

Orbiting the Big Dog

American Journal of Evaluation
2016, Vol. 37(1) 139-147
© The Author(s) 2015
Reprints and permission:
sagepub.com/journalsPermissions.nav
aje.sagepub.com



Donaldson, S. I. (Ed.). (2013).

The Future of Evaluation in Society: A Tribute to Michael Scriven. Charlotte, NC: Information Age Publishing. 181 pp. Hardcover: \$85.99, ISBN 978-1-62396-4528. Paperback: \$45.99, ISBN 978-1-62396-4511.

Reviewed by: Linda Mabry, *Washington State University, Vancouver, WA, USA*

DOI: 10.1177/1098214015579910

The 2011 Stauffer Symposium at the Claremont Colleges was intended to honor Michael Scriven for a half century of unflinching and unstinting leadership in evaluation. *The Future of Evaluation in Society: A Tribute to Michael Scriven* is the published record of this gathering. But, as editor Stewart Donaldson tells it, the guest of honor refused to attend until the spotlight turned to the future of evaluation. Scriven only partly succeeded in adjusting the focus: In this volume, the contributing authors extravagantly admire the title character before offering their visions of evaluation's prospects.

Under no obligation to recount his own contributions to a field that “gives us more to be humble about than to be impressed by” (p. 11), Scriven's own chapter is a dazzling demonstration of his worthiness of the honor he tried to blunt. Both drawing together and expanding upon ideas developed over a prolific career, he offers a “cosmology of evaluation” (p. 15), its logical and epistemological nature, star wars and revolutions, and paradigm shifts—some still in process.

No one acquainted with his work can be surprised that Scriven closes his own chapter by inviting “criticisms, constructive or not, and the chance to improve my arguments and conclusions” (p. 41). But the book's co-contributors, rather than responding to the provocative view he articulates, acknowledge his vast and seminal body of work and sketch overlaps with their own efforts—connections, contributions, challenges, and disagreements honed over time.

Paradigm Shifts in the Cosmology of Evaluation and Science

With the caveat that, for most practicing evaluators, theoretical arguments are destined to fail the *So what?* test, Scriven contends that “there are plenty of evaluation theories” so exclusively aimed at methodology that they neglect to “discuss the foundations of the subject” (p. 16). Remedially, he guides readers on an historico-epistemological tour of his envisioned cosmology of evaluation as he considers a series of paradigm shifts leading to evaluation's acceptance as a distinct discipline in the universe of science, its emergence as an alpha or transdiscipline, and its progress toward becoming scientific inquiry's exemplar.

Framing Paradigm

Gazing into starry skies, Scriven spots a nova: the skepticism of antiquity bursting into empirical science. Thus released, bits of space debris speed off on trajectories that take them through a series of scientific revolutions. To trace evaluation's torturous path, Scriven subdivides Kuhn's (1962) notion of paradigms, distinguishing between a *framing* paradigm regarding the nature

of knowledge and a *scaling* paradigm regarding the complexification of knowledge over time. Evaluation's framing paradigm shifted when the field was accepted as a respectable scientific endeavor, but the shift in its scaling paradigm will remain unaccomplished until evaluation's multifaceted nature is better recognized.

Scriven puts evaluation's first paradigm shift in the context of the scientific revolution that redefined truth: ancient Greek empiricists disentangling knowledge from myth; Renaissance astronomers distinguishing it from religion; and, still orbiting Aristotle, 18th-century logical positivists insisting that facts be verified. Just as Pluto was nudged out of planet status in 2006, the positivist view of knowledge and facts through the 19th century "banned evaluation from scientific legitimacy . . . [an] analog of the Inquisition's attack on the emergence of the new astronomical cosmology . . . [that] was completely mistaken" (p. 19). In the 20th century's consideration of qualitative approaches to inquiry, Max Weber and others implicitly rejected the positivist perspective of a value-free science—but it was another half century before evaluators joined the debate.

Positivists rejected unconfirmable "evaluative facts," failing to note that "*factual* in common usage just means well-supported, or supported beyond reasonable doubt" (p. 23). To join the scientific oeuvre, evaluators needed to show that value-laden criteria of quality were more than "simply a matter of taste or preference" (p. 18). This they did by demonstrating the "feasibility of objective evaluation" in studies of "major educational changes . . . including federal programs like Head Start" (p. 20). Or did they? Even today, the astral dust remains unsettled. Scriven mentions a recent meteor shower that left craters in the evaluation community: the U.S. Department of Education's (2005) funding preference for quantitative and experimental designs, presumably because of their generalizability and supposed objectivity; the American Evaluation Association (AEA)'s (2003) protest when the government proposed its funding preference; and disavowal of AEA by some of its quantitative-leaning members.

