THE INCLUSION OF STAKEHOLDERS IN EVALUATION: BENEFITS AND DRAWBACKS

Daniel B. Morris
University of British Columbia
Vancouver, British Columbia

Abstract: The current trend in evaluation research is to include stakeholders in the evaluation process. Central to this trend is the belief that stakeholder involvement empowers the stakeholders, increases the utilization of results, and increases the validity of the evaluation. This article examines both the theoretical foundation for this belief and its empirical support. Conditions under which these benefits are realized are analyzed and compared to situations that fail to gain these benefits.

Résumé: La tendance actuelle dans l'évaluation de recherche est d'inclure des intervenants dans le processus d'évaluation. L'idée centrale de cette tendance consiste en trois aspects: l'implication des intervenants habilite les intervenants, accroit l'utilisation des résultats et augmente la validité de l'évaluation. Cet article examine la fondation théorique de ce concept et son soutien empirique. Les conditions dans lesquelles sont réalisées les bénéfices sont analysées et comparées aux situations qui manquentd'obtenir ces bénéfices.

██████████ Many factors must be considered for an effective program evaluation. One issue that has generated controversy is the inclusion of stakeholders in the evaluation process. This article attempts to illustrate the diverse views of stakeholder participation though careful examination of published research. The approach taken here recognizes that the term "stakeholder" has had different definitions for different researchers over the years and that these operational differences continue to persist today. To address this inconsistency, researchers within both the objective-scientist and social-constructivist paradigms are moving toward explicitly defining how the term "stakeholder" is used (e.g., Collins, 2000; Scriven, 1997; VanderPlaat, Samson, & Raven, 2001). Indeed, Scriven (1997) distinguishes between clients, consumers, payers, stakeholders, and audience, providing definitions for each term while acknowledging that the terms are not mutually exclusive. However, using an explicit definition for stakeholder in this article would not reflect the

inconsistencies still evident in the literature. Therefore, a conscious and deliberate decision was made to define stakeholder broadly in this article ranging from token to active participation by recipients, distributors, or financial supporters of program services.

It is generally agreed that a primary motivation for including stakeholders in an evaluation is based on the values held by the evaluators and funding agency. More specifically, the value of empowering stakeholders by encouraging active participation is central to this view (Fetterman, 1999). Also, greater participation of stakeholders is thought to increase the utilization of the program evaluation results, and the validity of these results (Brandon, 1998; Greene, 1988). The main drawbacks of stakeholder participation tend to be quite pragmatic in nature. Additional time, personnel, and expenses required are cited as drawbacks. However, it has also been demonstrated that stakeholder participation can lead to feelings of lack of power by participating stakeholders, disregarding of results by decision makers, and questionable validity (Lawrence, 1989; Mercier, 1997; O'Brecht, 1992). Thus, stakeholder participation in evaluations can influence the *same* outcomes both positively and negatively. This article shall explore each of these outcomes from a theoretical perspective and then analyze empirical studies to determine under what conditions and to what extent these benefits and drawbacks are realized.

OBJECTIVE-SCIENTIST PERSPECTIVE

Until the 1970s, scientific objectivity was a central concern of program evaluation. Stakeholders, including recipients, distributors, and financial supporters of services, were considered sources of bias, interfering with the development of objective and valid program evaluations (O'Brecht, 1992). Program participants were viewed as "objects of inquiry" that could occasionally be consulted for clarification while management was the main audience of the evaluation (VanderPlaat et al., 2001). Any contact with stakeholders was limited and focused on obtaining facts concerning program functioning and effects (Brandon, 1998). Nevertheless, by the 1970s evaluators began to see stakeholders as a source of more than just factual information and began to actively seek out stakeholder involvement.

