The Canadian Journal of Program Evaluation Vol. 9 No. 2 Pages 131–149
ISSN 0834-1516

ALL PULLING TOGETHER: WORKING TOWARD A SUCCESSFUL EVALUATION

Alan G. Ryan
University of Saskatchewan
Saskatoon, Saskatchewan

Caroline Krentz
University of Regina
Regina, Saskatchewan

Abstract: Utilization of results is an important aspect of program evaluation. Naturalistic methods aim at increasing the utilization of research findings by providing stakeholders with a rich understanding of the situation being studied. In this article, we report on a study that caused us to examine the assumption of the power of naturalistic methods. We found that one key to utilization is to narrow the gap between the type of information provided by the study and the expectations of the stakeholders.

Résumé: L'utilisation des résultats est un aspect important de l'évaluation de programmes. Les méthodes naturalistes visent à augmenter l'utilisation des résultats des recherches en donnant aux clients une compréhension profonde de la situation étudiée. Dans cet article il s'agit d'une étude qui met en question cette hypothèse de la puissance des méthodes naturalistes. Nous avons constaté que la clé de l'utilisation est de réduire la lacune entre le genre d'information produite par l'étude et les attentes des clients.

Amidst continued reports of worthwhile program evaluations failing to bring about action, many suggestions have been offered whereby clients can utilize more of their expensive and hard-won program evaluation findings to improve their programs. Indeed, it is no exaggeration to say that the issue of the utilization of evaluation findings has been a preoccupation of program evaluators since the beginning of the modern-day profession (Weiss, 1972). Whole books have been written on the topic (Alkin, Daillak, & White, 1979; Patton, 1986), and high-profile debates have taken place at gatherings of evaluators (Patton, 1988; Weiss, 1988a, 1988b). In this article, we reflect on an evaluation study where the desire to enhance utilization of the findings caused us to examine anew our assumptions of how to conduct the evaluation. By way of introduction, some of the suggested approaches for enhancing utilization of program evaluation findings are categorized and briefly reviewed. The categories described below are inevitably broad, and many researchers concerned with utilization straddle several categories.

HOW MIGHT PROGRAM EVALUATION UTILIZATION BE ENHANCED?

Understand how decision-makers act on information. Theorists in the field of knowledge utilization have applied their analyses to the specific area of program evaluation utilization. Strong examples of this work are Kennedy's (1983, 1984) understanding of "working knowledge" or Weiss and Bucuvalas's (1980) examination of which characteristics of reports are most likely to cause administrators to act. Cousins and Leithwood (1986) have reviewed a wide selection of studies that seek to understand how decision-makers first process incoming information and then decide whether to act on it.

Actively cater to the decision-maker's needs. Patton (1986) is perhaps the best-known exponent of this school of thought. As part of his philosophy of evaluation, he advocates that when program evaluators communicate with decision-makers, they pay a great deal of attention to the idiosyncratic needs and habits of those decision-makers. The hope is that in so doing, the evaluator will better ensure that the findings of the evaluation study are acted upon. He suggests that program evaluators should carefully target their efforts to promote the use of their findings; in this way, he feels, decision-makers will come to "own" the evaluation findings. Although many other program evaluators have endorsed this approach, some (for example, House, in Alkin, 1990) have worried that it unbalances the focus of the evaluation study in favor of the decision-maker and at the expense of other stakeholders in the program being evaluated.

Adopt naturalistic program evaluation strategies. The emergence of naturalistic ("qualitative") research methodologies in program evaluation studies parallels their emergence in educational research in general (Guba & Lincoln, 1989; Lincoln & Guba, 1985; Ryan, 1988). The rise of naturalistic evaluation strategies is due in part to the *zeitgeist* of social science research in general (a widespread, although not universal, sense of dissatisfaction with the positivist ideal of research), but also in part to the desire of evaluators to see their work bring about greater program improvement. The emphasis on mutualistic meaning-making that is the cornerstone of naturalistic evaluation should, in the eyes of naturalistic evaluators, lead to all stakeholders having an understanding of the manifold dimensions of the program being evaluated and the corresponding views and beliefs of all participants in it. Given this rich understanding, the will to change the program so that it becomes even better should, according to naturalistic evaluators, be a part of the reaction of the stakeholders to the findings of the evaluation. Thus, the effect of the naturalistic evaluation study is to encourage all parties to the evaluation to share the meaning they attach to the program and, in turn, to learn its meaning for other stakeholders, in the expectation that this shared understanding will result in a more meaningful resolution of differing perceptions and a shared understanding of the way ahead.