Scaling Paradigm

Since its acceptance as legitimate scientific inquiry, evaluation has faced a new challenge. Just as astronomers have come to the "realization that the sun is merely one of billions of stars" (p. 14), Scriven urges his colleagues to shift their *scaling* paradigm. Doing so would involve recognizing "a very diverse lot" (p. 12) of constellations of evaluation practice, rather than telescoping our collective vision to *program* evaluation alone. Program evaluation should not even be prioritized, he argues, as this would merely be "one step less parochial" (p. 12). He notes neglected asteroids: "the 7Ps, the newbies, REDs, SEDs, PECs, the mighty evaluative disciplines of ethics and logic," "the *inferential disciplines*" (pp. 33–34, emphasis original) both academic and nonacademic.

Evaluation as Transdiscipline, Alpha Discipline, and Exemplar

The revolution thereafter would involve recognition of evaluation not only as a multifaceted field but as a "transdiscipline" (p. 13). Scriven introduced this term two decades ago to mark evaluation as both a discipline in its own right and one of the "important tools to other disciplines" (p. 36), those "disciplines that have an evaluative half—particularly logic, ethics, medicine, and engineering" (p. 22).

A "third Paradigm Shift" (p. 38) would elevate evaluation to "alpha" status whereby evaluators determine the acceptability of research claims in other fields. Because "science is not living up to the minimum standards" (p. 37), Scriven predicts that, with higher standing, evaluation will be called upon for quality control purposes in the peer review of scientific research.

In a fourth paradigm shift, evaluation would break the boundaries of science, becoming the "model for all the applied disciplines—the social sciences and well beyond" (p. 39).

Ethics, the Final Paradigm Shift

At the end of the series, Scriven predicts an interstellar collision. Having explicitly tracked the paradigmatic trajectory of evaluation from the original scientific nova, he implies a simultaneous journey for values: From the distinction between facts and values, to the acceptance of value-ladenness in science and in evaluative conclusions, to recognition of ethics as an alpha value.

Peering into the etymological heart of *evaluation*, values, Scriven asserts: “We all pay lip service to the idea that ethical values are the alpha values . . . ‘ethics’ here refers to the system of ethics based on the single axiom that all humans have prima facie equal rights” (p. 40). Or is this mere stipulation, Scriven’s personal ethical value? He seems resigned to a contradiction, his certainty that universal human equality is the primal ethical value alongside his awareness “that what we call ‘the ethical values’ are simply *our* cultural values” (p. 40, emphasis original).

Thus, our field’s premier logician has reached two logical ends, namely, evaluation as an alpha discipline and ethics as an alpha value. It is at the point of ethics that the values trajectory and the evaluation trajectory intersect. Here, Scriven sees evaluation giving way, its “deflection of the primary effort into moral education” (p. 40). He describes this future event as a paradigmatic happy ending, a grand final “shift of all educated people to a real (not merely professed) belief in that value” (p. 40).

Futures

In this way, Scriven’s view of the future crystallizes historical, theoretical, and philosophical thoughts drawn from a lifetime of practice. The book’s other contributors similarly draw from work that has claimed their minds and marked their careers.

In keeping with the topic, most contributors predict evaluation’s continuation or expansion. Melvin Mark sees continuing debates and growth into new areas and methodological understandings, the incorporation of new expertise—perhaps a prediction of our scaling of the scaling paradigm. Christina Christie sees linguistic refinements in a furthering of Scriven’s *Evaluation Thesaurus* (1991), leading incrementally toward clearer policy. Dan Stufflebeam hopes for more beneficial impact, placing his confidence in the continuous refinement of standards and training. Karen Kirkhart sees a stretching of Scriven’s Key Evaluation Checklist to include cultural competence and cultural validity. Stufflebeam and Michael Quinn Patton particularly note Scriven’s Faster Forward Fund to accelerate progress in evaluation.

