Theorists’ Theories of Evaluation: A Conversation with Jennifer Greene

Michele Tarsilla
Western Michigan University

Background: Evaluation is sometimes viewed as a professional practice rather than a discipline corresponding to a well defined set of theories. However, Shadish, Cook and Leviton (1991) were able to demonstrate that evaluators’ work does have theoretical foundations. In particular, the authors identified five main elements for evaluation theory and described the contribution made to each of them by seven of the most influential scholars in the field over the last five decades.

Purpose: This paper intends to further the discussion on evaluation theory, by examining some of the contributions made Jennifer Greene, one of the most influential figures in contemporary evaluation. The paper mainly focuses on Greene’s innovative ideas on each of the five main elements of evaluation theory.

Setting: Not applicable.

Subjects: Not applicable.

Research Design: Not applicable.

Data Collection and Analysis: The paper is the result of both a desk review of Jennifer Greene’s most relevant work on bias, objectivity, and advocacy in evaluation, and a phone interview with her. For the sake of accuracy, the text of the interview and the corresponding analysis were submitted to Greene for review prior to publication.

Findings: The author shows how Greene has incorporated the five principles into her own work and how this eventually influenced her practice. Greene sorts evaluation approaches based on the interests they serve and the values they promote. However, she seems to have developed her theory on valuing further over the years and today she claims that evaluators should never privilege anyone’s specific side in the course of their assignments. Second, as knowledge is mediated by evaluators’ perceptual frames, Greene believes that an unfiltered (objective) view of the world is not feasible. Third, she views evaluation as a force for democratizing public conversations about important issues. Fourth, although the evaluator’s relationship with program staff could be collegial, Greene believes that evaluators have no authority or responsibility for the program design and implementation. Fifth, Greene declares that advocacy in evaluation is inevitable and, as a result, evaluators should play a socially enfranchising role today.

Conclusions: The debate on evaluation’s main theoretical foundations is still relevant. For this purpose, the author recommends that the five main theories applied by this article to examine Jennifer Greene’s work should be used more systematically in the future to describe and analyze evaluators’ practice. Such theoretical categories would be especially beneficial in that they will provide some common ground of understanding among both practitioners and scholars on evaluation concepts and practices which experience has shown to be in constant evolution.

Keywords: evaluation theory, valuing, social programming, knowledge, practice, use
The idea to conduct an interview with Jennifer C. Greene, one of the most influential figures in contemporary evaluation, came to me after taking a full-day workshop on mixed methods which she taught at Carleton University in Canada in April 2009. While I had the chance to become familiar with her work last spring, I became particularly interested in her contribution to the ongoing discourse on bias and objectivity in evaluation after I started the Interdisciplinary Ph.D. in Evaluation (IDPE) program at Western Michigan University (WMU) in September 2009.

In order to prepare this interview, I first conducted a thorough review of Greene’s most relevant work on distance and advocacy in evaluation, and tried to summarize and classify some of her key perspectives under five main categories (Theory of Valuing, Theory of Knowledge, Theory of Social Programming, Theory of Use, and Theory of Practice). In doing so, I drew upon the original work of Shadish, Cook and Leviton (1991), who used those same five categories to describe the contribution of seven influential scholars (i.e., Donald Campbell, Michael Scriven, Carol Weiss, Joseph Wholey, Robert Stake, Lee Cronbach, and Peter Rossi) to the development of evaluation theory over the last five decades.

Then, based on my readings, I formulated a few questions on Greene’s thinking related to each of the five identified theories and had a chance to discuss them directly with her during a 1-hour phone interview held the morning of December 10, 2009.

Questions and answers are presented in five different sections based on the specific theoretical dimension to which they refer. In order to facilitate the reader’s better understanding of the answers provided, each of the five sections begins with a synopsis of Greene’s work related to each of the five dimensions.

