Perceptions of evaluation held by principals, superintendents, and school board members were investigated. The study's instrument, a brief survey, was sent to a nationwide random sample of administrators. Dependent variables focused on the following general areas of concern: the users' perceptions of evaluation, including how useful they find its results, its most important function for them, etc., and the users' perceptions of their interactions with evaluators, including the frequency of contact, the most common types of evaluators, and the users' preferred method of receiving evaluation information. For the purposes of the survey, program evaluation was defined as "the process of providing information about programs to administrators or school board members to help them make decisions regarding the programs." Tentative conclusions based on responses to the questionnaire included: (1) a majority of users find the evaluations of educational programs in their school system useful; (2) roughly half of the average decisions made are determined by evaluation information; (3) only 28 percent of respondents reported that the program effects they most care about can be directly measured; and (4) few users speak frequently with program evaluators about the programs they are studying. (Author/RL)
A Nationwide Survey
of Administrators' Perceptions of Evaluation

Jean A. King
Tulane University

Bruce Thompson
University of New Orleans

April, 1981

Prepared for the Annual Meeting of the American Educational Research Association, Los Angeles, California, April 13-17, 1981. The survey discussed here was conducted under a contract with the National Institute of Education. However, the opinions expressed herein do not necessarily reflect the position of that agency, and no official endorsement should be inferred. The authors wish to thank Dr. Ellen Pechman for her invaluable aid in conducting this study.
ABSTRACT

The study investigated perceptions of evaluation held by principals, superintendents, and school board members. The study's instrument, a brief survey, was sent to a nationwide random sample of administrators. Among other things, the results indicate that administrators consider evaluation to be useful, but believe that many program impacts can not be directly measured, and that administrators have limited personal contact with evaluators regarding program evaluation.

RELATED ERIC DOCUMENTS:


In a happier time not any years past, there emerged from educational writings after intense labor the Evaluator. In this golden era, the world vibrated confusedly with unexamined programs of a variety of colors and shapes, administrators, legislators, and even the public turned expectant eyes toward the mighty Evaluator, hopeful that at last they would learn what was truly good.

Heavily burdened with the arms of the social scientist and greatly challenged by so overwhelming a need, the Evaluator hurled himself into the task undaunted, for his heart beat with the native purity of a rational being, and he knew that, at last, logic, rationality, and the social science paradigm were to have their day. And so he set about his evaluation tasks happily, taking on all comers, CIPPing and MOing with ease, using goals, not using goals, being responsive or adversarial or whatever was necessary to provide the often elusive Decision Maker with knowledge. Soon evaluation shops opened up across the country, and, as business boomed, few programs escaped the menacing prod of the Evaluator's yardstick. Schools were set up to train more evaluators; wise men spent days and weeks building the better evaluation model; and everyone knew that something good must be happening because, in an age of energy shortages, so many evaluators were exerting so much energy in so many directions.

Then one day a wizened Academician, greatly impressed by the bloom and buzz of the Evaluator's productions, whispered to a youthful Schoolman—new to the trade and known to be an occasional decision maker—that the Evaluator, by now somewhat
haggard from lack of sleep, but smiling still as if from divine inspiration, bore a striking resemblance to an inspirational character from an ancient book. The Schoolman, somehow missing the Messianic character of the Evaluator, was taken aback only briefly before commenting that, though he had seen no windmills in recent days, he too was reminded of the unhandy Don Quixote.\textsuperscript{1} The Academician gasped audibly at such blatant heresy and set out promptly to prove the young Schoolman wrong.

But everywhere he went, to his dismay he found only dusty volumes of untouched results, rented statisticians attacking others' methodology, and real-world decision makers muddling through with as little—or as much—rationality as they always had, a rationality that resembled his, but was not entirely identical. Slowly, inevitably, the Academician opened his eyes to the grim though obvious reality: despite the Evaluator's continued busy-ness, the world was not yet all that different. The evaluation millennium had not arrived, and most everyone was still in doubt as to what was, finally, truly good.

Like the Academician in the story, evaluators in recent years have been forced to acknowledge that their brand of rationality does not thrive in the real world. Despite the fact that more evaluations are happening now than ever before in history, school people have not found the results of evaluations as valuable to their functioning as we once would have hoped. Pressured decision makers will not necessarily use information evaluations provide them, even when it would be highly appropriate for them to do so,\textsuperscript{2} and even when they do use
information, they may not use it as earlier belief would have suggested, in the instrumental, go/no go fashion we formerly expected.

