Review of Educational Research
Fall 1986, Vol. 56, No. 3, Pp. 331–364

Current Empirical Research on Evaluation Utilization

J. Bradley Cousins and Kenneth A. Leithwood
The Ontario Institute for Studies in Education

ABSTRACT. This paper reviews empirical research conducted during the past 15 years on the use of evaluation results. Sixty-five studies in education, mental health, and social services are described in terms of their methodological characteristics, their orientation toward dependent and independent variables, and the relationships between such variables. A conceptual framework is developed that lists 12 factors that influence use; six of these factors are associated with characteristics of evaluation implementation and six with characteristics of decision or policy setting. The factors are discussed in terms of their influence on evaluation utilization, and their relative influence on various types of use is compared. The paper concludes with a statement about implications for research and practice.

Ordinary knowledge provides the basis for decision and action in most organizations. Such knowledge, derived from practical experience, is usually widely shared, sensitive to context, and comprehensive. By contrast, knowledge derived from social science methods tends to be context independent and, of necessity, to be selective rather than comprehensive. At best, such knowledge supplements ordinary knowledge.

The above argument (developed by Lindblom & Cohen, 1979) is a compelling one. On the one hand, it justifies the generation of knowledge using social science methods, and on the other, it imposes restrictions on the use of such knowledge. However, evaluators and others with a stake in social science have generally failed to recognize this argument, generating unrealistic expectations about the value of knowledge derived from social science methods and minimizing the application of such knowledge. Our purpose in this review is to assess what factors influence the use of evaluation data. Four questions guided our inquiry:

What are the methodological characteristics of empirical studies for investigating evaluation use and its determinants?

How have dependent variables been operationalized in these studies?

What orientations toward independent variables have been adopted?

What factors have been shown to affect the nature of use of evaluation results?

This research is sponsored in part by the Ontario Ministry of Colleges and Universities and the Ontario Ministry of Education. The opinions expressed here, however, are solely the responsibility of the authors. The editorial assistance of Hugh Oliver and Mary Stager is gratefully acknowledged. Previous versions of this paper were presented at the 1985 annual meeting of the American Educational Research Association and the 1985 joint meeting of the Evaluation Network, Evaluation Research Society, and Canadian Evaluation Society.

Framework

When social science methods are used to answer questions posed by decisionmakers, the outcomes are typically called *evaluations* or *evaluation studies.* In practice, such studies serve a broad array of decisionmaker needs, only some of which can be anticipated. For the purposes of the present review, we assumed that the role of the evaluator was to carry out the evaluation study and to communicate the results to the decisionmaker(s).

We defined *evaluation results* broadly as any information associated with the outcome of the evaluation, for example, data, interpretations, recommendations; such information could be communicated at the completion of the evaluation or as the evaluation was proceeding. The *target of evaluation* was assumed to be one or more of several possibilities, for example, curriculum, program, project, school, course, personnel, or student. Similarly, *decisionmakers* were assumed to be affiliated with one or more of several services, for example, education, health, mental health, or social services.

There are two conventional definitions of evaluation *use* or *utilization*, the dependent variable in the present review: (a) support for discrete decisions and (b) the education of decisionmakers. More recently, an even more basic conception of evaluation use has been described: that the mere psychological processing of evaluation results constitutes use, without necessarily informing decisions, dictating actions, or changing thinking. Each of these definitions was used in the present review.

Factors influencing decisionmakers' use of evaluation results were the independent variables in our study. Many taxonomies of such factors are available, claiming to provide useful frameworks for guiding evaluation practice, research on utilization, or both (see Alkin, Daillak, & White, 1979; Leithwood, Wilson, & Marshall, 1981; Leviton & Hughes, 1981; Weiss, 1972). Our subsequent review suggests that any given factor proposed or shown to influence evaluation use can be viewed as belonging to one of two major categories: (a) characteristics of the *evaluation implementation* and (b) characteristics of the *decision or policy setting.* These hypothetical higher order factors are assumed to be correlated and to influence in combination, how and to what extent use will occur. This conceptualization is similar to that proposed by Weiss.

Method

For two reasons, our literature search was restricted to the 1971–1985 period. First, research on evaluation utilization is comparatively new and it is unlikely that many relevant empirical studies were reported before 1970. Second, our interest was in the current state of knowledge. The 65 empirical studies eventually identified for review[1] exhausted, as far as we know, the empirical literature about evaluation

[1] There were 58 separate publications; two of the documents (Brown, Newman, & Rivers, 1980; Newman, Brown, & Littman, 1979) each contained three statistically independent studies. Newman, Brown, and Rivers (1983) reported four studies. Some of their subjects participated in two of these studies. Also, it appears that the analyses reported in two articles by Kennedy (1983, 1984) were drawn from the same data base, although no specific mention was made of this by Kennedy (1984); however, we judged the emphasis on analyses to be sufficiently distinct in the two studies to warrant the inclusion of both in our sample.

use for this period, although there are no doubt fugitive papers that remain undetected.[2] *Educational Resources Information Center* (ERIC), *Psychological Abstracts, Sociological Abstracts,*[3] *Dissertation Abstracts International*, and bibliographic follow up were sources we explored for identifying the studies. In addition, we searched a bibliography compiled by King, Thompson, and Pechman (1980).

Table I lists the studies, along with various descriptive characteristics. Because they derive from several settings (education, mental health, and social services), our findings may be generalized to a range of such settings.

Some studies in the sample touched on the use of tangential evaluation information such as social science data (Caplan, 1976; Florio, Behrman, & Goltz, 1979) and data-based decision support systems (Pauley & Cohen, 1984; Williams & Bank, 1984). Kennedy (1983, 1984) studied the use of "evidence" that sometimes exceeded the bounds of evaluation results. Each of these studies, however, provided an investigation of evaluation use sufficient to warrant inclusion in the sample.

Results

Methodological Characteristics

The 65 studies employed retrospective, longitudinal, and simulation research designs (see Conner, 1981 for similar distinctions). *Retrospective* studies focused on previous evaluations, relied for data mainly on the memories of decisionmakers, sponsors, and/or evaluators, and sometimes resorted to anecdotal accounts (e.g., Brickell, 1976; Carter, 1971; McGowan, 1976; Osterlind, 1979). Such studies had

---

[2] After submitting this review for publication, we learned of three new studies published in *Educational Evaluation and Policy Analysis*, 1985, *7* (4). These were: W. K. Ripley, "Medium of presentation: Does it make a difference in the reception of evaluation information?" (pp. 417–426); McColskey, W. H., Altshield, J. W., and Lawton, R. W., "Predictors of principals' reliance on formal and informal sources of information," (pp. 427–436); and Brown, R. D., Newman, D. L., and Rivers, L. S., "An exploratory study of contextual factors as influences on school board evaluation information needs for decisionmaking," (pp. 437–445). Ripley examined the effects of a factor we discuss as "Communication Quality" on the potential for use (defined in a manner we subsequently discuss as problematic in many simulation studies). He compared the effects of written reports, audio cassette reports, and video tape reports on three types of "use." McColskey et al. examined the effects of a factor we discuss as "Personal Characteristics" on principals' use of formal and informal sources of information for discrete decisionmaking. Those characteristics examined were principals' leadership orientation, open-mindedness, perceptions on autonomy, and prior training in social science research methods. Each characteristic was significantly related to the extent of use made of both formal and informal sources of information. Brown et al. inquired about the effects on several variables antecedent to utilization, factors we label "Decision Characteristics" and "Political Climate." Several of the dependent variables in this study were typically considered independent variables in other studies of use (e.g., need for more time, confidence in personal opinion). Findings argue for the importance of decision characteristics and political climate in explaining the nature and extent of evaluation use.

[3] We commissioned a computerized search for items published from 1971 to 1984 in these sources. The following keywords were used: evaluation utilization, data use, test use, decisionmaking, and knowledge utilization. A preliminary manual search in recent issues of these periodicals was helpful in identifying appropriate keywords and uncovering articles missed by the subsequent computer search.

TABLE I

Evaluation Utilization Empirical Studies by Descriptive Characteristics

Study	*N*	Sample	Design	Dependent Variables	Independent Variables	Theory	Instruments
1. Airasian, Kellaghan, Madaus, & Pedulla (1977)	47	Second grade teachers from a nation-wide sample in Ireland	Longitudinal: experiment	Magnitude and direction of student ranking change	Standardized test administration	None mentioned	Rating forms
2. Alkin, Daillak, & White (1979)	5	Program evaluations: educators, directors, decision makers; documents	Retrospective: multiple case study	Mainstream and alternative uses	Evaluation approach, credibility, org. factors, etc.	Theory of evaluation use	Interview schedule, questionnaire; document analysis
3. Alkin, Koescoff, Fitz-Gibbon, & Seligman (1974)	42	Title VII projects: federal monitors, project directors, reports	Retrospective: field study	Use rating, funding level, composite indexes, etc.	Evaluation and auditor report, and project factors	Theory of evaluation use	Evaluator, auditor data sheets, and questionnaire
4. Alkin & Stecher (1983)	66	Principals, prog. coordinators, resource personnel in 22 elem. schools	Retrospective: multiple case study	Strength and weakness identification, support for school decisions	Compet'g info., decision phase, evaluation and decision type, org. role, etc.	None mentioned	Topic-centered interview format
5. Barnette & Thompson (1978)	208	Secondary school teachers in instructional development chairperson role	Retrospective: survey	Ratings and probes on improvement in instruction	Evaluation and teacher characteristics	None mentioned	Questionnaire
6. Becker, Kirkhart, & Doss (1982)	70	Title I teachers and teacher aides	Simulation: experiment	Ratings of report, mean knowledge score	Report style	Communication theory	Questionnaire
7. Bigelow & Ciarlo (1976)	30	Community mental health team leaders, program and special program directors in a CMHC setting	Longitudinal: field study	Report interpretations, actions planned and taken, decision criteria	Competing influences, prog. objectives, info. needs, findings	Management process model	Interview schedules, telephone questionnaires
8. Braskamp, Brown, & Newman (1978)	111	Administrators, superint's, supervisors, curriculum directors, teachers	Simulation: experiment	Agreement and report ratings, cost est. of the evaluation	Message source and content, audience type	Communication theory	Rating forms