### Theory of Valuing

According to Greene (1997), evaluators’ work, including the use of their findings, is heavily influenced by values. Values are so central in evaluators’ work that Greene herself sorts evaluation approaches based on the interests they serve and the values they promote (2004). Values—Greene adds—are so pervasive in evaluators’ practice that the most accurate information on evaluators’ values could be easily found in either the first or last section of every evaluation report where the evaluation questions and the criteria for making evaluative judgments are usually listed.

Although valuing has been a long established pattern of Greene’s thinking, her ideas and definitions associated with the term seem to have undergone three stages of change over time. In her early years, far from prescribing what the “right” or “best” values should be, Greene simply provided a general definition of values in an article published in the American Journal of Evaluation in 1997: “Values are particular regulative ideals, such as rational decision making, interpretive meaning or community activism” (Greene, 1997, p. 27).

In doing so, she left evaluators free to pick the values which they wanted to commit to (Greene, 1997). Greene cautioned evaluators, though, that opting in favor of one value or another should be based on the role that they wished to play in society rather than on a partisan stance toward a particular program or an
alignment with a particular stakeholder group.

Years later, Greene seemed to have developed her theory on valuing further and, by placing more emphasis on evaluators’ responsibility to take multiple stakeholders’ views into account (democratic pluralism), she stressed that evaluators should never privilege anyone’s specific side in the course of their assignments.

More recently, Greene (2009) revised her general definition of values and made her theory of valuing more prescriptive, as shown by the following exchange which I had with her during the phone interview.

TARSILLA: Jennifer, while your article on advocacy published on the American Journal of Evaluation in 1997 seemed to stress the importance of democratic pluralism, your overall theory of valuing appeared to be not prescriptive then. Is that a fair assumption to make? If yes, why was it so?

GREENE: Yes, you are right. My theory of valuing was not prescriptive back then. However, I would certainly say that my theory of valuing is quite prescriptive today. My position vis-à-vis valuing in evaluation has definitely evolved over time and this has been a challenging but inevitable process. After all, once you accept the intrinsic presence of values in evaluation and then become well aware of what your own values are and how these affect your work, you are naturally inclined to want to advance those same values in your evaluation practice. However, in order to do that, you need to make them more explicit in your work and you need to be eager to share them with others. You cannot be a trickster and promote your agenda secretly within the scope of your work. You need to be open and accountable to both your clients and the whole society, whose needs I believe evaluators should always take into account in their practice. Objectivity and neutrality in and of themselves are values, too, and I have seen colleagues pursuing them with passion and determination.

In her search of values affecting evaluators’ practice, Greene over the years has been able to identify a distinct set of core values which, in her opinion, constitute a key reference for all evaluators’ work.

TARSILLA: Jennifer, while your theory of valuing was not prescriptive in your early years, it seems that you have radically changed your position on this, as demonstrated by your recent work on the five main values which, in your opinion, evaluators should follow. Could you please comment on that?

GREENE: Well, I would rather talk of five main genres of evaluation practice guiding evaluator’s work these days — each advancing different value because each addresses the needs of a different audience. These needs include: (1) the efficiency interests of policymakers (Weiss), (2) the accountability and ameliorative interests of on-site program managers (Patton), (3) the learning, understanding and use (Cronbach); (4) the understanding and development interests of direct service staff and affiliates (Stake); and (5) the democratic and social change interests of program beneficiaries and their allies (democratic evaluation).

TARSILLA: Jennifer, on the one hand, you seem to maintain a certain equidistance from each of the five values discussed above. I also had the impression that you do not expect evaluators to
Michele Tarsilla

promote all of such values in their respective work. What is your reaction to that?

GREENE: Well, while these five values are all legitimate, some of them are not necessarily compatible with each other. Most evaluation practitioners adhere to one or some logical combination of these genres and the values they advance in their own practice.

TARSILLA: On the other hand, though, some of your work seems to privilege one particular value (i.e., democratic and social change) among the five which discussed above. How do you respond to that?