Patton urges more serious attention to timeliness and unintended consequences. Even more, he pleads for evaluation to grasp more complex understandings—outward expansion to encompass dynamic and intertwined contexts at the systemic and global levels, and inward expansion to recognize the personal and psychoemotional factors in understanding and using evaluation results. Ernie House and Jennifer Greene take sociopolitical perspectives, Greene seeing evaluation as working toward democracy, House grimly observing vested interests taking control of accountability and, with it, evaluation.

Two authors take a dismissive stance. Robert Stake contradicts Christie’s view of lexicon expansion claiming “the practice of evaluation does not depend upon the language used by experts” (p. 109), and undercuts the group’s enterprise claiming “the future is likely to be a repetition of the present” (p. 113). Rodney Hopson fills in miasmatic, contradictory details: “You will not be able to avoid the usefulness and ubiquity of evaluation [or] . . . to mislabel, misappropriate, misconceive, misapply, or misuse evaluation . . . or avoid instances of bias and conflicts of interests” (p. xv).

Orientation to People

These contrasting views of evaluation and its future involve or imply differing orientations to people, individually or collectively. While Scriven’s prosocial premonition that evaluation will

lead to ethics is offered at an airy philosophical level, the other contributors take positions closer to the ground.

Some focus on practitioners. Mark addresses what has been called *the guild* (Smith, 1998), foreseeing growth in the scope of the evaluation community's work, its rejection of governmental attempts to dictate its methodologies, its collective reflections in the form of debates and evaluation research, and its welcoming of new ideas from "boundary spanners" (p. 172), pointedly including Scriven himself. Kirkhart urges that the expansion include cultural sensitivity and competence.

Other contributors orient toward those stakeholders who implement findings (or not), implying a degree of managerialism. Stufflebeam laments the paucity of positive impact on persons and society and urges redress. Patton identifies systems and globalization as part of the macro-context of evaluation work, urging both broader and deeper attention. More explicitly managerial, he continues his long focus on the decision makers (or what he has called *primary users*) best positioned to make immediate use of evaluation results.

As noted, Greene and House look toward sociopolitics. Curling up from the human relationships she sees as critical for comprehending evaluands, Greene predicts that deepened human sensibilities will disperse as democratic tendencies. House's less positive view of evaluation's future draws on analyses of societal exploitation. On the basis of geopolitical trends, he predicts the hijacking of evaluation by vested interests as part of a subversion of regulation and accountability, masked to avoid public resistance; perhaps Hopson agrees. Set in the volume's generally optimistic context, House's view stands out, nearly as shocking as Machiavelli's (1532) clear-eyed realpolitik.

Perhaps in modesty, perhaps taking a postmodern turn, some contributors would delegate all or parts of the evaluation task to the untrained, a matter about which the book offers some implied arguments. Christie would leave the actual *valuing* to others, presumably those with access to evaluators' analytic results—a stance she notes that Scriven has disapproved, deeming it barely construable as professional. Stake goes farther in leaving the entirety of the evaluation endeavor to "all the people of the world" (p. 108), taking a positive view of informal evaluation—which, in an earlier chapter, Scriven dismissed as pre-scientific, pre-professional.

Internal Debates

Thus, before the book's covers can be closed, the debates Scriven invites and Mark forecasts are implied within its chapters. The contributors take up points that have found and eluded general agreement over time, offering glimpses into the thinking and history of their community of practice. My overview-plus-commentary task similar to Mark's, here I will snippet in my own arguments.

Evaluation Logic and Process

Scriven's tracing of paradigm shifts in evaluation significantly employs logic. Pitting Hume's positivism against Weber's interpretivism, Scriven painstakingly dissects the notion of *science* and its early exclusion of evaluation. Etymological arguments, centered on *facts* and positivism's insistence that they be *value-free*, involve his "shanghai" (p. 26) of the term *colligation* to denote the kind of thinking that renders evaluation claims sufficient to include evaluation in the lineup of sciences.

What of this, in the grand scheme of things? Does this stretching of our vocabulary qualify as a useful expansion of the lexicon as predicted by Christie, or irritate Stake as a widening of the gap between formal and informal practitioners, or merely annoy me as unnecessarily abstruse? That may be mere space dust, but there are also comet crashes. Is Greene alluding to Scriven's logic when she scathingly chides, "technical rationality [is] *not* the solution to the challenges of independent, credible, and fair evaluation practice," arguing instead that, "for evaluations of consequence, I would rather rely on a person with whom I can relate" (p. 124).

