

LARRY HARDESTY, NICHOLAS P. LOVRICH, JR.,  
AND JAMES MANNON

## Library-Use Instruction: Assessment of the Long-Term Effects

*The recognition by librarians of the growing importance of evaluating library-use instruction is steadily increasing, as evidenced by reports in the literature. However, much work has yet to be done which uses sophisticated evaluation techniques. This paper reports the follow-up of an earlier study by examining the long-term retention of library-use skills. Through use of pretesting and posttesting, control and experimental groups, aggregate and individual comparisons, multiple regression, and other techniques, the authors concluded that long-term possession of library-use skills is more highly related to library-use instruction than to either inherent intellectual ability or academic diligence. In addition, the authors discuss the appropriateness of quantitative and qualitative methods of evaluation and caution against taking for granted the effective use of evaluation.*

THE OFTEN QUOTED remark about the weather, which is typically but erroneously attributed to Mark Twain,<sup>1</sup> can be applied to academic librarians involved in library-use instruction: that is, there is a good deal of talking about evaluation, but few seem to be doing anything about it. Richard Werking, in his excellent review and critique of the literature evaluating library-use instruction, found published evidence of only a handful of examples.<sup>2</sup> He did note, however, a growing number of articles pertaining to the evaluation of library-use instruction programs and techniques, including a previous article by the authors.<sup>3</sup> These articles play an important role in demonstrating to academic li-

brarians the various techniques that can be used in library program evaluation, and in adding to the developing body of knowledge concerning the effectiveness of library-use instruction.

The earlier article by the authors focused on two particular goals: (1) documenting the effects of library-use instruction on the short-term acquisition of library-use skills; and (2) demonstrating a methodology that could be used successfully in such an evaluation.<sup>4</sup> The authors found that a sample of DePauw University students exposed to library-use instruction programs in their freshman year tended to score higher—to a statistically significant degree—on a paper and pencil test developed by the authors to measure library utilization skills than did a comparable group of students not exposed to library-use instruction. In fact, as measured by the test, the short-term gains of the freshmen were comparable to the library-use skills of graduating seniors.

Werking, in citing a number of librarians associated with library-use instruction, reported that a common complaint about such tests is "the significance of such short-term gains is not likely to be great."<sup>5</sup> As Werking correctly observes, the question of long-term retention of skills is a very important educa-

---

*Larry Hardesty is head of the Reference Department, Roy O. West Library, DePauw University, Greencastle, Indiana. Nicholas P. Lovrich, Jr., is director, Division of Governmental Studies & Services, Washington State University, Pullman. James Mannon is assistant professor of sociology, DePauw University. The authors acknowledge the support of the Council on Library Resources and the National Endowment for the Humanities through their joint College Library Program grant to DePauw University for 1977-1982, a grant that made possible the library-use instruction program and the evaluation described here.*

tional concern. In order to assess the question of long-term retention of library-use skills, the authors have conducted a follow-up of the earlier DePauw University study. The purpose of this article is both to report the results of this follow-up study and to explain the methodology employed so that other librarians may use it in conducting similar evaluations of library-use instruction programs.

#### SAMPLING GROUPS

The present study analyzes data on several samples of DePauw University students. For comparative purposes the authors included a base-line group of ninety-one DePauw University seniors in the 1977 graduating class who received no formal library-use instruction from a librarian while attending DePauw University. A second major sample group consists of 312 seniors in the 1980 graduating class who agreed to complete a questionnaire containing the library-use skills test reported in the earlier article. These 312 students represent a sampling return rate of 70 percent of the entire 1980 DePauw University graduating class, which was surveyed in the spring of 1980. The third sample group consists of a panel of 1980 seniors (eighty-two students) who received formal library-use instruction as freshmen in 1977 and whose scores were reported as part of the earlier study. They are a subset of the 312 seniors responding to the 1980 survey.

The availability of information gathered over approximately a three-year period makes the evaluation of the DePauw library-use instruction program interesting in a number of ways. At the most elemental level, the skill-possession scores of the 1980 seniors can be compared with those of their 1977 counterparts, the students who had no formal library-use instruction. Second, such data can be employed to address the question of whether the degree of exposure to formal library-use instruction is associated with the level of library-use skills. In this connection, it can be determined whether library-use skills are more closely related to library-use instruction than to other plausible predictors of skills possession such as basic intellectual capacity or academic diligence.

