

Book Review

American Journal of Evaluation
1-11

© The Author(s) 2015

Reprints and permission:

sagepub.com/journalsPermissions.nav
aje.sagepub.com



Forss, K., Marra, M., & Schwartz, R. (Eds.). (2011).

Evaluating the complex: Attribution, contribution, and beyond (Comparative Policy Evaluation series, Vol. 18).
New Brunswick, NJ: Transaction. 360 pp. (Hardcover). \$52.95, ISBN 978-1-4128-1846-9.

Reviewed by: Michael Quinn Patton, *Utilization-Focused Evaluation*, Saint Paul, MN, USA

DOI: 10.1177/1098214015569758

Evaluating the Complex: Attribution, Contribution, and Beyond offers an international perspective on evaluating large-scale policy initiatives. The editors have assembled a fine collection of individual chapters from experienced evaluators but, in my judgment, *the whole is less than the sum of the parts*. The title, *Evaluating the Complex*, would be more accurately stated as *Evaluating the Complicated*. Let me explain.

Chapter 2 by distinguished Australian evaluation theorist and practitioner Patricia Rogers (Funnell & Rogers, 2011) explains the nature and implications of distinguishing the complicated from the complex:

Interventions are usefully conceptualized as having complicated aspects if they have multiple components, which need to be brought together to achieve clearly specified outcomes. These might be different elements of the program (either occurring concurrently or occurring sequentially) or processes that operate differently in different contexts. Expertise is needed in each of the components and also in the coordination of the parts. Complicated aspects also describe well the notion of “causal packages,” where interventions achieve their results in combination with other interventions or favorable conditions, or where multiple different causal paths can achieve the results. It is essential to report on these in terms of “what works for whom in what contexts,” taking into account differential results for different types of participants and in conjunction with different implementation environments.

Interventions are more usefully conceptualized as having complex aspects where they are inherently dynamic and emergent, rather than following a path that has been tightly defined in advance to achieve tightly specified objectives. Planning is done with an overall vision in mind, but it acknowledges that plans will develop as different people become engaged and as the program unfolds. Community capacity-building programs have many emergent aspects as they work by identifying needs and strengths and then iteratively planning and implementing specific actions as priorities and opportunities change (p. 35).

In the interest of full disclosure, I subscribe to and find substantial value in distinguishing the complicated from the complex in this way (Patton, 2011, 2012, 2015). Complicated interventions “have multiple components, or only work in conjunction with other interventions.” In contrast, complex interventions are “nonstandardized and changing, adaptive, and emergent in response to changing needs, opportunities and understandings of what is working” (Rogers, table 1, p. 37). Using this distinction, the cases in this book are more complicated than complex, so the more accurate title for the book would be simply *Evaluating the Complicated*.

The Introduction advocates using the complicated versus complex distinction “as the first stepping stones in managing complexity, that is, a careful assessment of actually how complex the intervention and the evaluation task is, and the design of an appropriate response” (p. 15). But then, the

distinction is largely ignored throughout the rest of the book. The evaluations presented in the case chapters are actually more appropriately understood as complicated designs rather than complex ones, but that is as it should be, for the policies evaluated are more accurately described as complicated than complex. Failure to make the distinction leads to inappropriate conclusions and lessons about evaluating complexity because the cases are not actually about complex policies. That said, this is a common failure, one noted by Rogers when presenting the distinction in the book: “[T]here is widespread use of the term ‘complex’ to refer to interventions with multiple components (which this typology would refer to as ‘complicated’)” (p. 36).

The problem is particularly egregious and exacerbated in Marra’s concluding chapter in which she uses Rogers’ definition of the complicated to define what constitutes complex policies, namely, those “which (a) set multiple objectives through an integrated approach, (b) rely on multilevel governance structures and relations, and (c) call for assessing not only results but mostly the process of implementation” (p. 315). Let me reiterate. Those are the criteria of the complicated, not the complex. Complex interventions articulate a vision but not specific objectives; have no integrated approach or formal governance structure but self-organize as the work emerges; and both processes and results are nonstandardized, contextually variable, changing, adaptive, and emergent.

Marra further confounds the distinction by emphasizing that “the kind of policy initiatives included in this volume highlight both complicated and complex aspects. These policies are comprised of lots of parts—in line with the definition of ‘complicated’—and characterized by uncertainty and emergence—in line with the definition of the ‘complex’” (p. 316). But it is not simply that there is uncertainty and emergence that defines complexity. Even simple, standardized interventions manifest some uncertainty (e.g., variations in fidelity) and emergence (e.g., unanticipated consequences). The issue is *the degree of uncertainty and emergence*. Complexity is characterized by high and fundamental uncertainty because what will emerge is *unknowable and unspecified in advance*.