9. Brickell (1976)	10	Evaluation projects: associated personnel	Retrospective: multiple case study	10 episodes of evaluation use and method decisions	Project specific political factors	None mentioned	Participant observation
10. Brown, Braskamp, & Newman (1978)	95	Teachers and public school administrators	Simulation: experiment	Agreement, evaluator, report ratings	Message content	Communication theory	Rating forms
11. Brown & Newman (1982)	64	Teachers and public school administrators	Simulation: experiment	Agreement, info. suffic'y, evaluator ratings	Message source and content	Communication theory	Rating forms
12. Brown, Newman, & Rivers (1980) Study I	88	Education, business, social service professionals	Simulation: experiment	Budget needs assessment inventory	Message content	Communication theory	Rating forms
13. Brown, Newman, & Rivers (1980) Study II	84	Education, business, social service professionals	Simulation: experiment	Agreement, evaluator, info. satis'n ratings	Message content	Communication theory	Rating forms
14. Brown, Newman, & Rivers (1980) Study III	88	Education, business, social service professionals	Simulation: experiment	Agreement, evaluator, and info. satis'n ratings	Perceived need for evaluation, message content	Communication theory	Rating forms
15. Brown, Newman, & Rivers (1985)	178	School board members from nation-wide sample	Simulation: experiment	Information need, support for decision	Decision importance, setting, conflict, superin't opinion	Commu'n theory, conflict and social process models	Rating forms
16. Caplan (1976)	204	Political appointees and high-level civil servants	Retrospective: survey	Number instances of use, info. proc. style	Purpose, info. type, impact area, parochialism, etc.	None mentioned	Interview coding form
17. Carter (1971)	4	Evaluations with pending decisions and data available	Retrospective: multiple case study	4 episodes of evaluation use	Negative evaluation results	Power equalization theory	Naturalistic observation
18. Daillak (1983)	3	Evaluators in moderate large urban school district	Longitudinal: multiple case study	Programmatic decisions, impl. of rec's, staff devel't	Org. location of evaluator, org. resistance, etc.	None mentioned	Participant observation, informal interviews
19. David (1978)	30	Title I school districts: directors, project admin'rs evaluators, principals, teachers, superintend's	Retrospective: survey	Probes on prog. judgment and planning, feedback & select'n	Prog. char's, decision proc., eval. chars, & staff resist'ce	None mentioned	Interview schedule

TABLE I—*Continued*

Study	*N*	Sample	Design	Dependent Variables	Independent Variables	Theory	Instruments
20. Dawson & D'Amico (1985)	1	Evaluation of elementary, and secondary school effectiveness programs	Longitudinal: case study	Program improvement, learning re: prog. needs	Involvement of program staff in evaluations	None mentioned	Participant observation
21. Dickey (1980)	47	Title IV-C funded project directors, evaluation reports	Retrospective: survey	Multi-level impact rating	Decisionmaker and eval. characteris's	Two communities theory	Questionnaire, interview schedules
22. Dickman (1981)	156	Social researchers: entrepreneurial, social program, academic	Retrospective: field study	Index based on perceived prog. change	Evaluation method soph., org. location, prog. involv't	None mentioned	Questionnaire
23. Fetler (1982)	(4,500)[a]	Elementary school principals	Retrospective: survey	Checklist of 18 types of use	School achievement level	Evaluation use models	Use checklist and California Assess't Prog. (CAP) inst's
24. Floria, Behrmann, & Goltz (1979)	26	U.S. Congressional staff members (education)	Retrospective: survey	Reported use for policy delib'n, decision, importance rankings	Compet'g info., communication, timeliness, relevance, findings, etc.	None mentioned	Questionnaire and interview schedule
25. Glasman (1984)	273	Elementary school principals	Retrospective: field study	Usefulness and beliefs about use ratings	Purpose of use, principal attitudes and effectiveness	Principal's role value stance	Questionnaire and rating forms
26. Granville (1977)	157	Elementary and secondary school principals	Simulation: experiment	Single, discrete decision	Social, political pressure, data object'y	Evaluation and social theory	Rating forms
27. Goldberg (1978)	40	ESAA program administrators: assistant principals and teachers	Retrospective: field study	Helpfulness and quantity of info. ratings	Decision area, steps of decisionmaking	Stufflebeam model	Rating forms
28. Heldt, Braskamp, & Filbeck (1973)	36	Community college administrators	Simulation: experiment	Discrete decision rating	Constituency value position and prob. type	Decision theory	Rating forms
29. Herman & Dorr-Bremme (1983)	(917)[a]	Principals and teachers in 91 nation-wide school districts	Retrospective: survey	Test result usefulness ratings	Decision area, school level, and eval. info.	None mentioned	Questionnaire

30. Johnson (1980)	75	Social service decisionmakers	Retrospective: field study	Index based on conceptual and instrumental use	Higher order factors: organizational char's.	Havelock's research use framework	Interview schedule, questionnaire
31. Jordan (1977)	3,961	Superintendents and principals	Retrospective: survey	Reported benefit to schools, staff morale, community sup't.	Evaluation format	None mentioned	Telephone questionnaire
32. Kennedy (1983)	16	School districts: policy developers, principals, teachers	Retrospective: multiple case study	Reported info. proc'g, reten'n, and interpret'n	Working knowledge, info. needs	Information processing theory	Interview and observation schedules
33. Kennedy (1984)	16	School districts: Policy developers, principals, teachers	Retrospective: multiple case study	Reported use: indiv. concep'l, group, and direct applic'ns	Working and shared knowledge	Information processing theory	Interview and observation schedules
34. King & Pechman (1984)	1	Research and evaluation unit in education central office	Longitudinal: single case study	Inst'l and conceptual use and non-use of evaluation	Funding req'ts, personal gain organizational constraints, findings, etc.	Evaluation use model	Naturalistic observation
35. Leviton & Boruch (1983)	42	Evaluations: project directors, contractors	Retrospective: survey	Impact, reported decision and consideration	Type of eval. information	None mentioned	Interview schedules
36. Lorenzen & Braskamp (1978)	12	Community mental health center administrators	Simulation: experiment	Index based on usefulness ratings, decision outcome, budget sugges'ns	Decision type and information type	Communication theory	Questionnaire and interview schedules
37. Maher (1982)	4	Evaluations: program directors, school staff	Longitudinal: field study	Director sat'n, modif'n in operations, and eval. activ's	Evaluation strategy, evaluator location	Knowledge util'n, org. behav. theory	Interview schedules and on/off site log
38. McGowan (1976)	1	U.S. Office of Education EEP innovative program	Retrospective: single case study	Observed goal clarification, monitoring, funding support	Evaluation strategy and political pressure	None mentioned	Participant observation and informal interviews

TABLE I—*Continued*

Study	*N*	Sample	Design	Dependent Variables	Independent Variables	Theory	Instruments
39. Neigher (1979)	70	CMHC administrators and recognized authorities in mental health prog. eval.	Longitudinal: field study	Service prov'n, improvement and priorization of eval. activities	Evaluation requirements, target audience	None mentioned	Questionnaire
40. Newman, Brown, & Littman (1979) Study I	212	Education and business undergraduates	Simulation: experiment	Agreement, report, evaluator ratings	Message source, content, audience char's.	Communication theory	Rating forms
41. Study II	207	Business professionals and undergraduates	Simulation: experiment	Agreement, report, evaluator ratings	Message source, content, audience char's.	Communication theory	Rating forms
42. Study III	203	Education and business undergraduates	Simulation: experiment	Agreement, report, evaluator ratings	Message source, content, audience char's.	Communication theory	Rating forms
43. Newman, Brown, & Rivers (1983) Study I	87	Educational administration grad. students, workshop participants	Simulation: correlational study	Preferences for eval. rec's and evaluator role	Locus of control	Decision theory	Rating and ranking forms
44. Study II	72	Educational administration grad. students, workshop participants	Simulation: correlational study	Information usefulness ratings, eval. needs	Locus of control	Decision theory	Rating forms
45. Study III	30	Educational administration grad. students, workshop participants	Simulation: correlational study	Info. use ratings, eval. and info. needs, $ value of eval.	Locus of control	Decision theory	Rating forms
46. Study IV	8	Subsample from Study III	Simulation: observational study	Group decision dynamics	Locus of control	Decision theory	Bystander observation
47. Newman, Brown, Rivers, & Glock (1983)	283	Education and business undergraduates	Simulation: experiment	Information needs and values	Decision context	Dec'n and communication theory	Rating forms
48. Ory & Braskamp (1980)	71	University faculty members	Simulation: experiment	Report ratings and probes on use	Message content	Communication theory	Rating forms and questionnaire
49. Osterlind (1979)	1	Evaluation of politically volatile program	Retrospective: single case study	Observed staff reaction to evaluation	Eval. design selection, objectiv'y, staff feedback needs	None mentioned	Participant observation

50. Patton, Grimes, Guthrie, Brennen, French, & Blyth (1977)	60	Project officers, decision-makers and evaluators for each of 20 health prog. evaluations	Retrospective: field study	Reported initiative and change in decision priorities	Evaluation methods, personal, political factors	Organiza'l theory	Interview schedule
51. Pauley & Cohen (1984)	7	Community mental health center coordinators (plus case workers)	Longitudinal: field experiment	Reported satisfaction with MIS facil. system, data-based decisions, info. value, use, errors, productivity	Intervention of structured consultative MIS facil'n system	Data-based decision cycle	Rating and ranking forms, needs assessment inventory, interview schedule, activity log, client information forms
52. Rich (1979)	28	Congressional staff members with unemployment insurance expertise	Retrospective: field study	Serious consideration, support for policy	Control of research, object'y timeliness, etc.	None mentioned	Interview schedule
53. Rossman, Hober, & Ciarlo (1979)	7	Studies of impact: community mental health center	Retrospective: survey	Reported subjective utility, decision and program chg. measures, impact score	Dec. type, compet'g info., admin. level, pres. style, relevance	None mentioned	Interview schedule questionnaire, document analysis
54. Salmon-Cox (1981)	68	Teachers	Retrospective: survey	Probes on types of test use, beliefs about appropriate uses	Type of decision, other info., findings, test frequency	None mentioned	Interview schedules
55. Siegal & Tuckel (1985)	2	Evaluations of CMHC's: key actors	Retrospective: multiple case study	Serious consideration, understanding of operations, policy formulation	Findings, credibility, timing, resistance, communication	None mentioned	Interview schedule
56. Sproull & Zubrow (1981)	58	Educational administrators: superintendents, directors of curriculum and instruction	Retrospective: survey	Probes on test result use	Type of school syst., level of admin. comp. information	None mentioned	Interview schedules

TABLE I—*Continued*

Study	*N*	Sample	Design	Dependent Variables	Independent Variables	Theory	Instruments
57. Tash & Stahler (1982)	1	Community mental health center evaluation	Retrospective: single case study	Observations of enhanced communication, relevence, use of recommendations	Recommendations process strategy, staff involvement in evaluation	None mentioned	Naturalistic observation
58. Thompson, Brown, & Furgason (1981)	159	Teachers, public school administrators, business professionals	Simulation: experiment	Agreement, report, evaluator ratings	Message content, audience char's.	Communication theory	Rating forms
59. Van de Vall & Bolas (1982)	120	Social research projects: policy makers, social researchers, research report documents	Retrospective: survey	Index based on manifest, stage, certainty, latent impact	Organ'l decision structure, function of research	None mentioned	Interview schedules, coding forms
60. Webber (1984)	240	School trustees	Retrospective: survey	Usefulness ratings, probes on types of use	User exper'nce, demographic info.	None mentioned	Questionnaire, rating form, telephone interview schedule
61. Weeks (1979)	57	Evaluators	Retrospective: survey	Index based on knowledge of use	Org. location of evaluator, eval. methods, decision context	Organizat'l theory	Questionnaire
62. Wholey & White (1973)	[b]	Title I ESEA program evaluations (local, state, federal)	Retrospective: survey	Observed change, prog. effectiveness, planning, mgt., eval. development	Timeliness, relevence and findings of evaluation	None mentioned	Historical account
63. Williams & Bank (1984)	2	School districts: teachers, parents, principals, learning specialists	Retrospective: survey	Probes on usefulness of eval. info for spec'd decision, sat'n with info., process, use	Working knowledge, significance of dec'n school culture, perceived data utility	Info. proc. theory, organiz'l decision theory	Interview schedules
64. Windle, Majchrzak, & Flaherty (1979)	190	CMHC personnel	Longitudinal: field study	Program changes, type & emphasis of evaluation	CMHC develop'l stage, eval. activ'y, report style, eval. requirements	None mentioned	Questionnaire
65. Yeh (1980)	260	Kindergarten to sixth grade teachers	Retrospective: survey	Use for communication, instruc'l decisions, grade assign't.	Instructional practices, assessment, training, SES, etc.	None mentioned	Questionnaire, existing data