In addition to determining the relative degrees of association between skills possession

and academic background and instructional exposure among 1980 graduating seniors, multiple regression analysis can be utilized to determine how much variation in skill possession can be explained by each of the predictors while controlling the effects of the remaining determinants. Finally, the availability of panel data for more than eighty of the 1980 graduating seniors—data that include preinstructional, short-term postinstructional, and long-term postinstructional assessments of library-use skills—allows the direct testing of short-term and long-term library-use skills resulting from library-use instruction and the other predictors.

Because the central question of this evaluation pertains to the long-term effects of the library-use instruction, a brief explanation of the efficacy of a panel study is in order. A panel is a "special type of time-series technique; it measures some attributes of a given sample of people at several moments."<sup>6</sup> In other words, panel studies involve repeated observations of a sample of persons in order to assess changes over time. Panel studies are considered to have great statistical efficiency because individuals in the sample can be compared with themselves at various points in time, thereby reducing extraneous variability, and allowing for direct individual comparison. In short, panels are "useful for studying the effects of specifically introduced measures."<sup>7</sup> This method enabled the authors to select a sample of freshmen students in 1977, provide some of them with a series of library-use instruction sessions, and compare their scores on the skills test at three points in time—prior to the original instruction (1977), eight weeks after the instruction (1977), and as seniors in 1980.

#### QUANTITATIVE VERSUS QUALITATIVE EVALUATION

What follows is largely a quantitative analysis that utilizes statistical methods to investigate the subject of evaluation. Werking, in his 1980 article, is critical of such an approach for determining "proof" of effectiveness in the evaluation of library-use instruction.<sup>8</sup> Without denying the value of Werking's observations, the authors nevertheless believe they are justified on several sound grounds in using a quantitative ap-

proach. While *qualitative* evaluation is certainly legitimate in many evaluation contexts, *quantitative* evaluation is no less praiseworthy.

Quantitative evaluation has come under severe criticism, in part as an outgrowth of the results of the Westinghouse Learning Corporation's evaluation of the Head Start program.<sup>9</sup> Westinghouse's evaluators found, through using largely quantitative methods, that the effects of Head Start tended to fade when the children returned to poverty homes and ghetto schools, and this evaluation of a program—popular both in Congress and urban communities—met with sharp criticism, particularly with respect to the methodology used. The result has been that many educational-program evaluators now look to alternate methodologies, to techniques such as the qualitative assessments used in anthropology and sociology.<sup>10</sup> At least one observer has suggested that had the Westinghouse study found positive effects for Head Start, there would have been few questions raised about the adequacy of the quantitative methodology.<sup>11</sup>

No belittlement of the positive dimensions that qualitative methodology has brought to evaluation is intended; little is to be gained by a time-consuming and unproductive debate over qualitative versus quantitative methodology in the evaluation of library-use instruction. Reichardt and Cook, in their carefully reasoned examination of both methods, concluded that there was little reason to choose between them. They recommended that the researcher freely choose a mix of attributes from both types of methodological approaches so as to best fit the demands of the problems at hand.<sup>12</sup> In their view, the most telling and fundamental distinction between the two types of evaluative approaches lies along a continuum ranging from verification on one end to discovery on the other. According to Reichardt and Cook, quantitative methods have been developed most directly for the task of verifying or confirming established theories, while to a rather large extent qualitative methods have been developed primarily for the task of discovering or generating theories.<sup>13</sup>

As part of the overall evaluation of the library-use instruction program at DePauw University, both quantitative and qualitative

methods were used. Jerry Bakker, professor of chemistry at Earlham College and formerly the teaching-learning consultant at that school, well known for its library-use instruction program, conducted the qualitative part of the evaluation. The results of his evaluation, however useful, addressed primarily local concerns and are not included in any detail in this article.

For this public assessment of the impact of DePauw's library-use instruction program, quantitative analytical approach has a decided advantage. By employing statistics in the analysis of the effects of instruction and other factors upon library-use skills, we can communicate a good deal of information beyond our immediate setting. As Mueller has argued, "There is a continuity between common sense, which informally makes rough quantitative judgments, and statistics, which is not only a more formal and precise version of such knowledge, but also of more extended scope."<sup>14</sup> More specifically, while many in academic librarianship intuitively feel that library-use instruction is of considerable value in increasing library-use skills, quantitative measures can add precision and scope to such arguments. If one is particularly interested in sorting out the influence of other factors—such as student intellectual capacity, academic diligence or major field of study—statistical techniques can be indispensable in determining the direct effect of library-use instruction.

