

Being Blind in a World of Multiple Perspectives: The Evaluator's Dilemma Between the Hope of Becoming a Team Player and the Fear of Becoming a Critical Friend with No Friends

Michele Tarsilla
Western Michigan University

Background: A large number of evaluation theorists have debated the issue of objectivity and bias in evaluation over the last five decades. In particular, the degree to which distance from the evaluand enhances the validity and reliability of evaluation findings has been a prominent topic of discussion.

Purpose: This article has two primary objectives. First, it intends to present some of the positivist and post-positivist theories on distance that have dominated the evaluation discourse since the late 1960's, by also showing the limitations of their respective assumptions. Second, it describes a more recent evaluation theory on distance that is proving to help evaluators rapport with their evaluand more effectively, especially in the case of complex programs involving a large variety of stakeholders.

Setting: Not applicable.

Subjects: Not applicable.

Research Design: Not applicable.

Data Collection and Analysis: The paper is the result of both a desk review and a series of interviews with some major evaluation theorists aimed to compare and contrast some of the most relevant ideas on distance in evaluation expressed over the last five decades.

Findings: The author shows that evaluators today are still facing the dilemma of whether they should

seek proximity to or distance from the evaluand. However, the author identifies an increasingly popular evaluation approach (herewith referred as the pluralist approach as opposed to the niche approach) that promises to overcome the issue of distance in evaluation more successfully than any other earlier theory. The author dismisses the idea of absolute distance, predicated by both Scriven and Campbell. In doing so, he also shows that evaluators who are closer to the evaluand and the context contiguous to it tend to have a deeper understanding of the issues at stake and therefore enhance the overall quality of their evaluation. In addition, the author acknowledges that evaluators today have a new important role to play vis-à-vis their evaluand: mediating stakeholders' competing values and agendas for the sake of equity and social justice.

Conclusions: The author concludes that the proximity to the evaluand and the integration of multiple perspectives in an evaluation represent two of evaluation's most enriching—rather than detrimental—factors. The author also asserts that a truly participatory approach can effectively coexist with advocacy, so long as evaluators are able to clarify their stances vis-à-vis the social, economic, and political issues associated with their evaluand.

Keywords: *distance, bias, transformative evaluation, advocacy, pluralist and niche approaches*

When can an evaluation be considered objective, provided that objectivity really exists?¹ To what extent can the determination of merit, worth, or significance of a given product or program be regarded as bias-free? What is the optimal degree of distance from program stakeholders that evaluators should maintain in the course of their assignments? These are only a few of the questions that have animated the discussions among evaluators over the last decades.

On the one hand, those in the evaluation field who have embraced the positivist principles of objectivity sought a direct and unbiased observation of natural phenomena, have strongly defended the value-free and scientifically objective nature of evaluation. Such is the case of those evaluators who believe that methodological rigor (e.g., the use of randomized, controlled trials) warrants credibility to their work and enhances their intrinsic impartiality (Campbell & Stanley, 1969).

On the other hand, those evaluators who have espoused more post-positivist epistemological views have admitted that any human observation (including those conducted by evaluators) is always filtered through personal biases. Michael Scriven himself recognized quite candidly that direct and unbiased observations of reality are not possible (Scriven, 1976).

Quite an interesting idea, if it weren't for the fact that some evaluation theorists, including Scriven himself, developed it further and suggested that evaluators

should even isolate themselves from the evaluand (Scriven, 2000) in order to reduce the biases of their valuing. Scriven even coined the expression "blind is beautiful" (Scriven, 1967) to support the core assumption of his goal-free evaluation model: the greater the distance from a program being evaluated, the greater the evaluators' neutrality and the impartiality of their judgments.

Acknowledging that human observations and evaluative judgments are not infallible undoubtedly represented a milestone in the development of the evaluation theory. However, the recommended use of distance from the evaluand as a panacea for all personal biases appeared an oversimplification.