GREENE: I would not say that I privileged this or that particular value. In my opinion, the active choice for evaluation practitioners is usually not which values to promote in their work. The active, conscious choice is whose questions to address in this evaluation. Values do accompany these choices but choices do not seem to be consciously about values. Rather they are more consciously about evaluation purpose, audience, and key questions to answer.

TARSILLA: Jennifer, I particularly like your call to evaluators for “unmasking” both the influences and value stances which underlie their work but are too often overshadowed by the adopted methodology. You say: “It is time for evaluators to explicitly state the value commitments, programmatic assumptions, and political stances that underlie [your] chosen methodology.” However, my impression is that today methodologies, for example, randomized controlled trials (RCTs), as well as evaluation questions are increasingly prescribed by funders (e.g., via detailed request for proposals; RFPs), thus making it extremely hard for evaluation strategies to reflect evaluators’ personal values. Could you please comment on that?

GREENE: Yes, I see your point. It does happen indeed that ToRs [terms of reference] are quite detailed these days and often leave evaluators very little room to decide the methods to use or the evaluation questions to address. That is why I desist from accepting assignments whose suggested methodology I do not really feel supportive of. However, I am of the opinion that, as the contract is being negotiated, evaluators can strive to refine clients’ ideas on what the best methodology or tools to adopt should be. In other words, as highlighted during a seminar which I participated a few years ago, discretionary space for negotiating evaluation purpose, audience, questions, and then methodology still exists for the evaluator.

Theory of Knowledge

As knowledge is mediated by evaluators’ perceptual frames (comprised of experiences, interests, theoretical understandings, values, and beliefs), Greene (1997) states that an unfiltered (objective) view of the world is not feasible. In other words, “wholly disinterested knowledge claims are unattainable” (Greene, 1997, p. 27).

Social knowledge is both propositional and constructed (Greene & Caracelli, 2003). As such, Greene states that knowledge at times is the product of a continued and irreconcilable tension between popular paradigms (e.g., objectivity-subjectivity, inductionism-deductionism).
TARSILLA: Jennifer, if knowledge and understanding of the world are socially and culturally constructed, what is your personal view of RCTs which, by definition, reject constructivists’ ideas and aim at providing rather “objective” and scientifically sound evidence on social programs?

GREENE: Well, I acknowledge the merit of having a plurality of methods, methodologies, and paradigms to pick from when approaching an evaluation. I have deep respect for other ways of knowing. Although I do not particularly ascribe to the RCT approach (or rather to the ridiculous expectation of applying it to all situations and different types of program), I appreciate the principles of causality, generalizability and replicability underlying the work of those among us who use this specific approach. The RCT contribution is especially valuable if compared with the long established evaluation practice of focusing on short-term outputs. However, I also acknowledge the challenges of doing randomization in the real world, especially when the application of such method compromises the right to access services and goods among the most marginalized population groups living in the communities where RCTs are being conducted. Likewise, I can’t help recalling Campbell’s idea that randomized experiments should be used mainly for pilots or particularly innovative projects (and not systematically for existing programs) and that their findings should be used for scaling up demonstrable successful interventions. Evaluate with RCTs, said Campbell, only proud programs, that is, innovative programs that have demonstrated quality in prior studies.

TARSILLA: Jennifer, due to the increasing number of evaluators around the world and the growing emphasis on cultural competence as one of the evaluators’ most valuable tools, a diverse set of new “regulative ideals” seems to proliferate globally. In your opinion, is the ongoing globalization of the evaluation profession contributing more to the fragmentation or enrichment of the knowledge construction process?

GREENE: Globalization is definitely enriching the knowledge construction process these days. Regional evaluation associations, for instance, are allowing a growing number of evaluation professionals to contribute to the shaping of the current evaluation discourse which traditionally has been an exclusive domain of the West (United States, Canada, and Northern Europe).

Theory of Social Programming

Greene’s theory of social programming has been particularly influenced by a variety of evaluation approaches: democratic evaluation (House and Howe, 1999; MacDonald, 1978), fourth-generation evaluation (Guba & Lincoln, 1989), morally engaged evaluation (Schwandt, 1989, 1992) and participatory evaluation (Reiben, 1996; Weiss & Greene, 1992; Whitmore & Kerans, 1988).