Fortunately, however, the study of evaluation use has now come of age. The work of Patton et al. (1978) and Alkin, Daillak, and White (1979) has suggested an altered definition of use, one more realistic and suited to the highly politicized settings of most evaluations. Weiss (1979) has outlined clearly the strengths and weaknesses of various approaches to evaluation use studies. These theoretical developments, coupled with current empirical work on evaluation use--case studies of the type reported in Alkin, Daillak, and White (1979) and the ongoing simulation research of Braskamp, Brown, and Newman (see, for example, Braskamp & Brown, 1980)--represent a major and positive response to our disappointment that evaluation's golden age (if it ever existed) is today over.

The survey described in this paper marks an additional response in that it was designed first to provide information about people in one evaluation setting, the people who will ultimately use the information evaluators provide them, i.e. the LEA superintendents, principals, and school board members who in a very real sense control the effects of evaluations in public school districts. The second purpose of this survey was to further a discussion of the theory of evaluation use. The information collected, combined with the growing knowledge of factors affecting use, may help lead finally to improved practice in the public schools and to a more complete understanding of the
uses of evaluation.

What Is Known About LEA Evaluation Users

The literature to date has not by and large focused on LEA users. Two studies have examined the characteristics and practice of certain school district evaluation units. Lyon et al.'s Evaluation and School Districts (1978) studied the evaluation units of public school districts having 10,000 or more students, but mentioned administrative and board users only in passing. Similarly, Webster and Stufflebeam's 1978 address centered on the types of studies used in evaluation offices of large urban districts, rather than on the local users of their products. In addition, Caulley and Smith (1978) surveyed evaluation users in several state education agencies, but, again, did not examine local users in depth.

Little has been documented about the evaluation perceptions of administrators or board members nationwide. Those few studies dealing more directly with LEA users have tended to be relatively (and in an important sense necessarily) localized. David (1978), for example, studied the local uses of Title I evaluations in 15 Title I districts across six states, including as sources administrators, teachers, and parents. What she found was that although evaluations were used to meet state and federal reporting requirements, to provide some forms of feedback, and as a rough indicator of program effectiveness,

Title I evaluations do not seem to serve, as primary purposes, either as a basis on which to judge the program or as a guide to program improvement (p.19, emphasis in original).
She attributes this lack of use to evaluation's bad reputation and overtones of accountability for local users and also to their sense that evaluation results are irrelevant to program decisions—rather serious, but revealing news for evaluators. The results of her study, then, suggest the critical importance of examining local users in varied programs across the country.

Like David's study, the seminal case studies presented in Alkin, Daillak, and White (1979) do not provide information about local users across the country, but they suggest key variables for the study of evaluation use in LEA's. Their final list of categories for an analytic framework of evaluation use includes the following: pre-existing evaluation bounds; orientation of the users; evaluator's approach; evaluator credibility; organizational factors; extraorganizational factors; information content and reporting; and administrator style (p.235). Awareness of how these categories look and work out in LEA's would be one way to determine strategies for improving use.

Finally, a CSE project entitled "Evaluation Design: An Organizational Study" is examining districts having a "reputation for linking evaluation or testing with instruction" (Williams and Bank, p.7). Such a study again must examine fewer districts in depth, leaving the general sense of the LEA evaluation information users unknown.

The lack of information on LEA users marks a major gap in the literature of educational evaluation use, a gap that the present study sought in part to fill. As a revision of the old saw would tell us, evaluation use begins at home; and, if the use of
evaluation information does not improve at the local level, it seems unlikely that it can improve at all.

The Questionnaire and National Survey

The questions used in the survey were developed after an extensive review of the literature (see Thompson & King, 1981). The two independent variables were district size (small, enrollment less than 3000; medium, enrollment 3000-10,000; or large, enrollment over 10,000) and position in the district (principal, superintendent, or board president). The dependent variables focused on two general areas of concern: first, the users' perceptions of evaluation, including how useful they find its results, its most important function for them, and so forth; and second, the users' perceptions of their interactions with evaluators, including the frequency of contact, the most common types of evaluators, and the users' preferred method of receiving evaluation information.