This attitude is more than a shift in technique; naturalistic evaluation also contains a moral dimension that is inherent in the view of research from which it springs (Schwandt, 1989, 1991). Thus, in the move to naturalistic evaluation, the power relationship between the evaluator and the participants changes from being one of expert-client to a much more mutualistic situation where the evaluator and the participants are trying to understand each other's reality (Ryan, 1987). This change of perspective is expected to cultivate a moral dimension in participants' attitudes toward each other; in essence, naturalistic methods should promote mutual acceptance of the legitimacy of differing perspectives among the different stakeholders within an organization. This, at least, is the hope. Recently, we undertook an evaluation study where these assumptions were tested. Although we fully subscribe to the philosophical basis of naturalistic research and believe in its potential for enhancing utilization, we also had to come to accept that in this instance, at least, utilization depends as much on a client's expectations for an evaluation study as it does on the methodological underpinnings of the study.

THE STUDY

Torrance is the name we give to the small prairie town where this study was conducted. The director of education called for a project that would reveal the perceptions that the members of the community (the town itself and some of the surrounding area that is served by the high school) held of the school system under his supervision. In the system, there were several elementary schools, two middle-years schools, and one large comprehensive high school. The purpose of the project was to provide a “community report card” to the director, his school boards (there was a separate school board for the comprehensive high school), and the steering committee that had been established to oversee the project. The school system had earlier embarked on a long-term planning and renewal initiative and had already conducted several studies to these ends; this project was to provide a picture of the community’s views on the educational system—a missing piece of the necessary data for renewal—so that the reorganization could take these views into account.

This account sounds reasonably lucid now, but as evaluators we found that initially we had to spend quite a bit of effort in establishing the exact purpose of the study. In part this was due to the normal negotiation of understanding that accompanies the beginning of any program evaluation study in which outsiders are involved, but it was also due to the fact that the director and the committee had not reached agreement on what exactly was desired from the study or on the methodology that was to be used. This article will explore some of the issues that arose at the negotiation phase and also during and after the conduct of the study itself. We begin with a description of how the study was carried out; during this section we foreshadow some of the issues that arose, treating these in more detail in a later section. We will not present any results of the study except where doing so can illuminate a methodological point that we are trying to make.

Design and Procedures of the Study

The six phases of data collection and analysis are described briefly below. Phase 1, an open-ended survey of the community of Torrance, was the cornerstone of the whole project; everything else either flowed from it or was in reaction to it. Phases 2 and 3 represent our attempt to satisfy the concerns of the committee members who were more comfortable within the classical, experimental paradigm. In Phases 4 and 5 we explored in more depth some specific issues arising from the first three phases. Phase 6 presents our compromise between our belief that naturalistic studies need not present traditional recommendations (Ryan, 1993) and the expectation of some committee members who felt that recommendations were an essential part of the study’s findings.

Phase 1: Mail Survey of Torrance Households

In early December 1991, a blanket mailing to all residences in Torrance (over 7,000) was undertaken. A further 158 surveys were sent to homes of those current students of the comprehensive high school who live outside the Torrance boundaries. Each mailing consisted of a survey, a covering letter explaining the purpose of the exercise, a business reply envelope, and a form to be filled in if the respondent wished to take part in the subsequent group discussions of the issues (phase 5). Participation in the project was encouraged by an advertisement in the Torrance newspaper and through public service announcements on the local radio stations.