[a] Estimate based on information reported.
[b] No reported information regarding sample size.

limited reliability and precluded the manipulation of independent variables. *Longitudinal* studies examined the influence of data gathered prior to, during, and/or after evaluation implementation; although they sometimes involved the manipulation of independent variables, such studies failed to allow sufficient time to determine the impact of evaluation results. *Simulation* studies usually employed contrived evaluation reports to test, under highly controlled conditions, for effects on anticipated uses; although there was scope for manipulation of independent variables, contrived study conditions tended to lack important characteristics of actual decision settings. These three types of research design were used as a basis for organizing the review because they allowed for the identification of patterns associated with orientation toward dependent and independent variables.

Most studies were retrospective ($n = 35$) and were predominantly surveys, but included some single or multiple case studies. Sometimes longitudinal field studies ($n = 9$) involved quasi-experimentation (e.g., Pauley & Cohen, 1984). Simulation studies ($n = 21$), were almost exclusively experimental, generally employing random allocation of subjects to groups. Some of these studies used blocking variables to allocate subjects to groups (e.g., Newman, Brown, Rivers, & Glock, 1983); some used repeated measures designs (e.g., Braskamp, Brown, & Newman, 1978).

Study samples varied considerably in size (from 1 to 4,500; $Mdn = 60$) and in the unit of analysis (generally the individual respondent, but sometimes the evaluation project per se—Brickell, 1976; Carter, 1971; Leviton & Boruch, 1983; Rossman, Hober, & Ciarlo, 1979; Van de Vall & Bolas, 1982; or sometimes the domain of the evaluation—Alkin, Kosecoff, Fitz-Gibbon, & Seligman, 1974; Bigelow & Ciarlo, 1976; David, 1978; Kennedy, 1983, 1984; Tash & Stahler, 1982; Williams & Bank, 1984). Education respondents were typically teachers, principals, board members, and/or school system or district office administrators. Community mental health center (CMHC) personnel included program directors, team leaders, and case workers. Government officials were usually congressional staff members and policy-makers. Several samples were obtained from government social service agencies.

There was also considerable variability in instrumentation. Survey studies used questionnaires and interview schedules. Simulation studies favored rating forms. Other instruments included observation forms, participant and other naturalistic observation procedures, and content analysis forms.

Lastly, a variety of theoretical orientations have characterized research into evaluation use. The most frequent orientation in the sample was communication theory in which the question "Who says what to whom with what effects?" is of central importance. This orientation exclusively characterized simulation studies. Many of the remaining frameworks were either directly or indirectly associated with organizational behavior theory or decision theory. Some researchers relied on evaluation theories and models. There was no mention of any theoretical framework in 27 (42%) of the studies, consistent with the claim that evaluation utilization theory is not well developed.

Orientation Toward Dependent Variables

Table II presents the distribution of studies by orientation toward the dependent variable (i.e., evaluation utilization) and by study design. Three orientations toward

TABLE II

Orientation Toward-Dependent Variables by Study Design

Dependent variable orientation	Study design								
	Retrospective ($N = 35$)	Dependent variables		Longitudinal ($N = 9$)	Dependent variables		Simulation ($N = 21$)	Dependent variables	
		No.[a]	%[b]		No.[a]	%[b]		No.[a]	%[b]
Utilization as decision	Alkin et al. (1974)	7	100	Airasian et al. (1977)	6	100	Granville (1977)	1	100
	Barnette & Thompson (1978)	1	100	Pauley & Cohen (1984)	6	80	Heldt et al. (1973)	1	100
	Carter (1971)	4	100	King & Pechman (1984)	4	70	Lorenzen & Braskamp (1978)	2	100
	Dickman (1981)	1	100	Daillak (1983)	3	50	Brown et al. (1985)	1	10
	Johnson (1980)	1	100	Neigher (1979)	1	50	Newman et al. (1983a)[c]	1	10
	Van de Vall & Bolas (1982)	1	100	Windle et al. (1979)	1	50	Ory & Braskamp (1980)	1	10
	Weeks (1979)	1	100	Bigelow & Ciarlo (1976)	2	30			
	David (1978)	3	80	Dawson & D'Amico (1985)	1	10			
	Kennedy (1984)	7	80	Maher (1982)	1	10			
	Wholey & White (1973)	3	80						
	Leviton & Boruch (1983)	2	70						
	Yeh (1980)	5	70						
	Caplan (1976)	2	60						
	Glasman (1984)	5	60						
	Alkin et al. (1979)	1	50						
	Alkin and Stecher (1983)	1	50						
	Dickey (1980)	1	50						
	Fetler (1982)	9	50						
	Florio et al. (1979)	1	50						
	Goldberg (1978)	1	50						
	Osterlind (1979)	1	50						
	Patton et al. (1977)	1	50						
	Sproull & Zubrow (1981)	1	50						
	Rich (1979)	1	50						
	Rossman et al. (1979)	2	50						
	Williams & Bank (1984)	3	50						
	Brickell (1976)	4	40						
	Jordan (1977)	3	30						
	McGowan (1976)	2	30						
	Siegal & Tuckel (1985)	1	30						
	Tash & Stahler (1982)	2	30						

Utilization as education	Webber (1984)	1	100	Daillak (1983)	2	40	Becker et al. (1982)	1	10
	Salmon-Cox (1981)	2	70	Bigelow & Ciarlo (1976)	2	30	Thompson et al. (1981)	1	10
	Alkin et al. (1979)	1	50	Maher (1982)	2	30			
	Alkin & Stecher (1983)	1	50	King & Pechman (1984)	1	20			
	Fetler (1982)	9	50						
	Florio et al. (1979)	1	50						
	Patton et al. (1977)	1	50						
	David (1978)	1	30						
	Jordan (1977)	3	30						
	Leviton & Boruch (1983)	1	30						
	McGowan (1976)	2	30						
	Siegal & Tuckel (1985)	1	30						
	Kennedy (1983)	1	20						
	Williams & Bank (1984)	1	20						
	Yeh (1980)	1	20						
	Kennedy (1984)	1	10						
Utilization as processing	Kennedy (1983)	4	80	King & Pechman (1984)	1	20	Becker et al. (1982)	1	10
	Osterlind (1979)	1	50	Bigelow & Ciarlo (1976)	1	10			
	Rich (1979)	1	50	Pauley & Cohen (1984)	1	10			
	Rossman et al. (1979)	1	30						
	Siegal & Tuckel (1985)	1	30						
	Tash & Stahler (1985)	2	30						
	Kennedy (1984)	1	10						

TABLE II—*Continued*

Dependent variable orientation	Study designs								
	Retrospective ($N = 35$)	Dependent variables		Longitudinal ($N = 9$)	Dependent variables		Simulation ($N = 21$)	Dependent variables	
		No.[a]	%[b]		No.[a]	%[b]		No.[a]	%[b]
Utilization potential	Brickell (1976)	6	60	Dawson & D'Amico (1985)	5	70	Braskamp et al. (1978)	11	100
	Dickey (1980)	1	50	Maher (1982)	6	60	Brown et al. (1978)	12	100
	Goldberg (1978)	1	50	Neigher (1979)	1	50	Brown & Newman (1982)	16	100
	Sproull & Zubrow (1981)	1	50	Windle et al. (1979)	1	50	Brown et al. (1980, I)	1	100
	Glasman (1984)	3	40	Bigelow & Ciarlo (1976)	2	30	Brown et al. (1980, II)	3	100
	Tash & Stahler (1982)	3	40	Pauley & Cohen (1984)	1	10	Brown et al. (1980, III)	16	100
	Caplan (1976)	1	30				Newman et al. (1979, I)	3	100
	Jordan (1977)	3	30				Newman et al. (1979, II)	3	100
	McGowan (1976)	2	30				Newman et al. (1979, III)	3	100
	Rossman et al. (1979)	1	30				Newman et al. (1983b, I)[c]	3	100
	Salmon-Cox (1981)	1	30				Newman et al. (1983b, II)	46	100
	Wholey & White (1973)	1	30				Newman et al. (1983b, III)	64	100
	Williams & Bank (1984)	2	30				Newman et al. (1983b, IV)	1	100
	Yeh (1980)	1					Becker et al. (1982)	16	90
							Brown et al., (1985)	7	90
							Newman et al. (1983a)	9	90
							Ory & Braskamp (1980)	8	90
							Thompson et al. (1981)	7	90

[a] Estimate of the number of dependent variables operationalized according to particular orientation.

[b] Estimate of within-study (column) percentage of dependent variables operationalized according to particular orientation.

[c] Newman et al. (1983a) is Newman, Brown, Rivers, & Glock (1983); Newman et al. (1983b) is Newman, Brown, & Rivers (1983).

the dependent variable were evident in the studies: use as *decisionmaking*, use as *education*, and use as the *processing* of evaluation information. Some studies were unclear about type of utilization, focusing instead on utilization "potential."

In the context of support for discrete decisions, use was examined in relation to three different types of decisions having to do with program evaluation:

1. *Decisions about program funding.* For example, procurement, change in funding levels, renewal of funding, and initial funding requests (e.g., Alkin et al., 1974; Brickell, 1976; Dickman, 1981; Fetler, 1982; McGowan, 1976; Webber, 1984);
2. *Decisions about the nature or operation of a program.* For example, teaching of students, discrete actions of program staff, and staff efficiency (e.g., Carter, 1971; Florio et al., 1979; Herman & Dorr-Bremme, 1983; Siegal & Tuckel, 1985; Wholey & White, 1973);
3. *Decisions associated with program management.* For example, program planning, summative judgements of program impact, staff scheduling, and room arrangements (e.g., Glasman, 1984; Kennedy, 1984; Lorenzen & Braskamp, 1978; Maher, 1982; Sproull & Zubrow, 1981; Yeh, 1980).