In February, 1981, we mailed questionnaires to a random sample of public school district personnel in the following categories: 1005 to principals (roughly half elementary and half secondary); and 900 to superintendents and school board presidents (roughly 60% superintendents and 40% board presidents). Those surveyed were asked to complete the questionnaire and return it in the prestamped envelope we provided. The results discussed here are based on the questionnaires returned by early March (360 principals, 253 superintendents, and 31 board presidents) and are purely
descriptive. Our final analysis will be based on the additional responses we have since received, so these conclusions must be considered tentative.

An additional limiting comment must be made concerning the form the survey took. To insure a high rate of return from everbusy LEA users, the questionnaire consisted of a single sheet of short answer questions (multiple choice and rank ordering). Using this short and simple format represented a tradeoff; an open ended survey would no doubt have given richer results. We felt, however, that our purpose here was to cast a wide net in order to see what manner of fish we were catching. While our results are limited in this sense, they do provide suggestive data about three important groups of users in varied public school settings. The by-now standard "this is an area in need of further study" statement applies here, and we would hope these results would be of use in designing further studies.

The Good News and the Bad

The good news is that 60% of the LEA users surveyed—a clear majority—reported that they find the evaluations of educational programs in their school system either useful or very useful; only 14% reported that evaluations were minimally useful to them. Of the three types of information users selected, a larger percentage of principals marked that evaluations were minimally useful (20% compared to 8% for superintendents and 7% for board presidents). The biggest contrast among the three groups was evident in large districts where 79% of the
superintendents said they found evaluations either useful or very useful, compared to only 45% of the large district principals who marked either of those responses. It should also be noted that 37% of these principals marked that evaluations were only minimally useful to them. While many users thus find evaluations useful—and our format unfortunately did not allow them to explain in what ways—principals, and especially principals in large districts, may feel that the benefits of evaluations are not reaching down the organizational chart to their schools.

The second bit of good news we found is that, regardless of the district size, most users report that they do use the information received from program evaluations in their decision making. Overall, these LEA users reported that roughly half of the average decision they make (48%) is determined by evaluation information. Even allowing for the forced nature of such an item, this would still seem to provide further evidence for the alternate view of use given in Alkin, Daillak, and White (1979). Of considerable interest, of course, is the source of the other 50% of information determining decisions; the political context factor cited in Patton (1978) deserves detailed study.

These findings are especially interesting in light of two additional findings, the bad news to accompany the good. First, only 28% of the LEA users responding wrote that the program effects they most care about can be directly measured. Well over two thirds (72%) felt that these effects can only be measured indirectly (42%) or not at all (30%). The superintendents and principals reflected these general values with two exceptions;
only 20% of large district superintendents agreed that effects cannot be measured, whereas 37% of small district principals felt that professional judgment would be required. What may be contrasted here is the accountable atmosphere the large system superintendents live in with the small system principals' fears that what their schools are achieving may somehow slip through the cracks of an evaluation. These data support David's view (1981, p. 32) that "evaluation results are usually found lacking because important goals are not measured" (emphasis in original). If users' mind sets predetermine the non-effects of an evaluation, a resultant lack of effects is predictable.

Equally interesting (and equally speculative) is the finding that, of the small number of school board presidents responding, only 10% felt that program effects could not be measured either directly or indirectly; if this were shown to be true in a larger sample, then schools boards, feeling comfortable using the results of evaluations, may need to be educated to their administrators' concerns. On the other hand, administrators may need to be taught the benefits of what can be effectively measured in schools. In both cases, knowing these attitudes in advance may help evaluators deal more effectively with LEA clients.

The second dose of bad news should not be unexpected: few LEA users speak frequently with program evaluators about the programs they are studying. Only 2% reported daily contact, and the majority (57%) marked the least extreme category, i.e. they reported speaking to evaluators less than once every two weeks.
From talking with local school people, we feel confident that even this response is overly generous and that few LEA users have extensive contact with the people evaluating their programs. This supports Jane David's contention (1981, p.38) that for Title I programs "... in almost every district, there is little connection between program staff and evaluation staff." Lyon et al. (1978) reported similar findings in their survey of large district evaluation units. The problems inherent in this are self evident, if we assume that to know evaluators is to love evaluation—or at least to use its results more wisely.