Our efforts to determine the nature of the questions for the survey began with a dialogue with the steering committee about the naturalistic evaluation process. We opened our presentation by describing the beginning phase as an open-ended process designed to elicit community perceptions of the school program in a way that would capture these perceptions as faithfully as possible. Then, through their participation in the refinement of the questions and the accompanying letter of invitation, committee members began to formulate and share their views about the school program and the evaluation process. It was clear that among the committee there were differing perceptions of the program as well as uneven support for the naturalistic approach to evaluation. Nevertheless, the open-ended approach and survey questions received committee approval.

The resulting survey instrument was of an open-ended design, with three very broad questions being posed and a space being provided for the respondent to raise other issues. The three questions were: "How relevant and up-to-date are the courses that students take?", "How well are the courses being delivered in the Torrance public and comprehensive schools?", and "How well prepared are school leavers to enter the workforce or go on to higher education?" This open-ended style is not a common format for wide-distribution surveys; it was adopted, in preference to the usual rating scale response to a list of supplied items, in order to elicit respondents' true perceptions, unchannelled by any prior framing by the survey design. In addition, the open-ended format reduces the amount of interpretation that needs to be assumed during the analysis; a paragraph written to describe one's feelings on, say, how well our students are prepared for the workforce is much more revealing than a checkmark against "well prepared." In the latter situation, the evaluator has no way of knowing what was in the responder's mind that made the choice appropriate. We received 192 replies (2.3%), of which 160 provided usable information.

Analysis of Phase 1 Returns

Each return was given a number and then read. The written comments were coded into categories, question by question (Lincoln & Guba, 1985). Most answers tended to include a general statement together with some elaborating specifics. For example, one respondent replied to the question on how well the courses are being delivered by writing, "It appears to be adequate. There is always room for more one to one between teacher and student. The teachers seem to be rushed to complete all of the course in the given amount of time." The first sentence was categorized under the code "Adequate," the second under "Some learners need more personalized help," and the third under "Teachers are rushed to complete course in allotted time." Ultimately, between 50 and 70 categories were needed for each of the three questions and the open-ended section. The first author carried out this analysis; the second author duplicated the coding for a randomly selected number of response sheets in order to ensure that the categories were not merely artifacts of the categorizer's imagination. After the coding had been carried out for all responses, the categories were collected under *issues*, those topics that seemed to be the most pervasive and of potential interest to the committee. Each issue was then written up, with reference to the original response sheets. Ten randomly selected responses were kept back, unread, to serve as a check on the final description of the issues. It was found that when these ten were analyzed, they could be coded according to the previously derived system, and their contents were indeed captured by the descriptions of the issues.

Phase 2: Likert-Type Questionnaire to Torrance Households

This data collection phase was included to serve as a check on the completeness and validity of the data collected and analyzed in phase 1. Every 20th residential mailing was of a five-point, 30-item Likert-type questionnaire instead of the open-ended survey instrument described above. The items were derived from discussions between ourselves and some committee members on the likely responses that would be provided to the three open-ended questions of phase 1. The director also asked for the inclusion of a couple of items of specific interest to him; for example, an item on reporting procedures was added at his request. In addition, some opposites were included in order to check for the absence of response patterns. Finally, respondents were encouraged to write additional comments on their returns. Of the 395 questionnaires delivered to households, 41 were returned, slightly over 10%, of which 37 contained analyzable data.

Analysis of Phase 2 Data

The analysis of the written comments revealed similar patterns to those found in phase 1. Because of the difficulties of making inferences from data generated by people checking off their responses, interpretation of the results was deliberately kept fairly tentative. Our report concluded, "the Committee is probably justified in imputing the same overall feelings to the questionnaire respondents as were revealed by the written responses of the survey respondents."

Comparison of Phase 1 and Phase 2 Results

The analysis of the phase 1 open-ended survey gave rise to some 20 issues. We felt that the phase 2 questionnaire did not produce any additional issues to add to this list, with the "seeded" item on reporting procedures a possible exception. The responses suggested that some respondents were at least mildly dissatisfied with reporting procedures. The issue of reporting procedures did not arise as an issue in the open-ended survey (phase 1). It did arise, however, as an issue in some of the community group meetings (phase 5, below).