Some studies operationalized use in terms of extent of use, impact, and helpfulness of the evaluation (e.g., Alkin & Stecher, 1983; Dickey, 1980; Goldberg, 1978; Jordan, 1977; Ory & Braskamp, 1980; Patton et al., 1977; Rich, 1979; Van de Vall & Bolas, 1982; Williams & Banks, 1984). Other studies had respondents make discrete decisions based on given evaluative information (Granville, 1977; Heldt, Braskamp, & Filbeck, 1973). Examples of utilization viewed as support for discrete decisions were found in studies employing all types of research design.

In the context of use as education, rather than providing the direct basis from which discrete decisions are made, evaluation may enlighten decisionmakers by influencing, for example, perceptions of current and ideal program structure. This view of utilization (as education) was evident in studies of the effectiveness of such specific program elements as instructional materials and strategies (e.g., Alkin & Stecher, 1983; Bigelow & Ciarlo, 1976; Yeh, 1980). Studies examining the use of evaluation information to assist with such matters as staff concept development, confirming prior impressions, and improving morale were also considered to be examples of use as education (e.g., Alkin et al., 1979; David, 1978; Jordan, 1977; Leviton & Boruch, 1983; Maher, 1982; Patton et al., 1977). Most of these studies employed retrospective or longitudinal research designs. Only two simulation studies treated use from the educational perspective: Thompson, Brown, and Furgason (1981) had respondents report on the extent to which an evaluation report influenced their views; and Becker, Kirkhart, and Doss (1982) stated that information in an evaluation report was useful in pointing out program strengths and weaknesses.

Several orientations are consistent with the view that use occurs when evaluation results are merely processed or thought about by decisionmakers. In retrospective field studies, for example, Tash and Stahler (1982), observed requests for a manual based on evaluation recommendations and Rich (1979) and Siegal and Tuckel (1985) reported the extent to which recommendations were given serious consideration. Osterlind (1979) and Kennedy (1983, 1984) provided examples of other retrospective studies that defined use as processing; they observed staff reaction to

formal evidence. Pauley and Cohen (1984) observed the frequency of management information system (MIS) report requests in their longitudinal study; this variable was an indirect indication that the MIS information was being processed by CMHC staff. In more direct fashion, Bigelow and Ciarlo (1976) questioned respondents about their understanding of evaluation data. In yet another longitudinal study, King and Pechman (1984) described instances of intentional non-use of evaluation data. Finally, in the one simulation study we were able to find that treated use as processing, Becker et al. (1982) derived a mean knowledge score based on the respondent's comprehension of an evaluation report.

Irrespective of study design, many of the dependent variables in the studies were more properly considered as measures of potential for use rather than use per se. In retrospective studies, use was often operationalized, for example, as the user's satisfaction with the evaluation information received and estimates of its influence on teachers' decisions (e.g., Brickell, 1976; Caplan, 1976; Glasman, 1984; Sproull & Zubrow, 1981; Williams & Bank, 1984). These studies focused at least partly on variables that, although linked to evaluation utilization, could more properly be thought of as antecedents. Longitudinal studies shared the same focus, examining such dependent variables as evaluation type and emphasis (Neigher, 1979; Windle, Majchrzak, & Flaherty, 1979).

Several simulation studies included, as indications of use, ratings of agreement with evaluator recommendations (Brown, Braskamp, & Newman, 1978; Brown & Newman, 1982; Brown, Newman, & Rivers, 1980, Studies II & III; Newman, Brown, & Littman, 1979, Studies I, II, & III; Thompson et al., 1981). To conceptualize use in this way is to concede that the evaluator is the best judge of program effectiveness and that non-acceptance of his or her recommendations is tantamount to non-use of the evaluation information. Recommendations, however, may well be rejected by persons who have either benefitted from the evaluation in more general ways or at least have seriously considered the evaluation information. Furthermore, similar types of use may be advocated by decisionmakers, regardless of how they feel about the evaluation information, the evaluator, or evaluation audience information needs (see Brown et al., 1980, Study I; Newman et al., 1979, Studies I, II, & III; Newman, Brown, & Rivers, 1983; and Ory & Braskamp, 1980). Unlike other studies deliberately addressing utilization potential or antecedents of utilization as dependent measures, there was a tendency for simulation studies to refer to these as utilization measures per se: This orientation was inconsistent with the conceptualization of dependent variables in our framework.

In summary then, there was considerable variability in how dependent variables were operationalized in the studies reviewed. To a large degree, orientations toward dependent variables were consistent with our initial conception of legitimate alternatives; uses as support for decisionmakers, educational uses, and use as the processing of evaluation results. Several studies, however, examined the effects of factors (independent variables) on dependent variables that might better have been treated as factors themselves. Sometimes authors unwittingly treated such factors as measures of use. Table II shows that simulation studies were more prone to operationalize dependent variables in this way and to give less emphasis to use as education or processing.

Orientation Toward Independent Variables

Two categories of independent variable were examined in the studies. One category focused on characteristics of *evaluation implementation* (e.g., timeliness, relevance); the other focused on characteristics of *decision or policy setting* (e.g., political climate, competing information). Considerable manipulation of these independent variables was evident among simulation studies, whereas longitudinal studies involved little manipulation and retrospective studies none. An analysis of independent variables by research design (comparable to the analysis reported in Table II) failed to reveal any trends.

Within the two categories of independent variables, we explored twelve specific factors. Six of the factors were concerned with characteristics of evaluation implementation and six with decision or policy setting.

Factors concerned with evaluation implementation were evident in studies exploring such questions as: Is the evaluation methodologically sound and/or believable? Is the evaluation relevant to decisionmaker needs (both overt and covert)? Are the results of the evaluation presented in an intelligible way and are they consistent with decisionmaker expectations? Are the results available in time to support the decision process? These factors were as follows:

Evaluation Quality: Characteristics of the evaluation process including sophistication of methods, rigor, type of evaluation model.

Credibility: Of the evaluator and/or the evaluation process, defined in terms of objectivity, believability, appropriateness of evaluation criteria, and so forth.

Relevance: Of the evaluation to the information needs of the decisionmaker(s) in terms of the purpose(s) of the evaluation and the organizational location of the evaluator.

Communication Quality: Clarity of reporting results to the evaluation audience(s) in terms of style, evaluator advocacy of the results, and breadth of dissemination.

Findings: Positive, negative, consistent with evaluation audience expectations, value for decisionmaking, and so forth.

Timeliness: In the dissemination of evaluation results to decisionmaker(s).

Factors concerned with decision or policy setting were not confined to organizational characteristics but were extended to the information needs of all relevant audiences of the evaluation, embracing such questions as: What are the important personal characteristics of the audience and to what extent are users of the evaluation involved with the target of evaluation? What are the characteristics of the target itself? What resources are available? Are there sources of information such as political undercurrents or previous experiences that compete with emerging evaluation information? Who are the primary decisionmakers and what are their attitudes toward evaluation and/or research in general? How are decisions made and is this structure flexible? The six factors associated with these questions were as follows:

Information Needs: Of the evaluation audience(s), including type of information sought, number of evaluation audiences with differing information needs, time pressure, and perceived need for evaluation.

Decision Characteristics: Impact area, type of decision, program novelty, and the significance of the decision or evaluation problem, and so forth.

Political Climate: Political orientation of commissioners of the evaluation, de-

pendence of the decisionmaker(s) on external sponsors, inter- and intraorganizational rivalries, budget fights, power struggles, and so forth.

Competing Information: From sources beyond the evaluation (personal observations, staff, peers, etc.) bearing upon the problem and competing with evaluation data.

Personal Characteristics: Defined in terms of the decisionmakers' organizational roles, information processing style, organizational experience, social characteristics, and so forth.

Commitment and/or Receptiveness to Evaluation: Attitudes of the decisionmaker(s) toward evaluation, organizational resistance, open-mindedness, and so forth.

Figure 1 displays visually a framework for conceptualizing evaluation utilization. It is based on the combined analysis of dependent and independent variables explored in the studies reviewed. The utilization construct reflects the minimum requirement, namely, that evaluation results must be psychologically processed by the decisionmaker prior to conceptual growth or support for discrete decisions. Twelve specific independent variables or factors are clustered about one of two higher order, hypothetical dimensions. The broken bi-directional arrow separating the dimensions indicates that they are dynamic entities interacting with one another to produce effects on use. Our discussion of the relationship between independent and dependent variables is organized around this framework.

Relationships

In this section, the relationships between dependent and independent variables are examined in an attempt to make explicit the nature of the relationships, assess the relative importance of factors to evaluation use, and determine whether the relative importance of factors varies with type of use.

The term *observed relationship* (as used here) connotes a statistically significant relationship between a single independent and a single dependent variable in

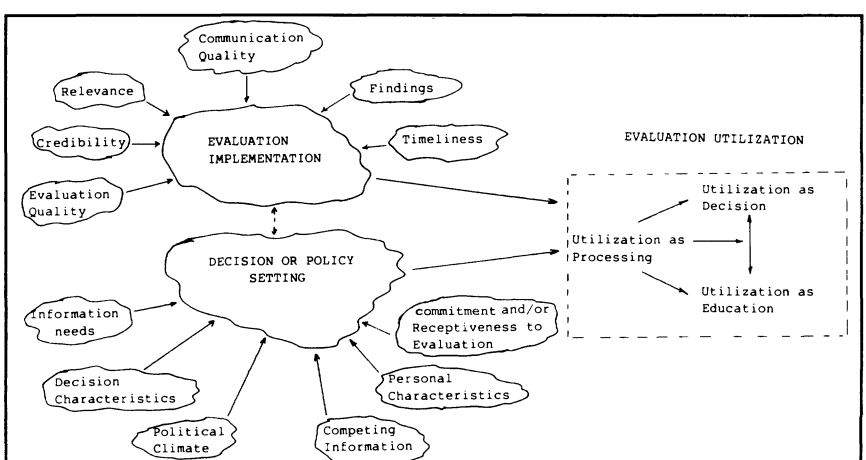

FIGURE 1. Evaluation utilization framework.

quantitative studies. As a minimum, if an independent variable were shown to interact with another to produce an effect, this would be considered evidence of an observed relationship (although it would not be counted twice if a statistically significant main effect for that variable were also evident). In qualitative studies, observed relationships are typically manifest in prose or narrative descriptions. The distinction between these general methodological orientations is important, not only because it is more difficult to identify observed relationships in qualitative studies but also because of differences in the treatment of *observed non-relationships.* An observed non-relationship means either that no indication of statistical significance between a single dependent and a single independent variable has been observed in quantitative studies or that there is not a relationship between a single independent and a single dependent variable in the narrative of qualitative studies. In qualitative studies, non-relationships are often not reported, especially in studies that do not employ a predefined conceptual framework or that use a grounded theory (Glaser & Strauss, 1967) approach to empirical investigation.

Table III was constructed by coding every observed relationship and non-relationship according to type of factor and type of use. Twelve codes were used for type of factor, corresponding to the six evaluation implementation and six decision or policy-setting factors previously described. There were four codes for type of use: use as support for decision, education, psychological processing of evaluation results by the decisionmaker, and potential for use to occur.

Using reported relationships and non-relationships as the unit of analysis, the following information was entered in each of the (12×4) rows of Table III to guide our discussion:

Number of Reports: Estimated number of observed relationships plus non-relationships.