#### A THREE-YEAR ASSESSMENT

In an earlier report on the library-use instruction evaluation program at DePauw University, it was shown that an important amount of short-term gain in library-use skills was associated with that school's library-use instruction. In comparisons contrasting instructed freshmen with both senior students of the 1977 class and noninstructed freshmen (as a control group), those students who were exposed to library-use instruction showed evidence of the positive effects of that instruction.<sup>15</sup> Although these results were important to note and document, they represent only the first step in understanding the possible effects of library-use instruction. More important than the question of short-term gain in skills, of course, is the question of the *lasting effects* of instruction. Moreover, can

we associate higher levels of individual skills in library-use with higher degrees of exposure to library-use instruction? Similarly, over the long run, are factors other than library-use instruction better predictors of the acquisition of library-use skills? In order to investigate these and related questions, the data collected in the original study was supplemented with additional follow-up library-use skills information collected in a survey of the 1980 senior class at DePauw University.

Taken together, the survey data collected at two points in time in 1977 among freshmen and the senior class, and the data collected among the seniors of the 1980 class, provide the basis for two kinds of analyses of long-term skills-acquisition effects of library-use instruction. First, such data allow the comparison of *aggregate* levels of skills possession among various groups of interest (e.g., 1977 seniors versus 1980 seniors, those in the 1980 senior class who received library-use instruction versus those who did not, etc.). Secondly, the existence of three measures of library-use skills taken at three points in time for a substantial group of 1980 seniors—constituting a panel study—allows the verification of hypotheses suggested by aggregate patterns of comparison at the *individual* level of analysis.<sup>16</sup>

In the area of aggregate comparison, perhaps the most basic question is that of overall effects: that is, did the library-use instruction given to some students in the 1977 freshman class result in raising the overall level of library-use skills of that class? If library-use instruction given to 1977 freshmen did result in the improvement of the aggregate level of skill possession of students in that class, it should be possible to show that the skill levels of 1980 seniors (the 1977 freshmen) are superior to those of 1977 seniors. Table 1 reports the results of such a comparison.

Table 1 reveals findings that fall in the predicted pattern. While the relatively small number of 1977 seniors, the differential effects of selectivity in return rates in the 1977 and 1980 surveys of senior students, and the disproportionality of cases in the three major areas of study make the use of inferential statistics inappropriate, it is informative to note that the direction of differences observed coincides with predicted differences, and that the areas where most use is made of library

TABLE 1  
COMPARISON OF LIBRARY-USE SKILLS  
AMONG 1977 AND 1980 SENIORS

Major Area of Study	Mean Scores*	
	1977 Seniors	1980 Seniors
	X no.	X no.
Humanities	14.62 (29)	15.80 (74)
Social science	14.91 (47)	16.34 (92)
Natural science	15.03 (15)	15.08 (64)

\*Mean scores on identical, twenty-item skills test by major area of study.

resources—the humanities and social sciences—are precisely those where the greatest differences are observed.

Any such aggregate comparisons are subject, of course, to the criticism that factors *other than* library-use instruction account for the observed effects. Perhaps other campus-wide influences or national student trends intervened between 1977 and 1980 to cause the 1980 seniors of DePauw University to have higher library-use skills than their 1977 predecessors—irrespective of any contact with library-use instruction. Similarly, it is possible that the 1977 and 1980 senior classes differ with respect to intellectual capacity and/or academic diligence, hence any difference in library skills scores in the aggregate are the result of such background differences rather than selective exposure to library-use instruction. In order to determine whether exposure to library-use instruction has the predicted effect upon library-use skills, it is possible to analyze the findings of the 1980 senior survey to discover if (1) the degree of exposure to library-use instruction is directly associated with level of library-use skills possession; and (2) the association between library-use instruction and skills possession is stronger than that between skills possession and other relevant dimensions of difference among students—such as intellectual capacity (as measured by the verbal portion of the Scholastic Aptitude Test) and academic diligence (as determined by grade point average). Table 2 sets forth the findings of the 1980 senior survey with respect to these two dimensions of comparison.