On the other hand, we were struck by how much richer were the data from the open-ended survey. We wrote in the final report that we felt the survey method had been vindicated as a way of gaining a meaningful understanding of the perceptions of the residents of Torrance.

Phase 3: Random Telephone Survey

A telephone survey was conducted in order to complement the coverage provided by the main mail survey. The mail survey was designed to serve as a forum for residents of Torrance who wished to give their views on the public and comprehensive school systems; it was never intended to be a statistically generalizable survey. Even so, it was felt that it would be reassuring to have a general feel for how representative the responses to the mail survey were; the telephone survey was designed to shed some light on this.

The telephone survey was administered to 31 randomly selected Torrance households. Because of the limitations of the telephone medium, the questions had to be less open ended than those of the mail survey. They were derived from the three basic questions of the survey, and expanded upon with additional probes designed to reflect some of the major issues that had emerged from the analysis of the mail surveys. The responses were analyzed using the same categories as had been established from the phase 1 mail survey. The categories proved adequate to the task, as very few additional points were raised.

Telephone interviews tend to elicit slightly more favorable responses to questions than do written surveys (Sudman & Bradburn, 1982). This is a result of social compliance that arises from the personal interaction that takes place over the telephone between the interviewer and the respondent. It is impossible to quantify this tendency. Certainly, it held true for the responses to this survey; indeed, so marked was the effect that we felt that the degree of approval was *truly* higher among those we telephoned than among those who wrote in. If so, then this is evidence that some residents who chose not to answer the written survey did so because they were content with the schools.

The telephone survey allowed us to gain some understanding into why the response rate to the mail survey was not higher. We had speculated that two reasons were likely: *(a)* people did not feel sufficiently knowledgeable about or interested in the school systems to make comments, and *(b)* people were satisfied with the systems as they were and so did not feel driven to reply to the survey. Both of these reasons received support through comments of participants in the telephone survey.

Phase 4: Student Interviews

Data from the first three phases informed the next step in the evaluation process. The second author suggested to the steering committee that students from all levels—elementary, middle years, and high school—be interviewed to ascertain their perceptions of the school program. This suggestion was well received by the committee, and members contributed approaches and questions they thought would stimulate student response.

Elementary and middle years students. Twenty-one children in kindergarten through Grade 5 and 18 children in Grades 6 through 8 were interviewed in an elementary and a middle years school, respectively. At the elementary school, the students took part in a discussion and then drew pictures illustrating their responses to two unfinished statements: "At school I like . . ." and "At school I wish . . ." The principal was briefed on the points raised by the children.

Each of the groups of the six students per grade in the middle years school was asked to jot down their responses to "Things I like about school" and "Things I would like to change." These were then recorded on chart paper and discussed in a group. After the interview sessions, we shared the chart paper sheets with the principal so that she could learn what the students had said.

High school students. In March 1992, 21 students from all four grades of the comprehensive high school met to discuss aspects of their school. We met in the library seminar room of the comprehensive, enjoying a pizza lunch before an hour-long discussion. The principal had arranged for the selection of the students. Ahead of the meeting time, the students were given a response sheet on which they were asked to note their thoughts on two questions: "How can the Comprehensive High School *best prepare* you for your future?" and "How can the Comprehensive High School improve the *quality* of your education?" Even though our meeting was held on the first day back to school after the February break, most students remembered to bring their sheets. We used a somewhat more sophisticated discussion format than the one designed for the elementary and middle years students. We stressed that no individuals or specific incidents should be mentioned, and assured the students that no information would be attributed back to individuals.

To respond to the first question, students were asked to identify ways in which the school *helps prepare, or could help prepare,* them for the world after school. Every point offered was recorded on sheets of chart paper; the version on paper was, thanks to the group discussion, often more focused than the original student statement.