Number of Relationships: Estimated number of observed relationships (i.e., number of reports minus number of observed non-relationships).

Number of Studies: Estimated number of empirical studies reporting a minimum of one observed relationship or non-relationship.

Prevalence of Relationship Index: $S/R \times O$, where S represents the number of studies, R is the number of reports, and O is the number of observed relationships.

Empirical Support: A listing of the empirical studies that reported a minimum of one observed relationship or non-relationship.

The prevalence of relationship index was created to compare the relative strengths of the relationships between dependent and independent variables in the studies—to serve a purpose similar to effect size statistics now widely used in conventional meta-analytic studies. Although this purpose was important in the present review, attaching a strength of relationship code at the level of the observed relationship would have been onerous, uneconomical, and of dubious reliability with much of our data. Indeed, in a number of the studies, the information required on which to base such judgments was simply not available, and the demands of computing an effect size statistic was not a compelling enough reason to limit the sample. Given our interest in achieving an exhaustive sample of the empirical literature, the prevalence of relationship index was a reasonable approximation of strength of relationship. The S/R component of this index may be thought of as a weighting factor that adjusts the number of observed relationships for the differential frequency of reports due to quantitative as opposed to qualitative studies.

TABLE III
Observed Relationships by Factor and Type of Use

Factor	Type of use[a]	Number of reports	Number of relationships	Number of studies	Prevalence of relationship index				Empirical support[b]
					Overall	Rank	Within use types	Rank	
Evaluation implementation factors									
Evaluation quality	Decision	66	49	26			19.3	1	1, 2, 3, 5, 11, 16, 19, 21, 22, 26, 31, 35, 36, 37, 48, 49, 50, 51, 52, 54, 55, 56, 59, 61, 64, 65
	Education	7	6	7	20.5	1	6.0	1	2, 19, 31, 35, 37, 50, 62
	Processing	3	3	3			3.0	2	49, 51, 55
	Potential	15	11	6			4.4	1	11, 31, 37, 48, 51, 56
Credibility	Decision	17	7	8			3.3	12	2, 3, 19, 20, 21, 22, 50, 63
	Education	4	3	4	6.8	11	3.0	8	3, 18, 20, 50
	Processing	1	1	1			1.0	9	7
	Potential	23	7	7			2.1	7	8, 11, 12, 13, 40, 41, 42
Relevance	Decision	34	26	14			10.7	4	2, 19, 20, 22, 24, 38, 39, 49, 50, 52, 53, 56, 59, 62
	Education	3	1	2	13.2	5	0.7	11.5	20, 50
	Processing	2	2	2			2.0	5	6, 7, 53
	Potential	2	1	1			0.5	10	37, 57
Communication quality	Decision	29	18	12			7.5	8	2, 7, 9, 20, 21, 24, 30, 49, 53, 55, 59, 64
	Education	4	3	3	7.7	8	2.3	10	6, 20, 58
	Processing	4	3	4			3.0	2	6, 7, 49, 53
	Potential	134	39	10			2.9	6	6, 10, 11, 12, 13, 24, 40, 41, 57, 58
Findings	Decision	37	29	17			13.3	2	2, 3, 7, 9, 17, 19, 24, 30, 35, 34, 49, 50, 52, 56, 59, 62
	Education	6	4	5	14.8	3	3.3	6	2, 19, 35, 50, 54
	Processing	2	2	2			2.0	5	49, 55
	Potential	0	0	0			0.0	11.5	
Timeliness	Decision	8	6	7			5.3	10	19, 21, 24, 50, 52, 55, 62
	Education	3	1	2	5.1	12	0.7	11.5	50, 54
	Processing	0	0	0			0.0	12	
	Potential	0	0	0			0.0	11.5	

				Decision or policy setting factors				
Information needs	Decision	12	12	8	7.3	9	8.0	7	2, 7, 16, 19, 23, 24, 30, 35
	Education	4	4	3			3.0	8	2, 19, 23
	Processing	1	1	1			1.0	9	32
	Potential	19	5	4			1.1	11	7, 9, 14, 54
Decision characteristics	Decision	44	29	20	16.4	2	13.2	3	3, 15, 16, 19, 21, 22, 24, 25, 26, 27, 28, 29, 36, 47, 54, 56, 60, 61, 63, 65
	Education	3	3	3			3.0	8	4, 24, 63
	Processing	1	1	1			1.0	9	53
	Potential	63	43	6			4.1	2	8, 15, 27, 36, 42, 47
Political climate	Decision	11	11	7	10.0	7	7.0	9	2, 9, 16, 26, 34, 35, 50
	Education	6	6	4			4.0	4	2, 34, 35, 50
	Processing	1	1	1			1.0	9	34
	Potential	9	9	4			4.0	3	9, 38, 39, 64
Competing information	Decision	14	14	10	12.6	6	10.0	5	4, 7, 15, 24, 26, 29, 36, 50, 56, 63
	Education	5	5	4			4.0	4	32, 33, 50, 54
	Processing	4	4	2			2.0	5	32, 53
	Potential	7	4	1			0.6	9	15
Personal characteristics	Decision	16	8	10	7.1	10	5.0	11	4, 5, 18, 24, 25, 36, 50, 56, 60, 65
	Education	4	4	4			4.0	4	2, 18, 50, 58
	Processing	1	1	1			1.0	9	53
	Potential	187	51	13			3.6	4	8, 16, 25, 40, 41, 42, 43, 44, 45, 46, 47, 53, 58
Commitment and/or receptiveness to the evaluation	Decision	15	14	10	14.1	4	9.3	6	18, 19, 20, 21, 24, 25, 49, 52, 59, 63
	Education	6	5	5			4.2	2	18, 19, 20, 24, 63
	Processing	3	3	3			3.0	2	32, 49, 55
	Potential	11	11	3			3.0	5	30, 57, 62

[a] Some studies report descriptions of use separately from discussions of factors that affect use. In such cases if more than one type of use was described the factor was assumed to affect all types, and relationships were coded accordingly.

[b] Study numbers correspond with Table I.

We now turn to a discussion of relationships associated with each of the 12 factors corresponding to the order of presentation in Table III. Relationships are described in the terms in which they were reported in the study. For example, the positive relationship to use of relevance of evaluation was reported both in terms of decreased relevance leading to diminished use or to non-use, and increased relevance leading to heightened use. Instances where results were inconsistent (i.e., presence of negative relationship or absence of relationship) are noted in the subsequent text.

Evaluation Implementation Factors

Evaluation quality. Over 40% of the studies reported observed relationships and non-relationships between evaluation use and quality of the evaluation. Dependent variables in the majority of these reports were use of evaluation as opposed to potential for use. Evaluation quality was typically defined as methodological sophistication, type of approach to the evaluation problem, or the intensity of the evaluation activities.

Weeks (1979), Van de Vall and Bolas (1982), and Siegal and Tuckel (1985) reported that increased methodological sophistication served to inhibit use (as defined in terms of use index variables and observations). Similarly, Yeh (1980) reported that instructional decisions were more likely to be based on less sophisticated teacher-made tests. Several authors, however, including Alkin et al. (1974), Granville (1977), Patton et al. (1977), Weeks, and Van de Vall and Bolas, showed that improved methodological sophistication was positively related to use in terms of support for decisions and the conceptual development of users. For example, early design and timing plans (Van de Vall & Bolas) and sophistication of sampling procedures (Weeks) were positively correlated with use index scores. Both Dickey (1980) and Dickman (1981) found no relationship between this factor and use, but Dickman observed that methodological consistency was directly related to perceived program change.

Evaluations with a focus on program process or implementation were generally found to be more useful than those dealing strictly with outcomes (Jordan, 1977; Leviton & Boruch, 1983; Windle et al., 1979). Other types of evaluation approach that enhanced either use or the potential for use included implementation of the A-VICTORY technique (Maher, 1982), a structured consultative intervention (Pauley & Cohen, 1984), and cost benefit orientations (Lorenzen & Braskamp, 1978). Favorable effects of the evaluators' approach and of the type of evaluation information were reported by Caplan (1976), David (1978), Alkin et al. (1979), Osterlind (1979), and Brown and Newman (1982).

Finally, Barnette and Thompson (1979) found that the intensity of the evaluation (defined as frequency of observation) led to enhanced use of the evaluation findings for instructional decisions. Airasian, Kellaghan, Madaus, & Pedulla (1977) found that test administration and subsequent reporting of results affected teacher decisions about ranking students. On the other hand, Salmon-Cox (1981) reported that increased frequency of standardized testing resulted in diminished use for instructional decisions.

Credibility. Over a quarter of the studies examined relationships between utilization and the credibility of the evaluator or evaluation process (usually defined in

terms of appropriateness of evaluation criteria). Some viewed evaluation credibility as a function of reputation (Alkin et al., 1979; Dawson & D'Amico, 1985); for others, credibility was associated with the title or sex of the evaluator or the evaluator's belief in his or her work (Braskamp et al., 1978; Dickman, 1981; Newman et al., 1979, Study II). Except for sex of the evaluator, about which results were conflicting, each of these views of evaluation credibility were positively related to use.

When evaluations were perceived by decisionmakers as having high face validity or when they were emphasized as important activities, use and the potential for use appeared to be greater (Brown et al., 1980; Williams & Bank, 1984); similarly, the collection of data perceived as inappropriate by decisionmakers was associated with reduced use (Daillak, 1983; David, 1978). Finally, Patton et al. (1977), failed to find significant relationships between the appropriateness of evaluation method and their broad definition of use.

Relevance. The 18 studies exploring the relationship between relevance and use defined relevance as either the extent to which evaluation was geared to the audience(s) or whether the evaluator was internal or external to the organization (organizational location); internal evaluators were considered to be more knowledgeable about their organizations' characteristics.

Most studies found that evaluations that reflected knowledge of the context in which the evaluations were to be used, appealed to preferences of the decisionmakers, sought consensus about the evaluation problem, or demonstrated insight into program operations and decisionmaking, were associated with higher levels of use (Dawson & D'Amico, 1985; Osterlind, 1979; Rich, 1979; Rossman et al., 1979; Van de Vall & Bolas, 1982). Evaluations that ignored these issues were associated with relatively low levels of use (David, 1978; Florio et al., 1979; McGowan, 1976; Neigher, 1979; Sproull & Zubrow, 1981; Wholey & White, 1973). Only Patton et al. (1977) found such issues to be unrelated to subsequent use.

As a measure of relevance, organizational location of the evaluator yielded conflicting results; it was found to be related to use in four studies (Alkin et al., 1974; David, 1978; McGowan, 1976; Van de Vall & Bolas, 1982) and unrelated in three (Alkin et al., 1974; Dickman, 1981; Weeks, 1979). All relationships that were reported suggested that internal evaluations were more useful than external evaluations.

Communication quality. One third of the studies inquired about the relationship between use and communication style, ongoing communication activities, and breadth of dissemination.