The results reported in table 2 once more indicate the presence of a significant effect upon library-use skills of library-use instruction. The use of two measures of degree of exposure to library-use instruction to esti-



TABLE 2  
LIBRARY USE SKILLS, LIBRARY-USE INSTRUCTION, AND ACADEMIC BACKGROUND CHARACTERISTICS:  
A COMPARISON OF DEGREE OF ASSOCIATION AMONG 1980 SENIORS (GAMMA)\*

Measures of Exposure to Library-Use Instruction				Measures of Academic Background			
<i>Number of Courses Taken at Upper Division Level Where Library Instruction Was Given</i>				<i>Scholastic Aptitude Test - Verbal</i>			
Skill Test Score†	None	One	Two or More	Grade Point Average	≤ 440	450-530	≥ 540
Low	77	18	1	Low	32	33	21
Medium	33	26	7	Medium	28	15	17
High	18	32	18	High	11	16	32
	gamma = .658				gamma = .279		
<i>Total Number of Courses in Which Library-Use Instruction Was Encountered (Freshman Year and Upper-Division Courses)</i>				<i>Grade Point Average</i>			
Skill Test Score	No. Courses	Freshman Only	Freshman and Upper-Division	Grade Point Average	≤ 2.7	2.8-3.2	≥ 3.3
Low	24	61	11	Low	44	27	24
Medium	4	33	29	Medium	25	26	15
High	6	17	45	High	13	22	31
	gamma = .624				gamma = .310		

\*Gamma is an ordinal measure of statistical association measuring one-way association. It utilizes information about one variable to tell something about a second variable. The higher the gamma score the stronger the association between two variables. See Michael Malec, *Essential Statistics for Social Research* (Philadelphia: Lippincott, 1977), p.137-46.

†Scores on the skills tests have been trichotomized into low (15 or less), medium (16 or 17), and high (18 or more) categories.

mate the effects of differential experience with library-use instruction results in virtually identical findings with respect to the predicted effect of library-use instruction. Whether one considers the total number of courses taken in which library-use instruction was provided, or whether one focuses only upon upper division courses wherein special bibliographical instruction by a librarian was part of the course of instruction, it is clear that degree of exposure to instruction is positively associated with possession of library-use skills. When a comparison is made of the degree of association (the gamma coefficients) obtained between instruction and skill possession and the background characteristics (SAT verbal and GPA) and skill possession, it is clear that library-use instruction is much more highly correlated with skill possession than either inherent intellectual ability or academic diligence.

It is possible, of course, that the relationship between exposure to library-use instruction and these background factors is biasing the observed results; that is, it could be that the likelihood of taking additional course work in areas where library-use instruction is likely to occur is correlated with intellectual capacity and/or academic diligence, hence indicating a spuriously high association between library-use instruction and possession of library-use skills. In order to check against this possibility, it is necessary to employ multiple regression analysis, a statistical process wherein the simultaneous consideration of instructional exposure and background factors can be accomplished and results can be obtained that indicate the *relative importance* of each factor in the determination of variation in library-use skills possession.<sup>17</sup>

Table 3 reports the results of a multiple regression analysis that employs SAT verbal test scores, grade point average, number of upper division courses taken wherein library-use instruction occurred, and total number of library-use instruction courses experienced to predict library-use skill scores among 1980 seniors.

The results displayed in table 3 indicate clearly that experience with library-use instruction is the most important source of variation in library-use skills possession. In terms of relative effects, the two indicators of exposure to library-use instruction rank highest

TABLE 3  
RESULTS OF MULTIPLE REGRESSION ANALYSIS\*

Results†		
Multiple R	.623	
R Square	.389	
Standardized Regression Coefficients‡		
	Beta	Statistical Significance
Total number of courses	.367	.001
Number of upper-division courses	.262	.001
SAT verbal	.162	.05
GPA	.159	.05

\*Relative effects upon level of library-use skills for 1980 seniors produced by exposure to instruction and academic background.

†Dependent variable = Library-use skill score

‡Independent variables = SAT verb 1, GPA, number of upper-division courses wherein library-use instruction occurred, and total number of courses since freshman year wherein library-use instruction occurred.