My argument regarding Scriven's logic involves his claim that we have long since shown the "feasibility of *objective* evaluation" (p. 20, emphasis added), a key point in his refutation of positivist insistence on verifiability. For me, the psycho-epistemological weakness is this: Since human powers of reasoning reside in *subjective* minds, no one is capable of *objectivity*—not the positivists, not the scientific community (who may agree but err), not even Scriven. Even if we encountered objective facts, we could recognize them only subjectively—and, therefore, idiosyncratically, such that only some (if any) of us could recognize them at all. I take this point to be the crux of Kuhn's (1962) argument regarding scientific revolutions and the historical overturnings of so-called objective facts or knowledge. Scriven, in choosing to use Kuhn's term, *paradigm*, inadvertently offers the pin that punctures his own balloon.

While he is attempting to inflate it, Scriven introduces a problematic distinction. On one hand, he calls attention to *inferential evaluations*—program and product evaluations that require "extended observation, testing, and precise measurement of crucial criteria . . . evaluation there is typically inferential" (p. 27). On the other hand, he claims there are *perceptual evaluations*—"although taxonomy usually involves complex lists of criteria to be checked and eliminated, it's sometimes a one-step recognition, especially for the professional. Evaluation is the same" (p. 27).

To float the distinction between inferential and perceptual evaluations, Scriven must convincingly show that, in the latter, perception and evaluation can co-occur. He attempts this with an example: an expert mango buyer who can identify ripe fruit at first glance, making a judgment in which perception and evaluation are inseparable. My argument is that inseparability is a temporal (and subjective) illusion. Maybe only an imperceptible nanosecond separates the *recognition* of the evaluand and the *assessment* of its quality, such that the evaluator himself or herself does not notice the disunity, but there is nevertheless a sequence.

These are nebulous quibbles, no doubt, to hardworking practitioners sorting through a galaxy of data, rival interpretations of it, and interactions with clients and other stakeholders. And even to me, they seem a bit off course since I, too, navigate largely by Scriven's star.

Accuracy and Utility

Patton and Stufflebeam discuss an issue that finds consensus among these (but not all) evaluators: the relative priority of *accuracy* and *utility* as evaluation standards. Stufflebeam, who exercised longtime leadership in the development of *The Program Evaluation Standards* (Joint Committee, 2011), recalls that the measurement principles of *validity* and *reliability* were uppermost in the minds of early practitioners during the emergence of the profession a half-century ago—and that such thinking naturally prioritized *accuracy* over utility. He observes that this precedence has gradually reversed in theory but not in actuality.

Scriven's role in this reversal is illuminated by Patton, who has championed *utility* in four editions of *Utilization-Focused Evaluation* (2008, 2012): "Scriven has consistently argued over the years that the evaluator's first obligation is to produce truth (so that it can be spoken to power)" (p. 53). But in the long deliberations during the most recent revision of the *Standards*, Scriven reasoned that prioritizing *accuracy* would "return evaluation practice to the days, especially in the 1960s and 1970s, when evaluators produced many technically sound reports that only gathered dust on shelves" (p. 52).

The *Standards'* continued positioning of *utility* as first in the sequence of four categories of standards represents a triumph of practical wisdom over logic. Accuracy is logically prior—appropriate *use* can hardly be predicated on *inaccuracy*—but nonuse has proven a chronic sore point professionally. The vexation finds voice here in Stufflebeam's main concern and his wish for the future: "our evaluations have resulted in too few beneficial impacts" (p. 83).

Perhaps these authors are, for the moment, presuming *accuracy*, since their chapters follow Scriven's discussion of the clash between *evaluation claims* and the *facts* required by a positivistic science. But elsewhere, many of the contributors have shown themselves quick to throw down evaluations they deemed inaccurate and evaluation approaches they considered likely to lead to error (e.g., Stufflebeam et al., 2001). For example, Stufflebeam (1994) has slammed David Fetterman's well-known *empowerment* approach (Fetterman, 1996); Stake (1986) has smacked Charles Murray's numbers-driven evaluation of a federally funded education program called "Cities in Schools;" and Scriven is quoted in this volume as calling Christie's willingness to delegate evaluative conclusions to stakeholders mere "foreplay" (p. 94) rather than good practice.

The other chapter contributors pick up on utility, impact, empowerment, and disempowerment, but none focuses directly on ethics, the *propriety* standards, where Scriven eventually arrives.