Evidence about communication style was inconclusive. Oral presentations of results (along with written reports), broadly framed and comprehensive recommendations, and nontechnical language contributed to higher impact scores, improved readability, and to greater awareness and appreciation of results (Becker et al., 1982; Bigelow & Ciarlo, 1976; Rossman et al., 1979; Tash & Stahler, 1982; Van de Vall & Bolas, 1982). At the same time, however, style was shown to have little if any influence on program decisions, knowledge of results, and other measures of use (Alkin et al., 1974; Bigelow & Ciarlo; Dickey, 1980; Florio, et al., 1979) or potential for use (Brown & Newman, 1982; Brown et al., 1980, Studies II & III; Newman et al., 1979, Studies I & II; Thompson et al., 1981).

By contrast, enhanced use appeared to be strongly associated with ongoing

communication and/or close geographical proximity between evaluator and decisionmaker (Florio et al., 1979; Johnson, 1980; Osterlind, 1979; Rossman et al., 1979). Advocacy by the evaluator of his/her results and forcefulness of communication was also associated with greater use (Brickell, 1976; Dawson & D'Amico, 1985; Johnson; Siegal & Tuckel, 1985; Van de Vall & Bolas, 1982). Finally, Van de Vall and Bolas found that dissemination breadth (e.g., reports geared to the public vs. scholarly journals) resulted in higher utilization scores, whereas Dickey (1980) found that this factor was unrelated to the utilization index that she constructed.

Findings. Nineteen studies provided evidence about the effect on use of evaluation findings usually defined in terms of congruence with decisionmaker expectations and value for decisionmaking. None of these were associated with utilization potential.

Most of these studies reported that when evaluation findings were congruent with decisionmaker expectations, acceptance and utilization increased; incongruent findings tended to be ignored and other information used instead (Bigelow & Ciarlo, 1976; Brickell, 1976; Carter, 1971; David, 1978; Johnson, 1980; Kennedy, 1984; King & Pechman, 1984; Osterlind, 1979; Wholey & White, 1973). The only studies in which unexpected findings led to higher utilization were those of Salmon-Cox (1981) and Van de Vall and Bolas (1982); according to Patton et al. (1977), the nature of findings was comparatively unimportant in predicting use.

Evaluation findings were reported to be of most use for such purposes as legislation, organizational development, and other types of federal and local decisionmaking, especially when the findings were practical and conclusive and when they identified alternative courses of action for decisionmakers (Alkin et al., 1974; Carter, 1971; David, 1978; Florio et al., 1979; Leviton & Boruch, 1983; Rich, 1979; Sproull & Zubrow, 1981). Siegal and Tuckel (1985) were alone in suggesting that general as opposed to specific recommendations were more likely to be given serious attention by decisionmakers.

Timeliness. Less than 15% of the studies examined the relationship between utilization and the timeliness of communication of evaluation results. The timely provision of evaluation results was positively related to a utilization index score (Dickey, 1980) and to use for student assessment (Salmon-Cox, 1981), programmatic decisions (David, 1978; Siegal & Tuckel, 1985; Wholey & White, 1973), and policy decisions (Florio et al., 1979; Rich, 1979). Patton et al. (1977), however, found that lateness of reports and timeliness of evaluation had little impact.

Decision or Policy Setting Factors

Information needs. Independent variables categorized under information needs were very varied. They included (a) intensity of information needs, (b) type of information required, and (c) variance in audience needs for information. Information needs were examined by 18% of the studies.

As decisionmakers perceived a greater need for information, so their use of evaluation results increased (Bigelow & Ciarlo, 1976; Caplan, 1976; Fetler, 1982; Johnson, 1980; Leviton & Boruch, 1983) and so did their tendency to agree with evaluation recommendations (Brown et al.,1980, Study III). Similarly, David (1978) found that high program stability was associated with the non-use of evaluation

results for program improvement. However, Brickell (1976) found that greater need had an adverse effect on the preparation of the evaluation report.

The type of information required by decisionmakers also influenced utilization. The need for information geared to widening program options had a positive effect on the use of evaluation for legislative change (Leviton & Boruch, 1983). Kennedy (1983) reported positive relationships between the need to examine strengths and weaknesses and the processing of evidence (data). Bigelow and Ciarlo (1976) and Salmon-Cox (1981) both reported that when required information was incongruent with the evaluation information, the relevance of the evaluation was reduced.

Variability in audience needs was found to be negatively related to the use of evaluation results for decision purposes by David (1978), Alkin et al. (1979), and Florio et al. (1979).

Decision characteristics. Over a third of the studies reported relationships between utilization and characteristics of the decision, for example, area of decision, decision context, and significance of decision.

Most of these studies examined evaluation use for different areas of decision. Use was reported to be high in such areas as instructional effects, program strengths and weaknesses, faculty self-improvement, facilities administration, and program administration. Evaluation was also reported to be of most use at early stages in the decisionmaking process (Alkin & Stecher, 1983; Florio et al., 1979). Decision areas in which utilization was low included, for example, program implementation, supplementary budgets, student problems, and promotion decisions (Caplan, 1976; Dickman, 1981; Goldberg, 1978; Heldt et al., 1973; Herman & Dorr-Bremme, 1983; Ory & Braskamp, 1980; Williams & Bank, 1984).

The relationship between the context of a decision and evaluation was examined by Herman and Dorr-Bremme (1983) who reported that test results were used for communication with parents and for instructional decisions more in secondary schools than in elementary schools. Yeh (1980) found that schools in high socioeconomic status regions more often used tests for reporting to parents than did low socioeconomic status schools. Sproull and Zubrow (1981) also found that private as opposed to public school system central office administrators were more inclined to use of tests for decisionmaking. Newman et al. (1979, Study III) found that ratings of the evaluator were higher when they were associated with business rather than with education evaluation reports. Weeks (1979) observed significant correlations between a use index and the number of decision participants (positive) and program goal specificity (negative). He also noted that four other measures of decision context were unrelated to use; likewise Dickey (1980) noted that three project characteristic variables were unrelated.

Highly significant decisions (according to Caplan, 1976) and decisions generating personal conflict (according to Brown et al., 1985 and Newman, Brown, Rivers, & Glock, 1983) were associated with high levels of evaluation use and information needs. But several nonrelationships were also reported.

Political climate. Political influences on the use of evaluation occurred at both the organizational and extra-organizational level. Only 10 studies observed relationships between political factors and evaluation use, and there were no observed non-relationships reported.

Within an organization, existing staff views, organizational arrangements, and rewards shaped the evaluation and influenced the extent of evaluation use (Alkin

et al., 1979; Brickell, 1976). Diminished use occurred if evaluation results were perceived as a threat by staff, if they felt "their hands were tied," if key staff members left the organization, or if there were internal debates and budget squabbles. Diminished use was also associated with interagency rivalries and with pressures on evaluators from program operators and directors (Brickell; King & Pechman, 1984; Leviton & Boruch, 1983; Patton et al., 1977). However, these studies noted that personal motives could lead to enhanced use and that different political factors sometimes neutralized each other, permitting greater freedom for the evaluator to work out evaluation strategies.

King and Pechman (1984) found that federal requirements for funding decisions enhanced the use of evaluation for accountability, and McGowan (1976), Neigher (1979), and Windle et al. (1979) described effects of this variable on shaping the evaluation. Patton et al. (1977) described the role of evaluation in reducing uncertainty for decisionmakers in power struggles with federal administrators. Other examples of extra-organizational political influences were identified by Granville (1977) and Alkin et al. (1979).

Competing information. When confronted with a problem, decisionmakers rely on various sources of information in addition to evaluation results. Fourteen studies observed relationships and non-relationships indicating in what way alternative sources of information competed with evaluation results for decisionmaking and problem solving.

Working knowledge, a term attributable to Kennedy (1983, 1984) and derived from personal experiences, beliefs, values, interests, and goals, was shown to be a powerful competitor with evaluation data. Several studies showed that knowledge derived from personal observation reduced the importance of evaluation for decisionmaking (Alkin & Stecher, 1983; Bigelow & Ciarlo, 1976; Herman & Dorr-Bremme, 1983; Salmon-Cox, 1981; Sproull & Zubrow, 1981; Williams and Bank, 1984). Such knowledge also facilitated the processing of formal evidence and the identification of relevant facts (Kennedy, 1983, 1984).

Other information shown to diminish reliance on evaluation results included program requirements and parent, staff, and peer input for school decisions (Alkin & Stecher, 1983; Florio et al., 1979), feedback from staff for curriculum and management decisions (Bigelow & Ciarlo, 1976; Sproull & Zubrow, 1981), teacher-made tests for assessment purposes (Salmon-Cox, 1981), peer support for program adoption (Granville, 1977), and support for a decision from supervisors (Brown et al., 1985; Rossman et al., 1979). Lorenzen and Braskamp (1978) also found that the processing of multiple sets of data reduced decisionmaker dependency on evaluation.

Personal Characteristics. Over a third of the sample reported relationships and non-relationships between evaluation use and personal characteristics of the decisionmaker including organizational role, training and experience, and leadership characteristics. Observations were made with respect to effects on the use of evaluation as well as the potential for evaluation use.

Findings about the relationship between a person's role in the organizational hierarchy and evaluation utilization were conflicting. A small number of studies supported the view that greater utilization occurred among those with positions higher in the hierarchy, for example, central office administrators as opposed to principals, and principals as opposed to teachers (Braskamp et al., 1978; Daillak,

1983; Rossman et al., 1979). However, Sproull and Zubrow (1981) reported contrary findings in teachers' (as compared with administrators') use of evaluation for curriculum decisions, whereas Florio et al. (1979) detected no differences among staff groups for policy decisionmaking. Several simulation studies investigated the effects of organizational role (Newman et al., 1979, Studies I & II; Newman, Brown, Rivers, & Glock, 1983; Thompson et al., 1981). Significant differences were found between school board members and administrators and between business and education administrators, in terms of such variables as reliance on personal experience, information needs, report and evaluator ratings, and agreement with evaluator recommendations. But several non-relationships with these variables were also reported.

The relationship between utilization and experience was also ambiguous. More experience was related to enhanced use for teachers and trustees (Webber, 1984; Yeh, 1980). But years of experience did not make a difference with teachers' use of evaluation for instructional judgments (Barnette & Thompson, 1978) or CMHC administrators' use of evaluation for decisionmaking (Lorenzen & Braskamp, 1978). Caplan (1976) found that educational background contributed to different information processing styles, and he discussed implications for potential uses of evaluation.

Leadership characteristics were shown to be positively linked to evaluation use in several studies. These characteristics included more skills and initiative (Alkin et al., 1979), higher levels of leadership, caring, and interest (Patton et al., 1977), and internal versus external locus of control (Newman, Brown, & Rivers, 1983, Studies I, II, III, & IV). But Glasman (1984) reported no differences between high and low effective principals in ratings of the use of achievement tests.

User commitment and/or receptiveness to evaluation. About a quarter of the studies examined the effect on utilization of the decisionmakers' commitment to the evaluation and/or their attitude toward evaluation. Most of these reports were connected to direct measures of use.