Students were then asked to identify *barriers* that prevent the school from helping them prepare for the future. These were also recorded on chart paper. All four grades were represented in the discussion groups; we did not separate out the comments by grade. A similar procedure was used to elicit student responses to the second question, about improving the quality of education.

Phase 5: Community Group Meetings

In preparation for the community meetings, the second author developed a process to assist the steering committee in selecting priority issues for the discussion. By this time, the first author had prepared a draft report that summarized the issues raised in the mail (phase 1) and telephone (phase 3) surveys, and the questionnaire (phase 2). These issues formed the basis of a consensus-seeking activity whereby committee members ranked and regrouped issues in terms of which ones would most repay additional examination from community members.

Earlier, committee members had deliberated over an appropriate discussion method for the community meetings. We were sensitive to the need to encourage dialogue about the issues while at the same time allowing all viewpoints a hearing. A force-field analysis approach was selected because it could accommodate insights about the existing program and those forces that were encouraging or discouraging change in Torrance's school system (Curriculum Branch, 1985; Sanders, 1977).

Thirty-five community members met on two evenings in the comprehensive high school. Participants comprised those who had indicated in their returned survey or questionnaire that they would like to meet for discussions, and members of the citizens advisory committee who had accepted our invitation. During the two-hour sessions, participants, divided into two groups, discussed two questions. All groups spent the first half of each meeting considering what the role of the school should be. For the second half of each meeting, one group considered the problem statement "The school system needs to improve teaching and learning" and the other considered "The school system must balance change and constraint."

Recording

For the question on the role of the school, participants noted their points on paper and then shared them, one point per person at a time, with the group. These were recorded on chart paper. The ideas on the chart paper were elaborated upon as the discussion proceeded.

For the second and third questions, participants were given a sheet of paper divided down the middle. One column was headed with a "Barrier" statement ("Barriers that prevent improvement in teaching and learning"; "Barriers that prevent a balance"), and the other column was headed with a "Forces" for statement ("Forces that encourage change in teaching and learning"; "Forces that encourage a balance"). Once again, members of the group shared their ideas, and these were recorded on chart paper. Each participant was then asked to consider the list of barriers on the chart paper and select *(a)* the barrier he or she most wanted to change, *(b)* the barrier that would be easiest to change, and *(c)* the barrier that would be most difficult to change.

Analysis

The information from all four groups was combined. The answers to the first question were interpreted in light of the coding system and the findings of the phase 1 Torrance survey. We chose this approach rather than carry out a separate analysis because we felt *(a)* the survey was the most representative presentation of Torrance community views on the schools and so should stand as the reference marker for the project, and *(b)* most participants had already contributed their views through their completion of the survey

Phase 6: Presentation of Pressure Points

The wealth of information obtained in the five phases of data collection was written up as a comprehensive case study. We had intended to leave the case study to stand on its own, and to allow the committee members to derive their own meaning from it and chart their next actions. This proposed course of action met with the approval of those members who were comfortable with entering into a situation via the reading of a case study report. However, other members of the committee very much wanted some recommendations from the "experts."

Accordingly, we compromised by identifying 23 of what we called "pressure points" in the system. Pressure points differ from recommendations in that they emphasized the data sources from which they came and played down the "expert evaluator" role in their development. We grouped them into three main categories: teaching, program, and administration. For each of the pressure points, we summarized in a one-page format the following:

- what the pressure point was,
- from which of the data sources it had emerged,
- what its context was,
- what the core problem causing the pressure point was, and
- some suggestions about what might be done to alleviate it.

An example of the format is shown as an appendix to this paper.

Withdrawal

In a very traditional way, our involvement with the project ended with the presentation of our report. As his next step, the director intended to incorporate our findings into the previously completed long-term planning and renewal exercises. He then anticipated employing a professional facilitator to guide the school system through a comprehensive renewal process. Our withdrawal was less of a wrench than is usual at this point in an evaluation study. We had involved the committee at every stage of the study, and they had had the report to study for some time before the final meeting. From our conversations with committee members, we felt that their involvement, together with the richness of the data presented in the case study, had made them ready to proceed on their own. We were not needed any more.