Competence

Professionalism is more prominent than ethics in their chapters. Although a desire to improve accuracy and competence drove development of the *Standards*, finding agreement about evaluator competence has proven at least as slippery as determining whether evaluation is science, as Scriven argues it is. Competing perspectives on this issue long delayed even an official definition of *evaluator* by the U.S. Bureau of Labor Statistics and blocked credentialing efforts repeatedly proposed to the AEA by various members—and established by the Canadian Evaluation Society. For example, the AEA's board of directors declined in 2003 to give its official approval to a proposed public statement regarding evaluator competencies so as not "to convey or imply approval of a proposal ultimately intended to develop credentialing criteria or procedures for either licensure or accreditation" (AEA board response to proposal, August 11, 2003).

On that occasion, the board argued that the purpose of the organization's public statements, the development of which was supported by a grant from the National Science Foundation, was "to develop statements on topics of *public* importance, while the [evaluator competencies] proposal appears to address a *professional* issue" (2003, emphasis added). Yet, a later board *did* approve a public statement regarding an identified subset of *cultural* competencies (AEA, 2011), a professional subcategory of the rejected 2003 proposal, whose incorporation into practice is promoted in this volume by Kirkhart.

Star Wars

Internal politics is, however, a small matter in the context of Scriven's cosmology. More imperative is House's concern about the corruption of evaluation by vested interests, about which Scriven's logical tour is silent. Scriven's philosophical prophecy has the appeal of logic carried to its natural, nearly irrefutable, end. Its gravity is undeniable. It's as if Plato's idealism and Aristotle's empiricism were finally reconciled, as if Einstein's (1916/1920), Hawking's (2001), and Hawking and Mlodinwo's (2010) quest for a theory of everything was in sight.

House pilots readers to a different world, one in which autonomous humans relentlessly pursue their own interests, hide the space junk they create from the prying eyes of evaluators and their ilk, consolidate and incorporate and globalize behind nebulae of hype. Plato and Aristotle are irrelevant in this moonscape, more recognizable to Sartre (1960) and Machiavelli (1532). From war zones, House reports,

Biases are deliberate. . . . many drug evaluations are being designed and conducted to deliver positive findings. . . . Vested interests have been shown empirically to influence medical research findings. (p. 65)

The financial event of our time has been the collapse of the markets in 2008 . . . In their assessments, the professional bond raters . . . realized there were serious issues. But . . . conflicts of interest biased evaluative judgments. (pp. 66 and 68)

Evaluators have their own material interests weighing significantly on one side of the outcome. Traditional research methods alone are insufficient to control such biases. (p. 69)

House is no Marx (1867/1996), no Lenin (1899/2004) inciting proletarian or professional insurrection. It's not just that Hopson's revolution will not be evaluated, it's that House foresees no revolution, no exit (cf. Sartre, 1944). Eclipsing Stufflebeam's confidence in standards and Scriven's thoughts of transdiscipline, House offers but a single hopeful possibility, in which he himself expresses little or no confidence:

All this requires action by the organized evaluation profession. Ultimately, we have to exercise serious professional review. We are past the era of stating standards for evaluators to follow and hoping for the best. Countervailing pressures are too strong. . . . [Scriven's] grand vision for evaluation presumes that the field of evaluation itself can maintain its own honesty and integrity by withstanding the corrupting forces . . . no easy task. (pp. 70–71)

In his work, Scriven has taken on the dark side (e.g., Rothman, 1991) but says nothing of this here. The earthly wars that agonize House do not figure as star wars in Scriven's cosmology. No common ground exists between Scriven's glimpse of the future and House's, no overlap between their starkly different universes, no shared issue about which they offer opposing views. There is no climactic debate. In the biggest intergalactic crash of the volume, Scriven has little to say about the societal climate change occurring below his theoretical clouds and, drowning in political–professional calamity, House can barely spare a paragraph for abstract logic.

Instead, they and their coauthors peer penetratingly into separate—completely separate—crystal balls and toss them into the readers' space. This is a natural result of separately conceived chapters, but not an entirely satisfying one. Perhaps at the symposium, speakers interacted; if so, some presentation of their discussions might have enriched their chapters. Instead, their disconnected thoughts come off as futuristic bowling with crystal balls in separate lanes.