Greene views evaluation as a force for democratizing public conversations about important public issues (Greene, 1997). Evaluation—Greene states—should work to enable full participation of all legitimate stakeholder interests in the conversation and in all relevant decisions about a particular program’s merit and worth, with democratic principles of
equality, fairness, and justice as guides to both the conversation and the decision-making.

In Greene’s opinion, the commitment to democratic pluralism constitutes the regulative ideal for this vision of evaluation and becomes its essential value commitment. As a result, evaluation and the program can work in concert to help democratize the conversation about equitable health care for the elderly, about generational and spatial destitution, about kids killing kids, about a safe and adequate food supply.

TARSILLA: Jennifer, since the poverty rate is on the increase and evaluations should have contributed to its reduction over the years, to what extent is it fair to assert that evaluators have failed in attending to their “social” role over the last decade? Is it anyone’s responsibility? Is such responsibility to be found outside of the “evaluators’ circle”?

GREENE: According to House, evaluators have the responsibility to alleviate social problems. Evaluators have been expected to contribute to the betterment of living conditions around the world. However, the dire data showing that many social problems still exist and that these are far from having been solved, certainly leads to the conclusion that evaluators have some responsibility for this. Evaluators are partly and not fully responsible for this, though. In my view, we evaluators are underselling our potential: that is where our responsibility lies. However, as we also provide services to decision-makers and these are really the ones driving the public agenda on almost all fronts, part of the responsibility is found at the political level.

TARSILLA: Jennifer, with the change of the political climate in the US over the last year, do you already see any substantial difference in the way evaluators are carrying on “public conversations about important public issues” nowadays, especially if compared with the days of the Bush Administration?

GREENE: Although I feel that some changes will be occurring in the future, I acknowledge that it will take some time before that happens. The current situation is very complex. The economic crisis is impacting all sectors. As a result, the public agenda, geared towards economic recovery more than anything else, is far from giving evaluation the importance which it deserves. Even so, some things are being done these days to facilitate the inclusion of evaluation in the political agenda, though. I particularly praise the work currently being done by the American Evaluation Association (AEA) Evaluation Policy Task Force.

Theory of Use

Although the evaluator’s relationship with program staff could be highly collegial, Greene accepts that evaluators have no authority or responsibility for the program design and implementation. Therefore, in Greene’s thinking, evaluators should not push for any particular program change. In other words, if the program staff decide to keep the program as it is, the evaluator should respect that decision. In such instances, the evaluation’s merit is to keep the issues alive in the ongoing evaluative conversation and to encourage repeated and ongoing reflection and analysis among those who run the program being evaluated.
Influenced by Lee Cronbach, Greene also believes that it is critical for an evaluator to understand the full contextual complexities of programs being evaluated (Greene, 2004) before using the evaluation to educate the “Policy-Shaping Community” (Mathison, 2005, p. 95). Hence, Greene frequently uses the term “educative evaluator” to describe the evaluators’ role.

In relation to a Kellogg Foundation evaluation she conducted, Greene reiterates that evaluators’ roles and responsibilities have boundaries and that she did not push for any particular program change in the course of that specific assignment. Greene indeed believed that neither she nor any member of her team had authority to do that (Fitzpatrick, 2001, p.87). In her interview with me, Jennifer Greene expanded on this topic further and reiterated that the evaluation function is quite distinct from the design and implementation of any program.

TARSILLA: Jennifer, to what extent do evaluators have authority or responsibility for the design and implementation of programs?