SOCIAL CONSTRUCTIVIST PERSPECTIVE

The origins of participant involvement in evaluation can be found in the social constructivist perspective employed in qualitative re-

search. This view has gained wide acceptance in the program evaluation community to the extent that the first utility standard in the program evaluation standards published by the Joint Committee on Standards for Educational Evaluation states:

> Persons involved in or affected by the evaluation should be identified, so that their need can be addressed. (Joint Committee on Standards for Educational Evaluation, 1994)

Utility standard 3 clarifies:

> Information collected should be broadly selected and address pertinent questions about the program and be responsive to the needs and interests of clients and other specified stakeholders. (Joint Committee on Standards for Educational Evaluation, 1994)

These standards make quite evident the transition regarding the role of stakeholders in program evaluation. Indeed, stakeholders are now considered within the ethics of program evaluation. Guideline two (of three) for ethical conduct by the Canadian Evaluation Society states:

> Evaluators are to act with integrity in their relationships with all stakeholders. (Canadian Evaluation Society, 1996)

Thus, the role of the stakeholder has gained prominence in evaluation studies. A brief overview of the reasoning behind this transition follows.

The basic premise behind social constructivism is that there is no single truth. Rather, multiple perspectives converge toward a commonly accepted reality. Indeed, these perspectives may change over time and place (Abma, 2000). Drawing on the responsive-constructivist approach promoted by Lincoln and Guba (1985) and Guba and Lincoln (1989), Abma (2000) encourages the ideal that stakeholders should be actively involved throughout the evaluation process as partners. Thus, all perspectives related to an evaluation are valued and should be actively sought from stakeholders to gain a complete picture of program rationale, impact, and alternatives (O'Brecht, 1992). Subsequently, stakeholder involvement using the principles of social constructivism should lead to a more valid evalu-

ation that empowers stakeholders and increases the likelihood of the utilization of results as stakeholder investment is greater than would otherwise be the case.

EMPIRICAL SUPPORT FOR STAKEHOLDER PARTICIPATION

Cousins and Earl (1992) propose several organizational and evaluator characteristics that must be met for a successful participatory evaluation. These characteristics are examined as they relate to empowerment, utilization of results, and validity of results.

Organizational Requirements

1. Evaluation must be valued by the organization.
2. The organization must provide the time and resources required.
3. The organization must be committed to organizational learning as a route toward improvement.
4. Primary users participating in evaluation must be motivated to participate.
5. It is necessary to realize that participants will not have sufficient research experience and knowledge initially. However, adequate knowledge can be provided with training.

Evaluation Requirements

1. The evaluator must have adequate training and expertise in research skills.
2. Evaluators must be accessible for participant activities.
3. Resources (i.e., support services, budget) must be available.
4. The evaluator must take on the role of a teacher about evaluation.
5. Evaluators must be motivated to work directly with participants.
6. Evaluators must have a tolerance toward problems and imperfections.

Empowerment

Empowerment is the increased feeling or sense of power stemming from a given action, in this case participation. Feelings of being listened to, being taken seriously, and making an important contribu-

tion are characteristic of empowerment. Moreover, studies have demonstrated that participation leads to empowerment (e.g., Greene, 1988; Hunter & Gambell, 2000). In each of these instances participation was lengthy, deep, and situated within a context that actively sought and appreciated participation. Indeed, all the characteristics for a successful participatory evaluation described above were met. When these requirements are met the following advantages are gained (Lawrence, 1989):

1. Diagnosis of consensus and differences in definition of program goals and objectives can be addressed.
2. Stakeholders will perceive that the evaluation is explicitly recognizing their needs.
3. Overly intrusive or inappropriate evaluation procedures can be avoided.

When all of the above characteristics are not satisfied, empowerment may not be realized. In fact, a sense of empowerment may be diminished. For example, in a case described by Mercier (1997), the organization expressed fear that the evaluation would discredit the program and compromise funding due to a lack of understanding the complexity of the work performed in the evaluation. Stakeholders questioned the legitimacy, quality, and conclusiveness of the evaluation with increasing anxiety as the evaluation progressed. The evaluation continued despite this uncomfortable climate. Rumours regarding early results increased the apprehension. Even the data collection process of participant observation and semi-structured interviews led to greater anxiety and feelings of intrusion. Committee meetings revealed inequality among participants. Participants were given insufficient time or opportunity to think over concerns or propose modifications. Some participants even voiced resentment at being asked to participate without deriving any power from the process. In short, although participatory and qualitative in nature, this evaluation was very disempowering.

