just as well account for the observed difference in results.
- 40 Designing research studies in education, in which a large number of interrelated variables can influence the results, is much more difficult than in the physical sciences. Demonstrating that variables other than the hypothesized cause are less probable causes of an effect can be accomplished by (1) controlling these variables or (2) randomizing them.
- 41 In the case of the methods study, differences in students must either be controlled or randomized, if such differences are to be discounted as causes. For example, the influence of possible sex differences in learning could be controlled by assigning an equal number of boys and girls to the groups using the two methods. Many other variables, however, which might be related to the effect could not be controlled by the researcher (e.g., the ability of the parents to help the students with their homework). Still other variables which influence the effect might not even be realized by the researcher. The only way to determine the probability that variables of this type might account for any observed effect is to assign the students at random to the classes using the two methods. Such random assignment can, of course, be combined with control (e.g., if it were decided that the same number of girls was to be included in the two classes, the determination of which girls were to be assigned to each of the classes could still be made by random methods).

1. Michael Scriven, "The Philosophy of Science in Educational Research," *Review of Educational Research* 30: 426-27, December, 1960.

2. *Ibid.*, p. 426.

- 42 It is possible that girls achieve more (or less) than boys with either of the methods studied. It is also possible that method N_1 is more effective with girls than with boys, and that method M_2 is more effective with boys, although for the class as a whole, the methods appear to be equally effective. Proper statistical analysis of the results should include an investigation of these possibilities.
- 43 Unless random assignment of students is used, it is not possible to demonstrate that systematic differences between the groups on some variable (e.g., parents ability to help with homework) was not the cause of any observed difference in results. If randomization is used in assigning the students to the classes, statistical analysis can be utilized to determine whether the observed difference in results is greater than the variation which can be attributed to the random (or chance) assignment of the students to the two groups.
- 44 Differences in teachers may also be responsible for differences in results between the groups using methods M_1 and M_2 . If only two classes are compared, each with a different teacher, it is impossible to know whether results are due to the method or to the teacher. To determine the extent to which differences in results are attributable to differences among teachers, at least two teacher must use each method studied.
- 45 Another frequent alternative explanation for differences found between groups which are taught by different methods is the fact that pupils who receive special attention, as do those who are being taught by a new method, tend to improve their performance for this reason alone. Many studies in education involve comparing results for a "new" method used with one group of students, with results for the "old" method used with other students. Unless some provision is made for insuring equal enthusiasm in the two groups of students and their teachers, the "new" method may yield better results and hence appear to be better than the old method when the difference is attributable, in part, to the greater interest on the part of teachers and students who are in "the spotlight" as part of the new program.
- 46 The preceding discussion should have made it clear that establishing cause-and-effect relationships in education is very difficult. It is generally much easier to see that a research study has not demonstrated that the claimed cause is, in fact, the most probable cause, than it is to design and execute research that will prove the cause-and-effect relationship.
- 47 A frequent error in educational research is the failure to determine whether improvement in one area has been obtained at the expense of a loss in another area. For example, focusing students attention on one desirable objective may influence them to spend more time outside of school on their homework in that area, and less time on some other area. Unless measures are obtained which reflect all important educational objectives which may be affected by the research, it is impossible to know if a desirable outcome has been accompanied by some other undesirable outcome.

The Role of Published Research in Improving Education

- 48 Earlier reference was made to three ways in which published research might be helpful to an educator who was trying to improve some educational practice. These were (1) by suggesting methods which could be used in a local study, (2) by suggesting hypotheses which could be tested in a local study, and (3) by directly providing the solution to the problem. The difficulty of establishing cause-and-effect relationships which can be generalized to local situations makes it clear that direct solutions to local problems will seldom be found in published educational research. Such research will usually be of most value in suggesting research methods and hypotheses to be used in local studies.
- 49 Some educational researchers take the position that educational research should be directed towards discovering generalizations (laws) which can be applied in almost all educational situations. While it

is highly desirable to have such broadly applicable generalizations, educational research is not likely to yield such results until a much higher level of financial support is channeled to competent research staffs with opportunities to control conditions with large numbers of subjects. At the present time, and probably for decades to come, the educator who has a local problem to solve is unlikely to find answers available in the research literature. In most cases, educators can obtain from published research only suggested hypotheses and research procedures which will help them in conducting local studies on local problems.

- 50 The action taken by an educator who is trying to improve an educational practice should depend upon the risks involved in making incorrect decisions. If no great cost or risk is involved, educational practices may be modified on the basis of research conclusions which cannot properly be generalized to the local situation, or on the basis of no research at all. If considerable cost or risk is involved, however, a local research study should be conducted before modifications in educational practices are instituted.
- 51 The goal of all who conduct research, whether it is aimed at producing broad generalizations, or solving local problems should be to make the quality of the research as high as possible, so that the conclusions will stand up under the type of critical evaluation advocated in this book.

Bibliography of Guidelines for Evaluating Educational Research¹

Books

- Best, John W. *Research in Education*. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1959, 278-81, 307.
- Bixler, Harold H. *Check Lists for Educational Research*. New York: Bureau of Publications, Teachers College, Columbia University, 1928, xii + 118 pp.
- Borg, Walter R. *Educational Research: An Introduction*, New York: David McKay Company, Inc., 1963, 326-45.
- Mouly, George J. *The Science of Educational Research*. New York: American Book Company, 1963, 503-504.
- Rummel, J. Francis. *Introduction to Research Procedures in Education*. (2d ed New York: Harper & Row, Publishers, 1964, 268-69.
- Travers, Robert N. W. *An Introduction to Educational Research*. (2d ed.) New York: The Macmillan Company, 1964, 82-88.
- Van Dalen, Deobold B. *Understanding Educational Research*. New York: McGraw-Hill Book Company, Inc., 1962, 366-76.

Periodicals

- Dvorak, Earl A. "General Guide to a Study of Research Reports," *Peabody Journal of Education* 34: 141-44, November, 1956.
- Farquhar, William W., and Krumboltz, John D. "A Check List for Evaluating Experimental Research in Psychology and Education," *Journal of Educational Research* 52: 353-54, May, 1959.
- Symonds, Percival M. "A Research Checklist in Educational Psychology," *Journal of Educational Psychology* 47: 100-109, February, 1956.
- Van Dalen, D. B. "A Research Check List in Education," *Educational Administration and Supervision* 44: 174-81, May, 1958.

1. From: *A Cross-Section of Educational Research* by Edwin Wandt, Professor of Education, California State College at Los Angeles: David McKay Company, Inc. New York.