O, Wayfarer

Unable to resist a good metaphor, I imagined, from Scriven's introduction of his cosmology of evaluation, that I was time traveling in the vastness of outer space. Aloft, I marveled at Scriven's big bang portrayal of the explosion of science into human consciousness, the infinitesimally slow evolution of evaluation as scientifically acceptable, the survival of the fittest logics and methods in the meteor shower of evaluation practice and debate. I felt the gravitational pulls and tensions in the fabric of space-time that he and the other contributors present, wary of being swallowed up in House's black hole.

It is risky to travel in space, this space. These highly trained and experienced authors stepped up to a challenge Mark aptly described: “Making predictions is easy. What is difficult is making well-founded predictions that turn out to be reasonably accurate” (p. 161). Different as their forecasts are, some of them will inevitably miss their landings; they know this. It is less risky for the cosmonaut reader, of course, but good to be prepared. Being well versed in the work of each of the authors helps to locate them in the evaluation universe and to recognize the adjustments they are even now making in their trajectories.

But agreeing on one consensual point involves no risk whatever: Michael Scriven is the hypergiant star—Canis Major, the big dog—spanning the full length of the history of modern evaluation

and illuminating more aspects of our craft than any other. We all agree with him and disagree with him; we are all in orbit around his ideas and professional integrity. The view through his telescope, and those who have clustered around it here, can be exhilarating.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- American Evaluation Association. (2003, November). *American Evaluation Association response to U.S. Department of Education notice of proposed priority regarding scientifically based evaluation methods*. Author. Retrieved from www.eval.org/doctype.htm
- American Evaluation Association. (2011). *Public statement on cultural competence in evaluation*. Retrieved from <http://www.eval.org/ccstatement.asp>
- Einstein, A. (1916/1920). *Relativity: The special and general theory* (R. W. Lawson trans.). New York, NY: Holt.
- Fetterman, D. M. (1996). *Empowerment evaluation: Knowledge and tools for self-assessment and accountability*. Thousand Oaks, CA: Sage.
- Hawking, S. (2001) *The universe in a nutshell*. New York, NY: Bantam.
- Hawking, S., & Mlodinow, L. (2010). *The grand design*. New York, NY: Bantam.
- Joint Committee on Standards for Educational Evaluation. (2011). *The program evaluation standards: How to assess evaluations of educational programs* (3rd ed.). Thousand Oaks, CA: Sage. [earlier editions in 1988, 1994]
- Kuhn, T. (1962). *The structure of scientific revolutions*. Princeton, NJ: Princeton University Press.
- Lenin, V. I. (1899/2004). *The development of capitalism in Russia*. Honolulu, HI: University Press of the Pacific.
- Machiavelli, N. (1532). *De principatibus* ([*The prince*], written 1513). Florence, IT: Giunta.
- Marx, K. (1867/1996). *Das kapital (Capital: A critique of political economy)*. Chicago, IL: H. Regnery.
- Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Thousand Oaks, CA: Sage.
- Patton, M. Q. (2012). *Essentials of utilization-focused evaluation*. Thousand Oaks, CA: Sage.
- Rothman, R. (1991, September 4). NAEP board fires researchers critical of standards process. *Education Week*. Retrieved from <http://www.edweek.org/ew/articles/1991/09/04/01naep.h11.html>
- Sartre, J. -P. (1944). *Huis clos [No exit]* (trans. S. Gilbert). New York, NY: Vintage.
- Sartre, J. -P. (1960). *A critique of dialectical reason: Volume I, Theory of practical ensembles*. Paris, FR: Éditions Gallimard.
- Scriven, M. (1998, November). *An evaluation dilemma: Change agent vs. analyst*. Paper presented at the annual meeting of the American Evaluation Association, Chicago, IL.
- Smith, N. (1998). Professional reasons for declining an evaluation contract. *American Journal of Evaluation*, 19, 177–190.
- Stake, R. E. (1986). *Quieting reform: Social science and social action in an urban youth program*. Urbana: University of Illinois Press.
- Stufflebeam, D. L. (1994). Empowerment evaluation, objectivist evaluation, and evaluation standards: Where the future of evaluation should not go and where it needs to go. *American Journal of Evaluation*, 15, 321–338.

-
- Stufflebeam, D. S., Patton, M. Q., Fetterman, D., Greene, J. G., Scriven, M. S., & Mabry, L. (2001, November). *Theories of action in program evaluation*. Panel presentation at the annual meeting of the American Evaluation Association, St. Louis, MO.
- U.S. Department of Education. (2005, January 25). Scientifically based evaluation methods (RIN 1890-ZA00). *Federal Register*, 70, 3586–3589.