Of the three types of users, the superintendents, as might be expected, are somewhat more regular than their colleagues at speaking with evaluators; only 48% of them report contact less than biweekly, compared to 63% for principals and 79% for board members. It must be noted that principals in smaller districts report somewhat more contact than do their peers in larger systems, 55% marking less than biweekly contact, compared to 64% in medium sized districts and 73% in large. Both role and size of district may affect the frequency of contact with evaluators. These contact results are not surprising, although the fact that these same users report that they find evaluations useful and that they do use their results in decision making suggests the power of the information that they are getting, if not in speaking with evaluators, then, perhaps, in working with written reports. The good news discussed here can make us hopeful; the bad news suggests areas in which to begin improving practice.
Additional Results and Implications for Theory

As noted above, the purpose of this survey was twofold: first, to provide data on LEA evaluation users; and second, to begin to shed light on elements of a theory of evaluation use. Three items relate directly to these elements, specifically addressing the question of the evaluator's characteristics, the reporting of evaluation results, and the modes of evaluation use.

One variable potentially affecting how LEA users respond to evaluators is the evaluator's approach, the particular combination of political and technical skills he or she presents when relating to users. Meltsner (1976) has suggested a fourfold typology of policy analysts in the bureaucracy: the technician, having high technical and low political skills; the politician, having the opposite; the entrepreneur, high in both skill areas; and the pretender, low in both. In an empirical study, Thompson (1980) found support for the presence of two of these types—the entrepreneur and the pretender—in the evaluation department of a large urban school district.

In our study, when asked to label the most common type of program evaluator in their systems, LEA users gave the following responses: entrepreneur, 49%; politician, 26%; technician, 17%; and pretender or placeholder, 8%. The importance of political skills in LEA settings is evident in that 75% of the respondents labeled their most common evaluators as high in political skills. The relatively equal importance of technical skills is reflected in the 66% rating for those categories with high technical skills. In contrast to Thompson's finding, these data do not
support the presence of placeholders in LEA's, although this may be because the variables were operationalized differently in the studies; additional work is needed to clarify this.

Response differences were present on this item for board presidents and for large districts. Two categories accounted for 87% of the evaluator types identified by the small number of board presidents who responded--60%, entrepreneur; and 27%, technician--perhaps reflecting the different perspective of the non-educator. Like the school board presidents, the perspective of large school districts' personnel may differ from that of administrators in other systems. The majority of both superintendents and principals in large districts identified the entrepreneur as their most common type of evaluator (62% and 57% respectively); both political and technical skills may be needed for evaluators to survive in these settings. Because it is relatively easy to manipulate, the evaluator's approach to evaluation deserves further attention.

A second variable meriting study is that of the reporting of evaluation information. We have documented elsewhere the importance of informal contacts between evaluators and LEA users (Thompson and King, 1981). To determine these users' feelings of the relative importance of such contacts, users were asked to rank the following factors important to the use of program evaluation in their school systems: evaluation reports which identify deficiencies and discuss possible actions to correct them; evaluation personnel who meet with administrators to explain the reports; and administrators who recognize the merits
in an evaluation report and push for responsive action. The first factor is self evident; evaluating programs and writing reports go hand in hand and always have. The second factor reflects the reported values of repeated and informal contacts between program personnel and evaluators. The first factor refers to what Patton et al. (1978) label the personal factor, and the literature suggests its importance.

Although a disappointing number of respondents failed to answer the item, those who did (27%) provided somewhat surprising results. Superintendents (N=68) agreed quite strongly that the most important factor was the evaluation report (69% first rankings), followed by the concerned administrator (47% seconds), followed by interactions with evaluation personnel (50% thirds); this order was also the modal response (41%). The principals' responses (N=99), while not as entirely consistent, certainly showed the same pattern; the modal response (35%) was the same, and 60% rated the reports as most important while 49% rated interactions with personnel as least important.

These results suggest that LEA users do want reports and that they may be less aware of the importance both of the personal factor and of their contacts with evaluation personnel. The work of Braskamp, Brown, and Newman then takes on greater importance as detailing report variables that may influence readers' reactions. Holley's multi-media approach to evaluation reporting (e.g. Lee and Holley, 1978; Holley, 1979) becomes important as well because it not only enhances reports, but trains evaluation audiences to the value of contacts with the
evaluator. The notion of collaborative evaluation outlined in Gray (1979) has not yet spread to LEA's if the report remains the critical product of an evaluation in the eyes of the users.