In particular, I find it controversial that Scriven, while urging evaluators to be empiricists and logicians, called upon them to render their evaluative judgments with the caveat (unavoidable in my opinion) of not mixing personal opinions with the logical resolution of their evaluative claim.

I also do not believe that seeking distance from the evaluand is either realistic or beneficial to the evaluator's work for the following four reasons.

First, if it is true that "multiple accounts of reality help build a true picture" (Scriven, 2000), how is it possible for evaluators to "build the picture" (that is, to make a synthesis of all the collected information and provide an evaluative judgment) according to their personal understanding of what stakeholders' needs are rather than based on the multiple perspectives provided by stakeholders?² And yet how is isolation

¹ Interestingly enough, a quick Google search conducted on September 30, 2009 seems to support the view that an "objective evaluation" can and does exist (450,000 related links were identified as opposed to the 327,000 found for the term "subjective evaluation").

² It is hard to imagine, for instance, that international development evaluators (allegedly in search of the public truth cherished so much by both Campbell and Scriven) could distance

from stakeholders contributing to evaluators' identification of program side effects, a skill that Scriven defines as extremely important for evaluation (Scriven, 1993)?

Second, if the purpose of evaluation is to determine whether needs have been met (Scriven talks of a needs-referenced evaluation) and if the evaluator's main objective is to unveil the public truth through his/her work, how could this be possibly done without either a direct dialogue with those who shape public truth (e.g., responsive and participatory evaluation) or a thorough understanding of those whose needs are still unmet (which, in my opinion, might require the evaluator to build a rapport of trust with stakeholders)? Clearly, so long as evaluators maintain their distance from the evaluand, for the sake of objectivity and impartiality, evaluators will not be able to render judgments that genuinely capture the true merit, worth and significance of a program.

Third, by cherishing the value of distance, evaluators risk elevating themselves above everybody else. As a result, far from being immersed in the reality of the evaluand and lacking the understanding of the environmental factors inseparable from it, evaluators end up being tempted to use unilateral or non-inclusive criteria (untestable or not subject to any form of scrutiny) in judging the value of a certain program or product. That would inevitably make evaluations biased and not appreciative of that "multiplicity of multiples" that Scriven defines as a critical component of any sound evaluation (Scriven, 1983).

themselves from program stakeholders and formulate their evaluative judgments after only a few sporadic contacts with either those who fund or coordinate such programs.

Fourth, by keeping their distance and making inferences about their evaluand based solely on observable behaviors, evaluators risk compromising the utilization of the evaluation findings. For instance, by overlooking program participants' intent, thoughts, feelings and opinions (Rosas, 2006), evaluators are less likely to capture the program's social validity (Hawkins, 1991) and, more importantly, to predict the feasibility and stakeholders' acceptability of their final recommendations (e.g., scaling up or modification of the program being evaluated).

Overall, Scriven's recognition of the finite and partial nature of human knowledge along with the need for taking competing perspectives into account when making evaluative judgments were relevant for the development of both evaluation theory and practice. However, Scriven's theories presented two main limitations, which some later thinkers (Stake 2004, 2006; Weiss 1977, 1996; Wholey 1983; Wholey & Newcomer, 1989) attempted to address in their work. First, the ambiguity of the evaluation purposes (e.g., what is an evaluator's main responsibility? Is it really to identify the public truth and unmet needs? If so, is the evaluator's isolation instrumental in achieving the envisaged evaluation purposes?). Second, the lack of clarity on how to apply those very same idealistic objectives in order to achieve the identified evaluation purposes (e.g., whose responsibility is it to turn evaluative judgments into concrete public actions with a social significance?)

In order to address the limitations of the early evaluation theories (popular in the 1960's and early 1970's), which viewed proximity to stakeholders as detrimental to the objectivity or the scientific nature of the evaluator's valuing, new perspectives

on distance in evaluation emerged in the late 1970's and 1980's. Thanks to such new paradigms, evaluators (finally relieved from the nonsensical pressure to be objective and value-free in their profession) started redefining the boundaries of their roles and responsibilities. For example, they quickly became used to the notion of advocacy in evaluation (Fitzpatrick, 2001; Greene, 2004; Mertens, 2008), something that was simply unthinkable in the past.