GREENE: We as evaluators have no right or responsibility to tell program managers to continue or suspend a program. The best we could do is to collect defensible data and use that to show what an intervention has been capable of attaining (or not attaining). Program managers in most case have more expertise about the field than evaluators. For example, if I were doing the evaluation of a school program, I would be very cautious about recommending the continuation or the suspension of the activities I have evaluated. Primarily, because I do not teach there. I do not have the right to tell other people what to do in their own program. What I could rather do, though, is to be a catalyst and facilitate discussion among those who have a vested interest in the program (almost like Patton does in his developmental practice). With respect to that, I would not say that the dissemination of findings is the only phase when such conversations should take place. Obviously, the conclusion of an evaluation is a very important moment to discuss with the client and all the other stakeholders what the next steps should be. However, I strongly believe that, for this public discourse to be successful, evaluators need to lay ground for communication and mutual sharing beginning with their first contact with the client and the rest of the stakeholders.

Greene’s focus is quite holistic. Her work includes outcomes, along with the content and process of the program. As a result, while the core of the evaluation may be to judge merit and worth (Scriven, 1967), Greene emphasizes the importance of understanding the process and program improvement. As a result, her evaluation purpose seems to be formative, rather than summative, and her primary audience for this formative information seems to be the program staff (Fitzpatrick, 2001, p. 95). Rather than answering particular questions, Greene attempts to bring about change, not simply use, by establishing an “ongoing evaluation conversation about value-based issues important to the program” (p. 95).

TARSILLA: Jennifer, how would you respond to those who see your evaluation work as mostly formative?

GREENE: I do not feel that this is an accurate description of the evaluation work I have done in all these years. Data on program outcomes can only be...
meaningfully interpreted and understood if there are also high quality data on program design and implementation. Therefore, my work does not necessarily privilege formative evaluations per se but rather as comprehensive evaluations (Shadish, Cook, and Leviton) in the legacy of Lee Cronbach. The Kellogg evaluation referenced in Fitzpatrick, for instance, had both program improvement and outcomes attainment questions for that evaluation.

TARSILLA: Jennifer, in the executive summary of your evaluation of the Natural Resources Leadership program, you stated that the program’s aim was “to improve the management of natural resources, and to enhance rural economic development while maintaining or improving environmental quality,” and that “the program was successful in realizing its aims.” However, only a few participants actually benefited from that specific program. In this case, where is the borderline between the search for public good (democratic pluralism) and the good of very few marginalized individuals, as in the case of the Kellogg Foundation program?

GREENE: Well, public good is not simply the greatest good for the greatest number of people. That is quite an instrumental vision of such principle. I strongly believe that pursuing the public good should focus on taking care of those who are at the margins of society. Going back to the example you asked me about, I recollect that although the program objectives had been fulfilled, the findings from that evaluation were not fully exploited as the program staff whom my team and I worked with were mostly mid-level professionals and therefore they did not have decision-making power within the organization. Retrospectively looking at it, we should have sought the participation of the latter to a much larger degree.

**Theory of Practice**

Greene’s approach to evaluation is responsive, in Bob Stake’s sense of the word, in that the evaluation tries to become attuned to the issues and concerns of people in the context of any given program (Fitzpatrick, 2001). Greene ends up declaring that advocacy in evaluation is inevitable (Greene, 1997) and, as a result, she predicates that evaluators should play a socially enfranchising role today (Greene, 1997):

> Evaluators today should make special efforts to seek out and include those voices and perspectives that are often overlooked or excluded from the evaluative conversation—because they are not invited, because their views are unpopular, or because they don’t have the language or verbal fluency to be heard. Often excluded are the least enfranchised groups and individuals within a program setting, who are usually the intended participants in a social program, their families, and their surrounding communities (p. 29).

The evaluator—Greene says—is a sort of public scientist who should participate and be engaged in and not distanced from public affairs, someone who should strive to contribute to and not remain insulated from discussions and actions about public issues. Greene adds that evaluators need to establish open borders with the program being evaluated, borders that invite multiple crossings and visits (1997).

In order to enhance the pluralism in their own work and make multi-interested
evaluative claims, evaluators could also adopt participatory methods and facilitate the dialogue among program staff and participants toward shared understandings and collective actions (e.g., Pursley’s participatory evaluation of a network of four family support centers located in an upstate New York or the cluster evaluations conducted by the W.K. Kellogg Foundation). According to Greene, the evaluator, for instance, should always share his/her write-up with the client and get the feedback of the involved stakeholders along the process.