§Relative predictive power of independent variables.

and next highest in the ordering of standardized regression coefficients (indicators of degree of impact upon the dependent variable of one predictor after the intervening contributory effects of all other predictors have been controlled) for the four variables entered into the regression analysis. SAT verbal scores and grade point average do not rival the effects of total number of courses taken in which library-use instruction is obtained as a predictor of level of library-use skills possession.

The analyses developed up to this point indicate very clearly the possibility that important effects are associated with library-use instruction. However, the possibility persists that an *ecological fallacy* may be associated with the exclusive use of aggregate data and collective comparisons. That is to say, the aggregate association between instruction and skills possession may not derive from individual effects.<sup>18</sup> In comparing various subgroups (e.g., highly exposed versus freshman-instructed only, high grade point average versus modest grades, etc.) to determine the degree of association with skills possession demonstrated by one or another factor, it is always possible that the groups being compared are dissimilar with respect to one or more important factors. One way to remedy this problem in the study of factors associated with change due to instructional effects is to study *the same persons* (as opposed to differ-

ent groups of persons) over time. This panel study technique is often employed to determine both the direction of effects due to instruction and to assess the absolute amount of change occurring where it is possible that students might both gain and lose skills or information at varying rates.

Not only does the use of a panel study technique allow one to check for the hidden effects of intervening factors, but it also allows the researcher to distinguish between short-term and long-term gains in skills or information. By taking measurements of skills possessed before instruction, shortly after the conclusion of instruction (eight weeks), and a considerable time after instruction (three years), it is possible to identify both short-term and long-term effects of instruction, and it is possible to determine what factors are associated with both short-term and long-term changes in skills possession levels. Table 4 reports the results of such an analysis. It includes a listing of measures of association (Pearson correlation coefficients) for the four major factors investigated above—two measures of exposure to library instruction, a measure of intellectual capacity, and a measure of academic diligence.

Table 4 adds further evidence to the argument that library-use instruction is an effective means of enhancing library-use skills. In the area of academic background factors it can be seen that there is a modest degree of association between both grade point average and SAT verbal test scores and short-term changes in library-use skills, but that neither

factor is associated with long-term library-use skill scores to a statistically significant degree. In contrast, long-term changes in library-use skills are highly associated with both measures of exposure to library-use instruction. These findings indicate that neither intellectual capacity per se nor diligence in the pursuit of good grades will produce a degree of learning of library-use skills that can rival the amount of skills acquisition that is provided in library-use instruction. It is important to note that library-use instruction can be shown to have effects superior to those of academic background in both aggregate comparisons and the panel study setting, a fact that adds greatly to the contention that library-use instruction has firm value and lasting effects.

#### USE OF EVALUATION RESULTS: A CAUTIONARY NOTE

Werking has expressed particular concern with respect to the use of quantitative evaluation results as "proof" of a library-use instruction program's success.<sup>19</sup> This is certainly a legitimate concern, and the authors want to insert a cautionary note into this article for those planning to conduct evaluations in order to gain support for their programs. In their previous article on the DePauw University library instruction program, the authors reported use of the results of their evaluation to successfully gain administrative support for a grant proposal to continue the library-use instruction program.<sup>20</sup> Such use of evalu-

TABLE 4  
FACTORS ASSOCIATED WITH SHORT-TERM  
AND LONG-TERM CHANGES IN LIBRARY-USE SKILLS\*

	Short-Term Changes			Long-Term Changes		
	Corr. Coeff.	No. Cases	Stat. Sig.	Corr. Coeff.	No. Cases	Stat. † Sig.
<i>Measures of Academic Background</i>						
SAT, verbal	.19	77	.05	.16	77	Not sig.
GPA	.19	82	.04	.08	82	Not sig.
<i>Measures of Exposure to Library-Use Instruction</i>						
Number of upper-division courses	.11	82	Not sig.	.38	82	.0002
Total exposure to instruction, freshmen through graduation	.30	82	.003	.41	82	.0001

\*Panel study results of correlations between changes in skill level, exposure to library-use instruction, and academic background (Pearson Correlation Coefficients).

†Result listed as not statistically significant if *p* is greater than .05.