Building on this initial confusion, much of Marra’s concluding chapter makes the case for realist evaluation as the way to deal with the complexities of establishing cause and effect. Indeed, the book is preoccupied throughout with concerns about cause and effect, thus the subtitle *Evaluating the Complex: Attribution, Contribution, and Beyond*. So let’s turn to how the failure to understand and apply the distinction between the complicated and complex leads to basic errors in conceptualizing causality under conditions of complexity.

Complexity and Causality

In the Introduction, Forss and Schwartz emphasize nonlinearity as a defining characteristic of complexity (pp. 12–13) and so it is. Small actions can generate large reactions and vice versa. Complexity theory directs our attention to critical mass, tipping points, vicious and virtuous circles, and the unpredictable interactions among a large number of variables in complex dynamic systems. But nonlinearity is not more important than emergence, adaptation, dynamical (up and down) patterns, uncertainty, and unpredictability. Indeed, these characteristics operate together simultaneously, interdependently, interactively, iteratively, and unpredictably. The result: traditional notions of attribution are irrelevant. Yet they are so hard to give up. Thus, the editors and several authors take on the counterfactual problem asserting that “the question of what could have happened if the policy (program, strategy, project) had not been undertaken is always critical” (p. 24). So creative approaches to assessing counterfactuals are advocated. What is not said is this: *The whole idea of counterfactuals in complex dynamic systems is nonsense*. There are millions of possible counterfactuals. Every minute creates new counterfactual possibilities. There are simply too many unknowns interacting in unknown and unknowable ways to make talk of counterfactuals in any way meaningful. Ah, but for complicated situations, where cause and effect are separated over time and space, but knowable,

counterfactual scenarios can be constructed. Contribution analysis also makes sense for complicated interventions but complexity is simply too, shall we say, complex, to make pulling apart the contributing factors feasible and meaningful. Bottom line, if you find yourself falling back on old habits of counterfactual thinking, attribution and contribution analysis, and the search for casual mechanisms (realist or otherwise), you are in a complicated not complexity mind-set.

David Snowden is one of the leading thinkers about complexity. He was the Director of IBM's Institute for Knowledge Management and the founder of the Cynefin Center for Organizational Complexity. He has long used the distinctions between simple, complicated, and complex. At the center of his conceptualization of complexity is that cause and effect are **not** knowable in advance and can only be made somewhat coherent in retrospect; even then, whatever patterns are discovered do not repeat because they are inherently unique to the complex interactions that produced them (Snowden & Boone, 2007).

Evaluating the complex (vs. evaluating the complicated) begins with the premise that the language, conceptualization, and measurement of causality as traditionally undertaken in evaluation are irrelevant. When you find, as you do throughout this book, Herculean efforts to adapt causal inference to complexity, you can be assured of two things, namely, (1) those engaged in the endeavor are undertaking complicated analysis and thereby not dealing with complexity and (2) Hercules has met Sisyphus and can only join in trying to push the causal rock up the mountain—but even the two of them together will not succeed. It cannot be done—by definition.

Complexity Theory

Jos Vaessen, author of the chapter on “Challenges in Impact Evaluation of Development Interventions: Randomized Experiments and Complexity,” tries to solve the causal attribution problem through theory. First, the chapter reviews the strengths and weaknesses of randomized experiments (REs), but very little of that review is actually germane to evaluating complexity. He laments the polarization between adherents and skeptics of REs and strives for balance. The result, in my judgment, is the kind of balance manifest in according creationist and evolution theories equal credibility and respect. He notes the “polarization” of the debate about whether REs are the preferred method for impact evaluations, lamenting that such “polarization leads to ‘argument mining’ with proponents bringing up (arguably valid) arguments in defense of REs, while adversaries pick their favorite (and arguably valid) arguments against REs. Clearly, this is not the way forward on the path to knowledge growth.” (Vaessen, p. 307) Really? The history of science and “scientific revolutions” are marked by polar (opposing paradigm) debates as documented thoroughly by Kuhn (1970). Either the earth is the center of the universe or it is not. Either germs cause disease or they do not. Either the earth is warming dangerously or it is not. The debate is polarized because there are polar views. Either randomized controlled trials are the gold standard for impact evaluation or they are not.

Vaessen's vacillating effort to be diplomatic and balanced leads to the conclusion that REs “are complex endeavors designed to simplify complex processes of change leading to effects” (p. 308). That REs are complex is quite disputable and to claim so is to ignore most any meaningful definition of complexity. But that REs simplify complex processes is a fair and accurate conclusion. And therein lies the rub. By turning the complex into the simple, REs distort and conceal rather than reveal the dynamics of complexity.

Vaessen hopes to mitigate the substantial limitations and deficits of REs for evaluating complexity by bringing theory to bear.