While Greene (Greene & Caracelli, 2003) acknowledges the relevance of paradigms in evaluation, she also reaches the conclusion that what really drives evaluators’ inquiry is not the orthodox application of this or that specific methodology but rather the creative mix of methods dictated by a concrete need for information intimately dependent on the context where the evaluand is found: a mixed method way of thinking (Greene, Benjamin, & Goodyear, 2001). This mostly consists of a methodological eclecticism where the primary concern is fitness for purpose.

A strong supporter of mixed methods (Greene, 2006, 2007), Greene strongly believes that the combination of qualitative and quantitative methods would greatly enhance the successful completion of any evaluator’s assignment (e.g., through mini case studies on success participants’ success stories, interviews with key program staff and semi-structured surveys). In Greene’s words, “scientifically rigorous plans” should be accompanied by the use of less conventionally rigorous approaches which still have value in many circumstances (2004). However, evaluators should not place excessive emphasis on the methodology since the merit of an evaluation—Greene agreed with Cronbach on this—is determined not by the form of the selected inquiry but by the relevance of the gathered information.

In conclusion, Greene adds that it might be particularly beneficial for evaluators to use a reflective methodology. Greene says that a continuing cycle of reflection and analysis should characterize every evaluator’s work. Evaluators should always be able to answer the following questions based on their work: What have we learned? What does it mean? What do we as evaluators think about this? What do stakeholders think about it? In order to do that, some kind of partnership and collegiality with the stakeholders should be established. By doing so, although not a member of the program team, the evaluator would not end up being an outsider (Greene, 2001). Evaluators could even consider including a section on “Evaluators’ Insights and Analysis” in their final report and this might provide them with the opportunity for fruitful exchange with other experts in the same area as the program which they evaluated.

TARSILLA: Jennifer, a large number of evaluators have espoused the idea of using mixed methods in their practice over the last decade. How do you personally see this idea being put in practice? What are the weaknesses (if any) of the current evaluations which allegedly use mixed methods?

GREENE: I agreed with you when you say that mixed methods are widely used in evaluation. Using mixed methods makes sense and responds to the needs of practical and methodological people who are not entrenched in the use of one particular approach. The conceptual development of mixed methods then
simply followed the practice. That having been said, I admit that the utilization of mixed methods is still quite limited. What I see is lacking today the most in the use of mixed methods is the clear identification of the reasons why this specific method is being used as well as the intentions or objectives which its use is supposed to fulfill. While I praise the use of mixed methods in evaluation, I am also cognizant of the fact that they are not always the preferable methods to adopt in evaluation. That is the case, for instance, of evaluation with very limited resources or limited expertise in both qualitative and quantitative methods. The content of the evaluation could also influence the use of mixed methods: mixed methods, for instance, are not usually adopted in the case of health evaluations mostly interested in addressing questions of epidemiological nature (e.g., the incidence of a particular disease) and not in investigating the way people are either accessing medical services or dealing with their specific illnesses.

TARSILLA: Jennifer, could you please provide an example of a particular evaluation you worked on in which the use of mixed methods really showed the added value or comparative advantage of such approach over the simple use of either qualitative or quantitative methods?

GREENE: Some recent work has been published which could provide some very good examples of mixed methods application. One is a book written by Thomas Weisner titled Discovery of Successful Pathways in Children’s Development. The other is an article in this book written by Debra Skinner (first author) which shows the successful use of GIS (geographic information systems) and traditional qualitative research methods.

TARSILLA: Jenifer, of all the qualitative research methods currently in use, is there any emerging one which you would urge young evaluators to become familiar with?

GREENE: I can’t think on any emerging tool at the moment. However, what I find really interesting is the type of mixes of already existing tools that have been done in the past in order to address evaluation questions effectively.

References