Commitment was often defined in terms of the extent to which the decisionmaker was involved in the evaluation. Greater involvement contributed to high levels of use in terms of an impact score (Van de Vall & Bolas, 1982), and of policy formulation (Rich, 1979), as well as enhanced staff learning about the program, programmatic decisionmaking and an increase in such factors as communication, relevance, information processing, and credibility (Dawson & D'Amico, 1985). Other examples of the positive effects of involvement on evaluation use and potential for use were provided by Johnson (1980), Tash and Stahler (1982), and Daillak (1983). Only Dickey's (1980) results were not consistent: in this study, users' involvement in the evaluation bore no relationship to a utilization impact score.

Users' attitudes toward evaluation appeared to have an unambiguous positive relationship with utilization (Dickey, 1980; Glasman, 1984; Kennedy, 1983; Osterlind, 1979; Siegal & Tuckel, 1985; Williams & Bank, 1984). In each of these studies, more favorable attitudes were associated with higher levels of use. Similar results were reported by David (1978) and Daillak (1983); negative attitudes toward evaluation and organizational resistance were associated with diminished utilization.

Relative Influence of Factors

A "prevalence of relationship" index was developed in order to answer three questions about the relative influence of factors: (a) What was the strength of factors over all types of use and potential for use? (b) Did factors differ in strength depending on type of use?, and (c) Which factors were most influential across varying types of use? To obtain an estimate of the overall prevalence of the 12 factors, the prevalence of relationship index (previously described) was calculated, based upon all types of use and utilization potential (see Table III).

The most prevalent factors were evaluation quality and decision characteristics, with overall index scores of 20.5 and 16.4, respectively. As we noted before, there was inconsistent evidence in the literature about the relationships between use and methodological standards or intensity of the evaluation; also, a moderate percentage of the reports that we identified in association with these factors were non-relationships (24% for evaluation quality and 32% for decision characteristics). The next most prevalent factor influencing overall use was evaluation findings, with an associated index score of 14.8; only 22% of these reports were non-relationships. We found some instances where negative findings contributed to enhanced levels of use, but for the most part, our observations were consistent in suggesting that positive results congruent with decisionmaker expectations tended to be used more.

The users' commitment and/or receptiveness to evaluation and evaluation relevance both yielded overall index scores of 13.2, and were the next most prevalent factors. Evidence associated with these factors was consistent and the percentages of identified reports that were categorized as non-relationships were comparatively low: 6% for commitment and 22% for relevance. Although we observed some inconsistent findings with respect to the influence of the decisionmaker's personal knowledge, information that competed with evaluation results was found to be the sixth most influential factor, with an overall prevalence of relationship score of 12.6. Ninety percent of the reports that we identified were observed relationships. The relative ranking of the six remaining factors can be seen in Table III.

The prevalence of relationship scores associated with each type of use and potential for use (see Table III) were rank ordered to determine whether the relative influences of factors varied as a function of type of use. The rank ordering of factors affecting utilization as decision and utilization as processing varied only marginally from the overall rank order. With utilization as decision, no factor changed in rank by more than two positions and the six most influential factors were the same as those mentioned in the overall ranking (Spearman Rank Order Correlation, $r = .92$, $p < .001$). This finding is not surprising given that the majority of dependent variables were categorized according to this type of use. In the case of utilization as processing, most of the factors did not vary by more than two rank order positions from the overall or decision ranking ($r = .62$, $p < .05$ and $r = .56$, $p < .05$).

The rank ordering associated with utilization as education was different; there were no statistically significant correlations with overall, decision, or processing rankings. Evaluation quality was still found to be the most prevalent factor affecting this type of use. However, user commitment and/or receptiveness to the evaluation was the second most influential factor in determining the extent of staff or user conceptual development. Three factors, political climate, competing information,

and personal characteristics, were tied in rank and represented the next most influential factors at this level; each of these increased in rank relative to overall and decision level rankings. The most notable change was associated with personal characteristics of the users; this received ranks of 10 and 11, respectively, in overall and discrete decision rank orderings, suggesting that characteristics of decision or policy setting, particularly those associated with organizational resistance toward evaluation and administrative style, are important determinants of the extent to which users learn from an evaluation. The greatest negative shift in ranking was associated with relevance of the evaluation to decisionmaker needs; this dropped to a tie with timeliness in having the least influence on utilization as education. One possible interpretation is that evaluations that are relevant to decision-specific needs do not contribute substantially to unanticipated decisionmaker insights.

Findings associated with utilization potential were more difficult to interpret, primarily because the dependent variables (which we judged to be antecedent constructs) were exceptionally heterogeneous. They ranged from modifications in the evaluation process, through improved decisionmaker attitudes toward evaluation, to change in users' information needs. Again, quality of the evaluation ranked first in influencing such variables. The next most prevalent factors (in rank order) were decision characteristics, political climate, user personal characteristics, and user commitment and/or receptiveness to evaluation; each was associated with the decision or policy setting hypothetical dimension. A more fine-grained analysis, identifying the classes of dependent variable affected by these factors, is beyond the scope of the present review, but such an analysis could be worthwhile in shedding light on the way in which antecedent variables interact with each other to affect use.

Summary and Conclusions

A sample of 65 studies (covering 1971 to 1985) was analyzed as a basis for describing current research-based knowledge about evaluation use. The analysis inquired about methodological characteristics, the nature of dependent variables, and factors (independent variables) influencing the use of evaluation results. Four different operational definitions of use were evident: discrete decisionmaking, education or conceptual development, cognitive processing of evaluation information, and potential for use. The last of these orientations, we concluded, might better be considered as a set of antecedent variables in relation to the other three definitions of use.

Twelve factors influencing one or more types of use were identified from the review. Six of these factors, concerned with the implementation of evaluations, were: evaluation quality, credibility, relevance, communication, the findings themselves, and the timeliness of evaluations for users. The remaining factors, concerned with features of decision or policy setting, were: information needs of users, decision characteristics, political climate, competing information, personal characteristics of users, and user commitment and receptiveness to evaluation information. A preliminary framework suggesting relationships among the factors and the different definitions of the dependent variables was developed (Figure 1).

Our analysis showed that the relative influence of the 12 factors varied as a function of the type of use. Factors affecting the use of evaluation for decisionmak-

ing and the psychological processing of evaluation results were reasonably evenly distributed over the evaluation implementation and decision or policy setting categories. Factors associated with the latter category (user commitment, personal characteristics, competing information, political climate) were more influential in affecting conceptual gains associated with evaluation results. Improved evaluation quality (methodological sophistication, evaluation approach, evaluation intensity) appeared to affect all types of use favorably.

Overall evaluation use seemed to be most strongly evident when:

- evaluations were appropriate in approach, methodological sophistication, and intensity;
- the decisions to be made were significant to users and of a sort considered appropriate for the application of formally collected data;
- evaluation findings were consistent with the beliefs and expectations of the users;
- users were involved in the evaluation process and had a prior commitment to the benefits of evaluation;
- users considered the data reported in the evaluation to be relevant to their problems;
- a minimum amount of information from other sources conflicted with the results of the evaluation.

Studies in this review spanned a wide range of decision contexts and evaluation settings, thereby strengthening the generalizability of the findings. Nevertheless, the framework developed as part of the review was tentative and should be viewed rather as a stimulus for subsequent research. Questions that future research might address include the following:

Can the existence of the hypothetical, higher order categories—evaluation implementation and decision or policy setting characteristics—be verified? If so, which has the most impact on the use of evaluation results?

To what extent do factors vary in terms of their impact on decisionmakers' use of evaluation results? Which factors have the greatest impact? How do factors interact with one another to affect use?

Do the conditions under which utilization as decision occurs differ from conditions associated with utilization as education, or with utilization as processing?

At a more detailed level, there were large differences among the 12 factors in the amount of evidence about their influence on use. Even a cursory inspection of Table III indicates that considerable attention has been given, for example, to evaluation quality, communication quality, and decision characteristics. Subsequent research might be advised to focus on specific factors about which relatively little is known (e.g., timeliness, information needs, political climate, competing information) or about which results remain conflicting (e.g., decision characteristics, personal characteristics of the users). Furthermore, adopting potential for utilization as a dependent variable, particularly when interpretations are made in terms of utilization, appears problematic and might better be avoided in subsequent research.

In our view, the most important implication for practice concerns the conduct of evaluation studies. Results argue strongly for evaluation procedures that at the outset generate information helpful to users in carrying out their decisions. Results also suggest that evaluation users should be involved in ways manageable for them, in the planning and carrying out of the evaluation. Such involvement seems likely

to ensure the credibility and relevance of results, and to increase commitment to the evaluation process as a whole; it also seems likely to help in resolving problems of timeliness of reporting results and in responding to user information needs. Cooley (1983), an advocate of the general principal we are discussing here, described it as client-oriented evaluation.

Focusing evaluations on users' perceptions of the evaluation problem might be seen by some as unduly restrictive and unlikely to serve user education needs. Furthermore, user involvement in the evaluation process could lead to co-optation of the process, with loss of objectivity and a threat to the evaluator's integrity. These negatives must be weighed against the cost of widespread non-use of results. Involving the user in the evaluation process such that it becomes a forum for the mutual education of evaluator and decisionmaker may present a viable alternative to non-use or co-optation. It may also set the stage for decisionmakers to advance from discrete evaluations of discrete programs to systematic procedures for monitoring an organization's functioning more comprehensively. Cooley (1983) has suggested that such a step is crucial to improving the performance of educational systems.

References

Airasian, P. W., Kellaghan, T., Madaus, G. F., & Pedulla, J. J. (1977). Proportion and direction of teacher rating changes of pupils' progress attributable to standardized test information. *Journal of Education Psychology, 69*(6), 707–709.

Alkin, M. C., Daillak, R., & White, B. (1979). *Using evaluations.* Beverly Hills, CA: Sage.

Alkin, M. C., Koescoff, J., Fitz-Gibbon, C., & Seligman, R. (1974). *Evaluation and decision-making: The Title VII experience.* Los Angeles: Center for the Study of Evaluation.

Alkin, M. C., & Stecher, B. (1983). Evaluation in context: Information use in elementary school decision making. *Studies in Educational Evaluation, 9,* 23–32.

Barnette, J. J., & Thompson, J. C. (1978). A descriptive assessment of the effect of evaluations on instruction. *Studies in Educational Evaluation, 5,* 77–88.

Becker, H. A., Kirkhart, K. E., & Doss, D. (1982). Evaluating evaluation reports. *CEDR Quarterly, 15,* 5–8.

Bigelow, D. A., & Ciarlo, J. A. (1976). The impact of therapeutic effectiveness data on community mental health center management. In G. V Glass (Ed.), *Evaluation Studies Review Annual, 1* (pp. 371–382). Beverly Hills, CA: Sage.

Braskamp, L. A., Brown, R. D., & Newman, D. L. (1978). The credibility of a local educational program evaluation report: Author, source, and client audience characteristics. *American Educational Research Journal, 15*(3), 441–450.

Brickell, H. M. (1976). The influence of external political factors on the role and methodology of evaluation. *Evaluation Comment, 5*(2), 1–6.