Mercier (1997) indicated several factors that exacerbated the problems in this evaluation. First, stakeholders were involved relatively late in the evaluation process rather than at the outset. Further, stakeholders were not given much choice regarding how they were expected to participate. No clear indication of how the evaluation benefited all stakeholders was provided, nor was information consistently distributed to all members. Finally, Mercier suggested that stakeholder participation might not have been appropriate within the restrictive timeframe used. These factors correspond well with

the participatory requirements proposed by Cousins and Earl (1992). Thus, great care must be made in satisfying these requirements lest the process become disempowering.

Abma (2000) adds that conflict should be anticipated in participatory evaluations. If not anticipated, conflict can easily lead to feelings of disempowerment. The following recommendations are therefore made to manage conflict.

1. Make hidden conflicts visible.
2. Characterize both sides of the conflict as particular ways of looking at a situation.
3. Encourage people to suspend their positions and probe others for their reasoning in order to discover new possibilities.
4. Help opposing stakeholders gain insight into alternate perspectives by role-playing different positions.

When the evaluator is prepared for conflict and enables disputants to view alternate perspectives, empowerment can once again be realized. If stakeholder views are not considered to be of equal value, the participant evaluation process may actually increase the conflict and power differential among the groups, thereby polarizing participants and impeding any productive discussion (Brandon, 1998).

Utilization of Results

Cousins and Leithwood (1986) suggest that six factors affect the utilization of evaluation results.

1. Evaluation quality: Is the methodology considered sophisticated and appropriate?
2. Credibility: Is the evaluator reputable and credible?
3. Relevance: Is the evaluation geared toward its intended audience(s)?
4. Communication: Is communication sufficiently frequent, detailed, and of adequate quality?
5. Findings: Do the results correspond to those expected?
6. Timeliness: Is the evaluation produced sufficiently before decisions must be made?

Following some criticisms that "human factors" including evaluator willingness to involve users and establish rapport with these users

were omitted in the original model, an additional factor was added (Cousins & Earl, 1992).

> 7. Evaluator interaction: Does the evaluator demonstrate a willingness to involve users and establish good rapport?

Factors 4 and 7 are clearly linked to stakeholder participation.

Greene (1988) examined in more detail how participation is linked to utilization. She distinguished between three different types of stakeholder involvement: the very involved person (VIP), the somewhat or sometimes involved person (SIP), and the marginally involved person (MIP). The stakeholder type that is most prevalent will affect the utilization of the evaluation. The VIP perceives the decision-making process as open, consensual, and democratic. The process was enjoyable and considered worthwhile with plenty of opportunities to participate and no major obstacles to participation. The SIP perceives the process as valid and credible with stakeholder perspectives genuinely desired. However, the overall length and slow pace of evaluation made sustaining interest and participation difficult at times. The MIP felt uncomfortable participating and was not clear what would happen with the information provided. Written communications were frequent, confusing, and hard to understand, leading participants to feel frustrated and stupid. When stakeholders were either VIP or SIP, utilization of results increased. Greene (1988) suggests this increase was due to greater acceptance and ownership of results, greater sense of responsibility to follow through with the evaluation findings, and greater sense of the results being valid, credible, legitimate, and persuasive.

These benefits in utilization are not as likely when stakeholders are only mildly involved. This was demonstrated by O'Brecht (1992), who found that the inclusion of too many stakeholders can lead to minimal participation of any individual stakeholders. Indeed, much of the stakeholder involvement described merely summarized previously collected information, slowing the program evaluation process to the extent that the evaluation itself was trivialized and largely ignored (i.e., not well utilized). This poor utilization was due to a large extent to the lack of clarity and redundancy concerning each stakeholder's role in the evaluation process. Once roles were clarified, made less redundant, and communicated, stakeholders became more actively involved and utilization of results improved.

Validity of Results

> Brandon (1998) argues that stakeholder participation can enhance
> the validity of an evaluation. That is, more adequate and appropri-
> ate inferences can be made from the collected data and be reported
> in an evaluation. To ensure an evaluation is valid, the evaluator
> must possess a detailed understanding of the program, which is best
> obtained by involving program experts in the evaluation. These ex-
> perts have factual and practical knowledge, can simplify complex
> problems, and can identify what information is relevant for a deci-
> sion. By definition, stakeholders are program experts, and thus
> should be involved in the program evaluation. Program personnel
> are critical in identifying the problems the program is designed to
> address. Program beneficiaries are often aware of implementation
> and outcome details that personnel do not know.