Note: Short-term and long-term change scores are calculated on the basis of the difference (positive or negative) between the preinstruction skills score and the first and second skills tests for each respondent.

ation in decision making is neither automatic nor common.

Some evaluators contend that the main purpose of evaluation is simply to improve learning and instruction, and that all other uses are secondary or supplementary to this purpose.<sup>21</sup> However, if the ultimate purpose of evaluation is to contribute to decision making pertaining to the improvement of the effectiveness of library programs, the implementation of evaluation results is a critical consideration.<sup>22</sup> Carol Weiss has noted that while careful and unbiased evaluations should ideally improve decision making in a rather automatic fashion, evaluation is always a rational enterprise that takes place in a political context. The evaluator who fails to recognize political considerations "is in for a series of shocks and frustrations."<sup>23</sup>

Additionally, Werking, in citing the example of the abandonment of a teaching method *not* because it was ineffective, but because it was believed to be nonessential, points to an ever-present problem in using quantitative evaluation techniques to gain support for a library-use instruction program.<sup>24</sup> Library-use instruction programs are sometimes considered to be "amenities" by decision makers such as college administrators, classroom instructors, and library directors. According to Benjamin Bloom, however sophisticated and

elegant quantitative evaluations might be, they "are likely to have little effect if they are considered to be measuring trivial things which are not regarded as important by the students, teachers, patrons, and others."<sup>25</sup>

A certain amount of groundwork is necessary before any type of formal evaluation of a program is attempted. As noted by Howard Davis and Susan Salasin, newcomers to evaluation too often take effective use of evaluation for granted, with the result that evaluation results often end up being ignored.<sup>26</sup> Librarians interested in evaluating their library-use instruction programs would do well to recall the wise observation of Suchman: "Both the demand for and the type of acceptable 'proof' (of program effectiveness) will depend upon the nature of the relationship between the social institution and the public. In general, a balance will be struck between faith and fact."<sup>27</sup> Any librarian seriously considering the formal evaluation of his or her instructional program would be well advised to respect the limitations of both methodology and practical politics involved,<sup>28</sup> and take heart that in time well-conceived and rigorously conducted evaluations of program effects will have an increasingly important role in the management of college and university instructional resources.

#### REFERENCES

1. Charles Dudley Warner first wrote in an editorial in the *Hartford Courant* in 1890 that "everybody talks about the weather, but nobody does anything about it." Even back then it was so often attributed to Mark Twain that Charles Hopkins Clark, editor of the *Courant* wrote: "I guess it's no use; they still believe Mark Twain said it, despite all my assurances that it was Warner."
2. Richard Hume Werking, "Evaluating Bibliographic Education: A Review and Critique," *Library Trends* 29:153-72 (Summer 1980).
3. Larry Hardesty, Nicholas P. Lovrich, Jr., and James Mannon, "Evaluating Library-Use Instruction," *College & Research Libraries* 40:309-17 (July 1979).
4. A third goal of the earlier study was to document the effects of library-use instruction on short-term changes in attitudes toward the library. The earlier study was largely unsuccessful in addressing questions relating to attitudinal changes, and the authors concluded "that

either attitudes about libraries are far more difficult to influence or measure than are library-use skills, or the particular program at DePauw University was less effective in influencing attitudes than it was in influencing library-use skills." The same attitude-measure instrument was used in the long-term study reported in this article and the results and conclusions are the same as in the earlier study. For an interesting discussion of a long-term study of student attitudes toward library-use instruction, the reader is referred to Roland Person, "Long-Term Evaluation of Bibliographic Instruction: Lasting Encouragement," *College & Research Libraries* 42:19-25 (Jan. 1981). However laudable Person's efforts and support of library-use instruction, his conclusions must be accepted with some reservations. Most critical is his reliance on a low return rate (26 percent) from students, most of whom had originally volunteered to enroll in a library-use instruction course. With such self-