Brown, R. D., Braskamp, L. A., & Newman, D. L. (1978). Evaluator credibility as a function of report style: Do jargon and data make a difference? *Evaluation Quarterly, 2*(2), 331–341.

Brown, R. D., & Newman, D. L. (1982). An investigation of the effect of different data presentation formats and order of arguments in a simulated adversary evaluation. *Educational Evaluation and Policy Analysis, 4*(2), 197–203.

Brown, R. D., Newman, D. L., & Rivers, L. S. (1980). Perceived need for evaluation and data usage as influences on an evaluation's impact. *Educational Evaluation and Policy Analysis, 2*(5), 67–72.

Brown, R. D., Newman, D.L., & Rivers, L. S. (1985). Does the superintendent's opinion affect school boards' evaluation information needs? An empirical investigation. *Urban Education, 20*(2), 204–221.

Caplan, N. (1976). Social research and national policy: What gets used by whom, for what purposes, and with what effects? *International Social Science Journal, 28*(1), 187–194.

Carter, R. K. (1971). Client's resistance to negative findings and the latent conservative function of evaluation studies. *American Sociologist, 6,* 118–124.

Connor, R. F. (1981). Measuring evaluation utilization: A critique of different techniques. In J. A. Ciarlo (Ed.), *Utilizing evaluation: Concepts and measurement techniques* (pp. 59–77). Beverly Hills, CA: Sage.

Cooley, W. W. (1983). Improving the performance of educational systems. *Educational Researcher, 12*(6), 4–12.

Daillak, R. H. (1983). Evaluators in an urban school district: Organizational constraints upon evaluation influence. *Studies in Educational Evaluation, 9,* 33–46.

David, J. L. (1978, July). *Local uses of Title I evaluations.* Washington, DC: Department of Health, Education, and Welfare, Office of Education. (ERIC Document Reproduction Service No. ED 187–727)

Dawson, F. A., & D'Amico, J. J. (1985). Involving program staff in evaluation studies: A strategy for increasing information use and enriching the data base. *Evaluation Review, 9*(2), 173–188.

Dickman, F. B. (1981). Work activities, settings, methodologies and perceptions: Correlates of evaluative research utilization. *Knowledge: Creation, Diffusion, Utilization, 2*(3), 375–387.

Dickey, B. (1980). Utilization of evaluations of small scale innovative educational projects. *Educational Evaluation and Policy Analysis, 2*(1), 65–77.

Fetler, M. (1982, April). *Use of evaluation data and school achievement.* Paper presented at the meeting of the American Educational Research Association, New York. (ERIC Document Reproduction Service No. ED 218–343)

Florio, D. H., Behrmann, M. M., & Goltz, D. L. (1979). What do policy makers think of educational research and evaluation? Or do they? *Educational Evaluation and Policy Analysis, 1*(6), 61–87.

Glaser, B., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research.* Chicago: Aldine.

Glasman, N. S. (1984). Student achievement and the school principal. *Educational Evaluation and Policy Analysis, 6*(3), 283–296.

Goldberg, M. (1978, April). *New York City administrators' perceptions of the helpfulness of outside mandated evaluations on administrative decision-making in ESEA programs.* Paper presented at the meeting of the American Educational Research Association, Toronto, Ontario. (ERIC Document Reproduction Service No. ED 156–720)

Granville, A. C. (1977, April). *An experimental study of factors that influence school decisions: Implications for evaluation.* Paper presented at the meeting of the American Educational Research Association, New York. (ERIC Document Reproduction Service No. ED 139–824)

Heldt, V. W., Braskamp, L. A., & Filbeck, R. (1973). Effects of 'nonrational' influences on decision making in higher education: A simulation test. *Research in Higher Education, 1,* 163–172

Herman, J. L., & Dorr-Bremme, D. W. (1983). Uses of testing in the schools: A national profile. In W. E. Hathaway (Ed.), *Testing in the schools: New directions in testing and measurement*, No. 19 (pp. 7–17). San Francisco: Jossey-Bass.

Johnson, K. W. (1980). Stimulating evaluation use by integrating academia and practice. *Knowledge: Creation, Diffusion, Utilization, 2*(2), 237–262.

Jordan, K. F. (1977). Program improvement through school evaluation. *Educational Leadership, 34*(3), 272–275.

Kennedy, M. M. (1983). Working knowledge. *Knowledge: Creation, Diffusion, Utilization, 5*(2), 193–211.

Kennedy, M. M. (1984). How evidence alters understanding and decisions. *Educational Evaluation and Policy Analysis, 6*(3), 207–226.

King, J. A., & Pechman, E. M. (1984). Pinning the wave to the shore: Conceptualizing evaluation use in school systems. *Educational Evaluation and Policy Analysis, 6*(3), 241–253.

King, J. A., Thompson, B., & Pechman, E. M. (1980). *Evaluation utilization: A bibliography.* (ERIC Document Reproduction Service No. ED 207–984)

Leithwood, K. A., Wilson, R., & Marshall, A. R. (1981). Increasing the influence of evaluation studies on program decision making. In A. Lewy & D. Nevo (Eds.), *Evaluation roles in education* (pp. 357–383). New York: Gordon & Breach.

Leviton, L. C., & Boruch, R. F. (1983). Contributions of evaluation to education programs and policy. *Evaluation Review, 7*(5), 563–598.

Leviton, L. C., & Hughes, E. F. X. (1981). Research on the utilization of evaluations: A review and synthesis. *Evaluation Review, 5*(4), 525–548.

Lindblom, C. E., & Cohen, D. K. (1979). *Usable knowledge.* New Haven, CT: Yale University Press.

Lorenzen, G. L., & Braskamp, L. A. (1978). Comparative influence of political, cost/benefit, and statistical information on administrative decision making. *Evaluation and Program Planning, 1,* 235–238.

Maher, C. A. (1982). Utilization of program evaluation information in public schools: Perspectives and guidelines for school psychologists. *Journal of School Psychology, 20*(2), 113–121.

McGowan, E. F. (1976). Rational fantasies. *Policy Sciences, 7,* 439–454.

Neigher, W. D. (1979). Effort and influence: Enhancing the impact of mental health program evaluation at the agency level. In R. F. Rich (Ed.), *Translating evaluation into policy* (pp. 125–146). Beverly Hills, CA: Sage.

Newman, D. L., Brown, R. D., & Littman, M. I. (1979). Evaluator, report, and audience characteristics which influence the impact of evaluation reports: Does who says what to whom make a difference? *CEDR Quarterly, 12*(2), 14–18.

Newman, D. L., Brown, R. D., & Rivers, L. S. (1983). Locus of control and evaluation use: Does sense of control affect information needs and decision making? *Studies in Educational Evaluation, 9,* 77–88.

Newman, D. L., Brown, R. D., Rivers, L. S., & Glock, R. F. (1983). School boards' and administrators' use of evaluation information: Influencing factors. *Evaluation Review, 7*(1), 110–125.

Ory, J. C., & Braskamp, L. A. (1980). *Faculty perceptions of the quality and usefulness of three types of evaluative information.* (ERIC Document Reproduction Service No. ED 199–296)

Osterlind, S. J. (1979). What makes school evaluation work? *Educational Evaluation and Policy Analysis, 1*(1), 48–51.

Patton, M. Q., Grimes, P. S., Guthrie, K. M., Brennan, N. J., French, B. D., & Blyth, D. A. (1977). In search of impact: An analysis of the utilization of federal health evaluation research. In C. H. Weiss (Ed.), *Using social research in public policy making* (pp. 141–163). New York: D. C. Heath.

Pauley, P. A., & Cohen, S. (1984). Facilitating data-based decision making: Managers' use of data in a community mental health center. *Evaluation Review, 8*(2), 205–224.

Rich, R. F. (1979). Problem solving and evaluation research: Unemployment insurance policy. In R. F. Rich (Ed.), *Translating evaluation into policy* (pp. 87–110). Beverly Hills, CA: Sage.

Rossman, B. B., Hober, D. J., & Ciarlo, J. A. (1979). Awareness, use, and consequences of evaluation data in a community mental health center. *Community Mental Health Journal, 15,* 7–16.

Salmon-Cox, L. (1981). Teachers and standardized achievement tests: What's really happening? *Phi Delta Kappan, 62*(9), 631–634.

Siegal, K., & Tuckel, P. (1985). The utilization of evaluation research: A case analysis. *Evaluation Review, 9*(3), 307–328.

Sproull, L., & Zubrow, D. (1981). Standardized testing from the administrative perspective. *Phi Delta Kappan, 62*(9), 628–631.

Tash, W. R., & Stahler, G. (1982). Enhancing the utilization of evaluation findings. *Community Mental Health Journal, 18*(3), 180–189.

Thompson, P. A., Brown, R. D., & Furgason, J. (1981). Jargon and data do make a difference: The impact of report styles on lay and professional evaluation audiences. *Evaluation Review, 5*(2), 269–279.

Van de Vall, M., & Bolas, C. (1982). Using social policy research for reducing social problems: An empirical analysis of structure and functions. *Journal of Applied Behavioural Science, 18*(1), 49–67.

Webber, C. F. (1984). *School board member perceptions of the utility and importance of student evaluation information in Alberta.* Unpublished doctoral dissertation, University of Alberta, Edmonton.

Weeks, E. (1979). The managerial use of evaluation findings. In H. C. Schulberg & J. M. Jerrell (Eds.), *The evaluator and management* (pp. 137–155). Beverly Hills, CA: Sage.

Weiss, C. H. (1972). Utilization of evaluation: Toward comparative study. In C. H. Weiss (Ed.), *Evaluating action programs: Readings in social action and education* (pp. 318–326). Boston, MA: Allyn & Bacon.

Wholey, J. S., & White, B. F. (1973). Evaluations impact on Title I elementary and secondary education program management. *Evaluation, 1*(3), 73–76.

Williams, R. C., & Bank, A. (1984). Assessing instructional information systems in two districts: The search for impact. *Educational Evaluation and Policy Analysis, 6*(3), 267–282.

Windle, C., Majchrzak, A., & Flaherty, E. W. (1979). Program evaluation at the interface of program echelons. In R. F. Rich (Ed.), *Translating evaluation into policy* (pp. 45–58). Beverly Hills, CA: Sage.

Yeh, J. (1980). *A reanalysis of test use data* (CSE Report No. 143). Los Angeles: University of California, Center for the Study of Evaluation.

AUTHORS

J. BRADLEY COUSINS, Evaluation Consultant and Doctoral Candidate, Department of Measurement, Evaluation, and Computer Applications, Ontario Institute for Studies in Education, 252 Bloor Street West, Toronto, Ontario M5W 1V6. *Specializations*: Evaluation utilization, program and personnel evaluation.

KENNETH A. LEITHWOOD, Professor, Educational Administrator, and Head, Centre for Principal Development, Ontario Institute for Studies in Education, 252 Bloor Street West, Toronto, Ontario M5W 1V6. *Specializations*: Planned change, role of principal in school improvement, program and personnel evaluation.