> Nonetheless, Brandon (1998) cautions that if all stakeholder groups
> are not involved equitably, a single group may co-opt the evaluation
> process as it tries to maximize its own importance and worth, thereby
> invalidating the evaluation results. Should this occur, the results
> might be even more inaccurate than if no stakeholders were con-
> sulted at all. Indeed, it has been argued that documents and obser-
> vations will produce more objective and reliable results (see O'Brecht,
> 1992). However, even in these situations the documents and obser-
> vations accessed are informed by stakeholders and still subject to
> bias and questionable findings.

CONCLUSION

> As illustrated in the preceding discussion, stakeholder involvement
> in evaluations can be both beneficial and detrimental. Under ideal
> circumstances, in which all of the principles outlined by Cousins
> and Earl (1992) are met, participants are empowered, results are
> more likely to be utilized, and the validity of the evaluation is in-
> creased. However, in naturalistic settings where most evaluations
> occur, these principles are often not met. Therefore, the evaluator
> must compensate as much as possible for their omission. This may
> involve acting as an educator, a mediator, or a facilitator. Nonethe-
> less, it is worth re-emphasizing that sufficient time, money, and
> support services must be made available for these adjustments to
> be made, or the potential value of stakeholder evaluation will not be
> realized. Further, once there is greater consistency in defining
> stakeholders, the benefits and detriments of stakeholder participa-

tion explored in this article will be provided with greater substance. But even while the term stakeholder is used inconsistently, some valuable insights can be gained that should help generate discussion on this important topic.

REFERENCES

Abma, T.A. (2000). Stakeholder conflict: A case study. *Evaluation and Program Planning, 23*, 199–210.

Brandon, P.R. (1998). Stakeholder participation for the purpose of helping ensure evaluation validity: Bridging the gap between collaborative and non-collaborative evaluations. *American Journal of Evaluation, 19,* 325–337.

Canadian Evaluation Society. (1996). *CES guidelines for ethical conduct.* Ottawa: Author.

Collins, M. (2000). Critical perspectives on educational evaluation. *Canadian Journal of Program Evaluation [Special issue],* 111–135.

Cousins, J.B., & Earl, L.M. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis, 14*, 397–418.

Cousins, J.B., & Leithwood, K.A. (1986). Current empirical research on evaluation utilization. *Review of Educational Research, 56,* 331–365.

Fetterman, D.M. (1999). Reflections on empowerment evaluation: Learning from experience. *Canadian Journal of Program Evaluation [Special issue],* 5–37.

Greene, J.G. (1988). Stakeholder participation and utilization in program evaluation. *Evaluation Review, 12,* 91–116.

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

Hunter, D.M., & Gambell, T.J. (2000). Incorporating stakeholders in standard setting: What's at stake? *Canadian Journal of Program Evaluation [Special Issue],* 83–109.

Joint Committee in Standards for Educational Evaluation. (1994). *The program evaluation standards: How to assess evaluations of educational programs* (2nd ed.). Thousand Oaks, CA: Sage.

Lawrence, J.S. (1989). Engaging recipients in development evaluation: The "stakeholder approach." *Evaluation Review, 13,* 243–256.

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

Mercier, C. (1997). Participation in stakeholder-based evaluation: A case study. *Evaluation and Program Planning, 20,* 467–475.

O'Brecht, M. (1992). Stakeholder pressure and organizational structure for program evaluation. *Canadian Journal of Program Evaluation, 7*(2)*,* 139–147.

Scriven, M. (1997). Truth and objectivity in evaluation. In E. Chelimsky & W.R. Shadish (Eds.), *Evaluation for the 21st century: A handbook* (pp. 477–500). Thousand Oaks, CA: Sage.

VanderPlaat, M., Samson, Y., & Raven, P. (2001). The politics and practice of empowerment evaluation and social interventions: Lessons from the Atlantic community action program for children regional evaluation. *Canadian Journal of Program Evaluation, 16*(1)*,* 79–98.