The notion of advocacy itself developed over the years to the extent that the question today is no longer whether evaluators should promote values in the course of their practice but rather whom they should advocate for. Some of the most influential evaluation theorists today underscore the necessity for evaluators to play the role of critical friend to their clients (Fetterman, 1997), or to advocate in favor of their program participants (Greene 1997) or intended primary users (Patton, 1997). Others (Mertens, 2007) call upon evaluators to promote social justice and human rights among the most vulnerable and voiceless segments of the population that are significantly affected by the programs being evaluated. Far from being unobtrusive observers, as was the case with Scriven, evaluators are finally able to sit at the table with the evaluand (Caracelli, 2006) and become themselves recognized members of the groups being observed.

The development of such radically new perspectives on evaluators' distance (or rather proximity) undoubtedly has been very beneficial to many practitioners. By encouraging evaluators to become more cognizant of their own biases, such theories have helped evaluators to make their personal ideological stances vis-à-vis the evaluand more explicit in their own work so as to render more balanced and

defensible evaluative judgments. In particular, two main approaches seem to have contributed the most to the development of the theory of distance in evaluation over the last few decades: the niche and the pluralist approaches.

According to the niche approach evaluators, instead of engaging with all stakeholders (which would compromise the objectivity of the evaluation too much), should mainly focus on a specific group among them. Patton's utilization-focused evaluation (Patton, 1997), for instance, has urged evaluators to maintain trust and rapport with their "intended primary users" (mostly program managers) in order to avoid "isolated pearls of evaluation wisdom." Similarly, Weiss has focused on developing a dialogue with policy-makers (Weiss 1977, 1996); Wholey has privileged working with program managers (Wholey 1983; Wholey & Newcomer, 1989); and Stake has engaged himself with local administrators and project staff (Stake 2004, 2006).

Such evaluation theorists have encouraged evaluators to become close to one specific stakeholder group, rather than the totality of program participants. However, they failed to provide evaluators with the tools to address two major limitations associated with the niche approach: the more limited accountability and, as a result, the lower guarantee that evaluation findings will be used; and the lack of inclusiveness in the evaluation design and implementation.³

³ In other words, the theories in favor of a "segmented stakeholder involvement" did not provide an effective solution to the question of distance in evaluation. Is it hard, for instance, to praise the evaluation of a community youth program whose findings have a negative impact on the beneficiaries due to the evaluator's excessive focus on the program manager's objectives (e.g.

According to the pluralist approach, evaluators are called upon to become extensively involved with the program stakeholders and cognizant of the diversity found within the context where the evaluation is being conducted.

Such a promising approach is innovative for two main reasons. First, it confers evaluators with a new and interesting role: mediating stakeholders' competing values and agendas for the sake of equity and social justice. As a result, the integration of multiple perspectives in an evaluation becomes one of evaluation's most enriching—rather than detrimental—factors.

Second, the approach shows that, that a truly participatory approach can effectively coexist with advocacy, so long as the evaluators clarify their stances vis-à-vis the social, economic and political issues associated with their evaluand. This is case of transformative evaluators, who address power imbalances in society as part of their work by also pushing further for social change and equity (Mertens, 2009).

In conclusion, the debate on objectivity and bias in evaluation (e.g., which should be the optimal degree of evaluators' distance from program stakeholders or to what extent the determination of merit, worth, or significance of a given product or program can be regarded as bias-free) is far from being settled. In particular, new and increasingly popular evaluation approaches are stirring the discussion on these topics and spurring the elaboration of a new theory of distance in evaluation.

The pluralist approach is a good example of that. By opposing Scriven's

demonstrating the program effectiveness to the donor) rather than on needs and perspective of the local youth affected by the program.

notion of absolute distance and blindness in evaluation, pluralists call on evaluators to increase their proximity to the evaluand and claim that such practice promotes a better understanding of the internal dynamics and contextual factors associated with the evaluand.