The school boards did subsequently employ a professional facilitator to help resolve some problem areas in the school system (although the large-scale exercise envisioned by the director never came to pass). We have learned informally that our report, including the pressure points, was an important resource in that process.

WHAT WE HAVE LEARNED ABOUT USING NATURALISTIC EVALUATION

Our thoughts are collected under the headings of the stages of the study that best illustrate our thinking. The stages refer to the times during which we interacted with the committee.

Stage A: Negotiation of Design Stage

Relationship of our project to other studies. We first had to understand how this evaluation project related to other studies that had been carried out in the system. The school system had recently completed a comprehensive internal review that had set out a plan of action for the professionals within that system. At our initial meeting, there was some confusion among the director, the committee, and ourselves about why another study was needed so soon afterward and what status the results of our study would have vis-à-vis those of the previous comprehensive review. As discussions went along, we agreed that the community's perceptions of the educational system constituted the "missing link" that had not yet been addressed and that our study should therefore address.

Different agendas of committee members. The committee comprised the director, another central office administrator, consultants, board members, school administrators, two students (one from the high school and one from a middle years school), and community members. As the focus of the project became the perceptions of the community, the school administrators (in particular) became concerned lest the study become a soapbox for any members of the community who were upset with the school system or who had a narrow, specific point of view to foster. In particular, they were concerned that the project might degenerate into a "teacher-bashing" exercise. Other members of the committee felt that a project designed to uncover community perceptions should do just that, and wanted to let the chips fall where they may. In the end, we convinced committee members that our study would be sensitive to the need to balance these concerns—for example, we would not identify specific teachers who came in for particular opprobrium, but we would report on the community's perceptions of the quality of teaching, should that emerge as an issue—and the project proceeded.

Stage B: Presentation of the Design

We decided that we wished to use a naturalistic, sense-making approach to the task of uncovering the meaning that the community attached to its school system. Such an approach seemed to accommodate best the needs of the committee and the moral concerns of those members who were determined to protect their colleagues from personal attacks. Accordingly, we developed a virtually totally open-ended survey design that simply asked Torrance community members to tell us what they thought of the school systems. When we presented this plan to the committee, the reaction was bifurcated. Those members (and all of them were from the professional membership of the committee) who felt comfortable with research conducted within the naturalistic paradigm endorsed our proposal. Other committee members (all of the lay members and some of the professionals) were unconvinced by our suggestion for several reasons:

1. Asking community members to write their views would not result in a random sample of participants. Hence, we could not generalize our results to the whole population, nor could we apply statistical tests of significance.
2. Using an open-ended format would take control of the study away from the committee and would result in information that would not necessarily be of use to the committee.
3. If naturalistic methods result simply in a portrayal of the meaning the Torrance educational system has for community members and not a list of recommendations, what is the point of hiring evaluators? Surely the job of the evaluator is to tell the committee what to do to improve things.

Our reaction. First, the methodological issues proved amenable to discussion. We showed that randomization and tests of significance were not essential to the construction of a meaningful portrayal. We explained that inviting participants to contribute freely would result in the fullest disclosure of the community's concerns and that that information would be of most use to the committee, regardless of their preconceptions of the state of community satisfaction.

Second, to reassure the traditionally minded committee members, we added two further phases of data collection: the Likert-type 30-item questionnaire (phase 2) and the random telephone survey of nonrespondents (phase 3).

Third, we found that the role of the evaluator in naturalistic studies was more difficult to convey. The idea that our expertise should be used in the collection and synthesis of data but that the committee should take ownership by entering into the process as they interpreted that data and thus planned for the next steps was alien to many committee members. Most members agreed to suspend their skepticism and go along with the process that we had outlined and that the rest of the committee had accepted.

Stage C: Reporting the Results of the Survey and Questionnaires

The case study report from phase 1 was ready first. This was met with great enthusiasm by all committee members. They found that by reading the case study report (a substantial document of some 40 pages) they were indeed able to gain an understanding of what the community members were feeling about the school system. However, some members were still uneasy over the small numbers and nonrandomness of the responses.