A third variable of critical importance in the literature is the varying functions which program evaluation helps users accomplish, also called evaluation's modes of use (see Caplan et al., 1975; Knorr, 1977; Rich, 1977; Weiner, Rubin, & Sachse, 1977; Weiss, 1977; and Pelz, 1978). As the tale at the beginning of this paper suggests, we have rather humbly accepted the fact of late that the users of evaluations may not apply our results for immediate programmatic cures, i.e. that while the instrumental function (a go/no go decision) may be a sufficient characteristic of use, it is not a necessary one. Equally important are conceptual and symbolic modes (under various labels) by which, in the first case, evaluation results influence users' current thinking about a program or issue, and, in the second, where evaluation results are used to achieve users' goals (for example, garnering political support or substituting for a decision).

To determine indirectly which modes were most important to LEA users, we asked them to rank which of the following functions evaluation most helps them perform: identifying problems in program implementation so the problems can be fixed before they become serious; determining the basic problems confronting the school system as a whole; and persuading people to make program changes that need to be made. The first is one way the instrumental mode works out in practice; the second represents
the conceptual mode; and the third, the symbolic. This item severely limited the users' responses, and some may object to the choice we forced users to make. We, however, felt that the three activities were basic to evaluation as it takes place in LEA's and that the results would suggest enough questions to merit this risk.

Again, and probably owing to the nature of the item, the number of respondents who answered was disappointing (28%). The results, however, are suggestive. Of the three modes given, the conceptual (determining the system's basic problems) received the highest number of first rankings (42%), followed by the instrumental (identifying implementation problems) which received 34% firsts, and the symbolic (persuading people to make changes) which received 25%. Thirty-six per cent of the superintendents and 46% of the principals placed the conceptual mode first. The symbolic function received the most low rankings (43%), while the combined first and second rankings for the conceptual and the instrumental were 74% and 73% respectively. These numbers may provide support for the alternative notion of evaluation use discussed above; LEA users are interested not only in identifying problems in implementation, but in discussing the more general problems facing their districts. Researchers seeking result-stimulus and action-response are ignoring an important function that evaluation performs for local users. Additional research on the conceptual use of evaluation at the LEA level is needed.
Final Thoughts

If evaluators can no longer play the part of the savior, neither do they have to carry the lance of a Don Quixote. The initial data collected in this survey make it clear that LEA evaluation users--superintendents, principals, and school board presidents--find evaluations of value and do, at least in part, use evaluation information to make decisions. On the other hand, these data provide areas for future research and speculation. Over two thirds of the users surveyed felt that the program effects they most cared about could not be directly measured, while a majority reported that they had infrequent contact with evaluators. The variables of evaluator characteristics, reporting of results, and modes of use represent important starting places for additional study.

The importance of studying evaluation use in LEA's cannot be overestimated. While evaluation and policy studies at the national level are needed and may provide critical insights for local evaluation use, careful examination of the use, misuse, and non-use of evaluation information at the district and even the building level is also needed to help us come closer to the day when evaluations will more effectively improve programs for students. This survey marks a first step in that direction.
Endnotes

1. A thank you to Robert Wise (1978) for suggesting the Don Quixote image.

2. Michael Scriven (1981) would warn potential users to first make sure the evaluation results are worth using. This caveat must be remembered throughout this discussion.

3. Jerrilyn Andrews, Alan Guma, Jason Millman, and Ellen Pechman assisted us greatly by suggesting revisions in the first draft of the questionnaire.

4. For the purposes of the survey, program evaluation was defined as "the process of providing information about programs (e.g., Title I, a reading curriculum, etc.) to administrators or school board members to help them make decisions regarding the programs."

5. The percentages reported here have been rounded to the nearest point.

6. The word pretender has negative connotations which are unfortunate. As Thompson (1960) points out, pretenders may serve an important function in the social order of which they are part. We shall use the term placeholder rather than pretender.
Bibliography


6. Gray, P.J. The development and testing of a collaborative approach to education program evaluation (Doctoral dissertation, University of Oregon, 1979). (University Microfilms No. 80-5, 767)