Thanks to the pluralist approach and also based on the assumption that an evaluation will never be completely objective, evaluators are finally relieved from the nonsensical pressure to be value-free and impartial at all costs. However, that entails some new challenges for evaluators: first, to identify and reconcile the variety of values and often competing agenda associated with the evaluand; second, to be upfront about their own values before starting an evaluation, with the caveat that the promotion of personal value(s) is not an arbitrary choice but rather a corollary of the specific evaluation strategy being adopted (including its purpose, audience and key questions).

References

- Campbell D. T., & Stanley, J. C. (1969). *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand McNalley.
- Caracelli, V. J. (2006). The evaluator-observer comes to the table: A moment of consequential decision making. *American Journal of Evaluation*, 27(1), 104-107.
- Fetterman, D. M. (1997). Empowerment evaluation: A response to Patton and Scriven. *American Journal of Evaluation*, 18(3), 253-266.
- Greene, J. C. (1997). Evaluation as advocacy. *American Journal of Evaluation*, 18(1), 25-35.

- Greene, J. C. (2004). The educative evaluator: An interpretation of Lee J. Cronbach's vision of evaluation. In M. C. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences* (pp. 169-181). Thousand Oaks, CA: Sage.
- Hawkins, R. (1991). Is social validity what we are interested in? Argument for a functional approach. *Journal of Applied Behavioral Analysis, 24*(2), 205-213.
- Mertens, D. M. (2007a). Transformative considerations: Inclusion and social justice. *American Journal of Evaluation, 28*(1), 86-90.
- Mertens, D. M. (2007b). Transformative paradigm: Mixed methods and social justice. *Journal of Mixed Methods Research, 1*(3), 212-225.
- Mertens, D. M. (2008). Stakeholder representation in culturally complex communities: Insights from the transformative paradigm. In N. L. Smith & P. R. Brandon (Eds.), *Fundamental issues in evaluation* (pp. 41-60). New York, NY: Guilford.
- Mertens, D.M. (2009). *Transformative research and evaluation*. New York, NY: Guilford.
- Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text* (3rd ed.). Thousand Oaks, CA: Sage.
- Rosas, S. R. (2006). Nonparticipant to participant: A methodological perspective on evaluator ethics. *American Journal of Evaluation, 27*(1), 98-103.
- Scriven, M. (1967). The methodology of evaluation. In R. Tyler, R. Gagne, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39-83). Chicago, IL: Rand-McNally.
- Scriven, M. (1976). Evaluation bias and its control. In G. V. Glass (Ed.), *Evaluation studies review annual* (Vol. 1, pp. 101-118). Beverly Hills, CA: Sage.
- Scriven, M. (1993). Hard-won lessons in program evaluation. *New Directions in Program Evaluation, No. 58*. San Francisco, CA: Jossey-Bass.
- Scriven, M. (2000). Evaluation ideologies. In D. L. Stufflebeam, G. F. Madaus, & T. Kelligan (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (pp. 249-278). Boston, MA: Kluwer.
- Shadish, W. R., Cook, T. D., & Leviton, L. C. (1991). *Foundations of program evaluation: Theories of practice*. Newbury Park, CA: Sage.
- Stake, R. E. (2004). *Standards-based and responsive evaluation*. Thousand Oaks, CA: Sage.
- Stake, R. E. (2006). *Multiple case study analysis*. New York, NY: Guilford.
- Weiss, C. H. (Ed.). (1977). *Using social research in public policy making*. Lexington, MA: Lexington Books.
- Weiss, C. H. (1996). Excerpts from evaluation research: Methods of assessing program effectiveness. *American Journal of Evaluation, 17*(2), 173-175.
- Wholey, J. S. (1983). *Evaluation and effective public management*. Boston, MA: Little, Brown.
- Wholey, J. S., & Newcomer, K. E. (1989). *Improving government performance: Evaluation strategies for strengthening public agencies and programs*. San Francisco, CA: Jossey-Bass.