When the results of the 30-item, Likert-type questionnaire (phase 2) were reported, some committee members immediately started to scrutinize the percentage responses and to prioritize the reported reactions on the basis of those numbers. We cautioned that the numbers were not necessarily any more meaningful or precise than the data in the case study report—indeed, our analysis of the responses to the 30-item questionnaire pointed out some inconsistencies in the way that respondents answered similar items—but even so, for some committee members, the numbers took on great significance.

The findings of the telephone survey (phase 3), which showed that the randomly selected interviewees seemed to be even more content than, or uninterested in, the schools than those who returned the survey and the questionnaire, mollified those committee members who were previously unconvinced of the representativeness of the case study report.

Stage D: Presentation of Results of Student and Community Interviews

In some ways, we feel that this stage was a pivotal one for the study. The findings from our work with the students (phase 4) and those community members who had signed up to meet with us (phase 5) contained some of the most severe criticisms of the school system thus far. Students can be very pointed in their comments. We awaited the reaction of the committee with slightly bated breath, but were reassured by their reaction. Gone were any methodological quibbles; absent were any reservations on the merits of the criticisms. Instead, the committee worked diligently to incorporate these latest findings into their developing understanding of the community's perceptions of the school system. These new data enhanced the picture that had been established by the case study report and reinforced by the questionnaire and the telephone survey. Specifically, we feel that we can say this about the committee's attitude at this stage:

- The methodology of the study had become accepted.
- The moral basis of naturalistic research had reassured the committee that, regardless of the uncovering of any perceived deficiencies, the personal worth of all stakeholders would be respected.
- The fact that some of the community's perceptions matched those of the committee validated the legitimacy of other perceptions that they had not heretofore held.

Stage E: Presentation of Pressure Points

The generation of the pressure points (phase 6) proved to be quite a difficult exercise for us. We found, in assigning responsibility for action to some part of the school system, that we tended to end up piling all the issues into the lap of the director, the natural locus of power. We had to work very hard to keep the solutions to the pressure points sufficiently open so that other components of the school system could take ownership for their resolution. We felt vindicated in the effectiveness of our work when, upon presentation of the pressure points, committee members accepted as givens those we had identified and concentrated their efforts on developing action plans for the amelioration of the irritants. We felt that we had indeed derived the pressure points from the data, and that these data were the same as the committee members had internalized. We were particularly gratified that the committee did not see any reason for us to be involved further; we had given them an understanding of the community's perception of the school system, and now it was up to them to take action.

We did have one regret. Even though we emphasized that the pressure points we had identified only constituted our thoughts, we found that the committee accepted our set of pressure points as *the* set; they did not generate any further points of their own. This confirmed our suspicion (Ryan, 1993) that the generation of recommendations (or, as in this case, near-recommendations) was a task of the "expert" that would not be challenged by the client. Thus, despite our undoubted success at convincing committee members to embrace the methodological and moral precepts of naturalistic research, we were dismayed at how quickly that sense of ownership could be handed back to the expert outsider.

CONCLUSION

Any previous naturalistic evaluation studies that either of us had tried had been conducted in a situation where all parties accepted the assumptions, strengths, and limitations of naturalistic methodologies. Those studies were smooth sailing. In this instance, we had to convince some members of the committee that our approach was the appropriate one. In the process we, in turn, had to compromise the tenets of naturalistic evaluation by incorporating more traditionally founded data collection phases and reporting procedures. The additional data collection phases helped gain acceptance for the primary data collection instrument, the open-ended survey with its inductive analysis. The reporting procedure of the pressure points certainly helped the committee establish what it wanted to do next, but the price that was paid was a return to the expert-client relationship that we had worked so hard to replace with a mutualistic vision of the evaluation process.

We have learned that we must be sensitive to the magnitude of the conceptual shift that changing paradigms entails. It should not have been a surprise to us; after all, it has taken us quite a bit of time and effort to understand and adopt naturalistic research methodologies. Perhaps that had slipped from our memories. We are grateful to the committee of the Torrance project for reminding us. We will try not to forget the next time.