- selection and low return rate the reader does not know what a majority of the students who took the course thought, let alone what a majority of the student body might think of such a course. There are other concerns with the article, but this criticism is not intended to discourage Person and others from conducting such studies. It is intended to encourage the "better research" in the area of library-use instruction that both Person and the authors agree is needed.
5. Werking, "Evaluating Bibliographic Education," p.161.
  6. Julian Simon, *Basic Research Methods in Social Science* (2d ed.; New York: Random, 1978), p.185.
  7. C. A. Moser, *Survey Methods in Social Investigation* (London: Heineman, 1965), p.111.
  8. Werking, "Evaluating Bibliographic Education," p.166-67.
  9. *The Impact of Head Start: An Evaluation of the Effects of Head Start on Children's Cognitive and Affective Development* (Bladensburg, Md.: Westinghouse Learning Corporation, 1969).
  10. Francis A. J. Ianni and Margaret Terry Orr, "Toward a Rapprochement of Quantitative and Qualitative Methodologies," in Thomas D. Cook and Charles S. Reichardt, eds., *Qualitative and Quantitative Methods in Evaluation Research*, Sage Research Progress Series in Evaluation, V.1 (Beverly Hills, Calif.: Sage Publications, 1979), p.88.
  11. John W. Evans, "Head Start: Comments on the Criticism," in Francis G. Caro, ed., *Readings in Evaluation Research* (New York: Russell Sage, 1971), p.401.
  12. Charles S. Reichardt and Thomas D. Cook, "Beyond Qualitative versus Quantitative Methods," *Qualitative and Quantitative Methods in Evaluation Research*, p.19.
  13. *Ibid.*, p.17.
  14. John Mueller and others, *Statistical Reasoning in Sociology* (2d ed.; Boston: Houghton, 1970), p.3.
  15. Hardesty, Lovrich, and Mannon, "Evaluating Library-Use Instruction," p.313.
  16. The importance of individual-level verification of aggregate comparisons cannot be over-emphasized, particularly in the analysis of change along a temporal dimension. As a simple illustration one might consider the nightly newscast of electoral attitude changes occurring in the course of an election. Say a report of candidate preference in September shows candidate X with a 50 percent approval rating, and another survey in November shows the same result—50 percent approval. Could we conclude that no voters changed their preferences during the campaign? Could we assume that all voters changed their minds—i.e., all of those who favored X now favor another candidate and all of those who favored other candidates now favor X? Sadly enough, *either* of these explanations could be true—or neither could be true; perhaps an equal proportion (large or small?) of voters moved both toward and away from a preference for X. As is quite evident, without *individual* level comparisons wherein the changes in preferences of individuals can be observed varying over time, none of these hypothetical explanations of attitude change can be accepted.
  17. Hubert M. Blalock, Jr., *Social Statistics* (2d ed.): New York: McGraw-Hill, 1972), p.429-72.
  18. W. S. Robinson, "Ecological Correlations and the Behavior of Individuals," *American Sociological Review* 15:351-57 (June 1950).
  19. Werking, "Evaluating Bibliographic Education," p.166-67.
  20. Hardesty, Lovrich, and Mannon, "Evaluating Library-Use Instruction," p.316.
  21. Norman E. Gronlund, *Measurement and Evaluation in Teaching* (3d ed.; New York: Macmillan, 1976), p.13.
  22. Francis G. Caro, "Evaluation Research: An Overview," in his *Readings in Evaluation Research* (New York: Russell Sage, 1971), p.12.
  23. Carol H. Weiss, "Evaluation Research in the Political Context," in Elmer L. Struening and Marcia Guttentag, eds., *Handbook of Evaluation Research*, V.1 (Beverly Hills, Calif.: Sage Publications, 1975), p.13.
  24. Werking, "Evaluating Bibliographic Education," p.167.
  25. Benjamin S. Bloom, "Some Theoretical Issues Relating to Educational Evaluation," in Ralph W. Tyler, ed., *Educational Evaluation: New Roles, New Means*, Sixty-eight Yearbook of the National Society for the Study of Education, Part II (Chicago: Univ. of Chicago Pr., 1969), p.44.
  26. Howard R. Davis and Susan E. Salasin, "The Utilization of Evaluation," *Handbook of Evaluation Research*, p.624.
  27. Edward A. Suchman, "Evaluation Research," in Elmer L. Struening and Marcia Guttentag, eds., *Handbook of Evaluation Research*, V.1 (Beverly Hills, Calif.: Russell Sage, 1975), p.624.
  28. Larry Hardesty, "Instruction Development in Library Use Education," in Carolyn A. Kirkendall, ed., *Improving Library Instruction* (Ann Arbor: Pierian Pr., 1979), p.11-36.