REFERENCES

Alkin, M.C. (1990). *Debates on evaluation*. Newbury Park, CA: Sage.

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

Cousins, J.B., & Leithwood, K.A. (1986). Current empirical research on evaluation utilization. *Review of Educational Research*, *56*(3), 331–364.

Curriculum Branch, Education Department of Victoria. (1985). *Destination decisions: Decision-making strategies for school communities*. Melbourne: Author.

Guba, E.G., & Lincoln, Y.S. (1989). Fourth-generation evaluation. Newbury Park, CA: Sage.

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.

Lincoln, Y.S., & Guba, E.G. (1985). *Naturalistic inquiry*. Beverly Hills: Sage.

Patton, M.Q. (1986). *Utilization-focused evaluation* (2nd ed.). Newbury Park: Sage.

Patton, M.Q. (1988). The evaluator's responsibility for utilization. *Evaluation Practice*, *9*(2), 5–24.

Ryan, A.G. (1987). "Truth wrought in tales": An exploration of how case-study evaluations are received by those participating in the process. *Journal of Curriculum and Supervision*, *2*(4), 330–341.

Ryan, A.G. (1988). Program evaluation within the paradigms: Mapping the territory. *Knowledge: Creation, Diffusion, Utilization*, *10*(1), 25–47.

Ryan, A.G. (1993). An examination of the place of formal recommendations in naturalistic evaluations. *Evaluation and Program Planning*, *16*, 143–148.

Sanders, S.G. (1977). Force-field analysis: A functional management system. *Planning and Change*, *7*(4), 143–147.

Schwandt, T.A. (1989). Recapturing moral discourse in evaluation. *Educational Researcher*, *18*(8), 11–16.

Schwandt, T.A. (1991). Evaluation as moral critique. *New Directions for Program Evaluation*, *49*, 63–72.

Sudman, S., & Bradburn, N.M. (1982). *Asking questions: A practical guide to questionnaire design*. San Francisco: Jossey-Bass.

Weiss, C.H. (1972). *Evaluation research*. Englewood Cliffs, NJ: Prentice-Hall.

Weiss, C.H. (1988a). Evaluation for decisions. *Evaluation Practice*, *9*(1), 5–19.

Weiss, C.H. (1988b). If program decisions hinged only on information. *Evaluation Practice*, *9*(3), 15–28.

Weiss, C.H., & Bucuvalas, M.J. (1980). Truth tests and utility tests: Decision-makers' frames of reference for social science research. *American Sociological Review*, *45*, 302–313.

Appendix

CATEGORY: PROGRAM

Pressure Point #P8

REPORTING TO PARENTS ON STUDENTS SHOULD BE ACCOMPANIED BY CONSULTATION AMONG TEACHERS, PARENTS, AND STUDENTS

Documentation:

Survey ☒ Questionnaire ☐ Telephone ☐ Community ☒ Middle Years ☒ Comprehensive ☒

Context: Some parents of middle years and high school students are unclear about the reference points undergirding the reporting of student progress. They wonder about the extent to which their students' evaluations are self-referenced, criterion-referenced, or norm-referenced (although they do not use these technical terms).

Problem: Teachers need to make explicit the nature of their student evaluation program and communicate this to both students and parents.

What To Do:

1. Emphasize issues of student evaluation in teacher inservice.
2. Where appropriate, develop departmental, school, and system policies around student evaluation and reporting procedures.
3. Teachers should enter a dialogue with parents and students about evaluation and reporting procedures. This should be both formal (e.g., a written statement for parents and students at the beginning of the school year) and informal (ongoing reminders during interviews, at the beginning of major units, etc.)
4. Student evaluation policies and reporting procedures should be monitored so that any misconceptions or uncertainties can be dealt with quickly.

Action:	Who:	Timeline:	Indicators of Achievement:
1.			
2.			
3.			
4.			