REs are geared toward the question of whether an intervention works and do not shed much light on how and why interventions work. In order to look at the latter question, evaluation researchers need to look into the black box between output, outcome, and impact indicators and reconstruct the underlying

causality Two types of theory are of importance here. A reconstructed intervention theory should adequately represent the main assumptions of decision makers, target groups, and other stakeholders about causal pathways running from intervention to effects The intervention theory can be further enriched or tested by taking into account existing explanatory theories (from the social and behavioral sciences) on intervention contexts, processes of change, and potential effects

Substantive theories help to point the evaluator toward the relevant constructs and relationships between these constructs in order to make useful abstractions of the reality of a policy intervention, its intended effects, and the wider context in which it is embedded, aspects which subsequently can be tested through empirical research. Substantive theory from past (evaluation) research to some extent can help to anticipate these effects, which subsequently can be taken into account by means of additional data collection. A second important role for theory lies in reinforcing the causal analysis, the analysis of how, and to what extent changes in target variables can be attributed to a policy intervention. Relevant substantive theories can shed light on the nature of the causal relationships between interventions and (un)intended processes of change and help to rule out rival explanations for changes in target variables. Finally, there is an important role for theory in the interpretation of evaluation findings. Theory can provide a useful framework for helping us to understand why certain changes have come about or provide insights into the relevant (contextual) variables which are likely to influence patterns of results across settings. (pp. 299–300)

There's only one small problem with these high hopes for intervention and social science theory. Vaessen ignores *complexity theory*, which trumps both traditional intervention theory and substantive theory in complex dynamic situations. Complexity theory posits, as I stated earlier, that causality is not knowable in advance, can only be made somewhat coherent in retrospect, and once identified will be inherently unique to the complex interactions that produced it within a particular time and context. Complexity theory rules out randomized controlled trials as a viable method for evaluating complexity and rules out prediction-oriented, explanation-focused theories as irrelevant because complexity is its own explanation and prediction is not possible under conditions of complexity. These are clearly radical, paradigm-shifting premises, which may be why it is so tempting and easy to backslide from the complex to the complicated.

Evaluating the Complicated

I have argued that the overall framework and cases in *Evaluating the Complex* are really better characterized as *Evaluating the Complicated*. This conclusion is not meant as an attack but as a clarification. The book offers instructive examples of how to evaluate complicated interventions. The conceptual approaches to attribution, contribution, explanation, theory, working with policy makers, evaluator credibility, and rigor are all useful in evaluating the complicated. And I assert that the interventions discussed are complicated precisely because of how policy makers have formulated them: multiple partner, multifaceted, cross-sector collaborative efforts intended to affect identified outcomes through broad-based joint efforts. The political rhetoric accompanying the launch of such initiatives is all about cooperative endeavors aimed at agreed-on priorities for major systems change with accountability through specific results independently evaluated. Given the widespread failure of isolated, autonomous, siloed, and narrowly conceived project-level interventions based on linear and mechanistic change models, such large-scale, collaborative, multi-sector interventions represent a reasonable alternative strategic direction and, certainly, evaluators need to have appropriate approaches to evaluate the complicated. This book offers helpful advice and examples in that regard.

Focusing on the Evaluand

Having stated my overall judgment about the book's orientation, let me now backup and review why the editors and authors think they are evaluating the complex. Editors Kim Forss and Robert Schwartz observe in their Introduction that "doing an evaluation is more of a mess than it used

to be” (p. 7). They go on to assert that “evaluation tends to be more complex today than it was a decade ago” (p. 12). These challenges are not due to the nature of evaluation, however. They are due to the more complex policies and interventions evaluators are called on to evaluate. The editors state that the book is about “the situations where the evaluated object, ‘the evaluand,’ possesses the characteristics of complexity” (p. 12).

Sometimes the reader encounters words such as complex policy evaluation, but it is not the evaluation that is complex, it is the evaluated policy. An evaluation task may be difficult and we would argue that an evaluation of complex policies is more difficult than an evaluation of noncomplex policies, strategies, programs, and projects The complexity resides in the evaluated object and how the evaluation handles that process is a step toward understanding, recognizing, and making sense out of the complexity “out there.” (p. 13)

What, then, are the characteristics of a complex intervention or policy? Four are highlighted: (a) it is nonlinear; (b) there is an asymmetric relation between policy and management on the one hand, and impact and results on the other hand; (c) change may be fast or slow, and both at the same time; and (d) the degree of uncertainty and risk is very high. (p. 12)

Basically, the book focuses on evaluating overarching policy initiatives that involve multiple policy and program interventions that interact. Overarching policy initiatives can also be found at the supranational and even global levels in a broad range of areas, including environment, security, trade, immigration, and economic and social development. Transjurisdictional fiscal and monetary policies are now common features of overarching policy initiatives. Highly articulated strategies are, therefore, decided globally while they affect people’s lives locally. (pp. 1–2)

The Introduction traces the rise in such large-scale, multi-sector, and multifaceted interventions, including increased demand to evaluate them:

Of course overarching policy initiatives are not new. What appears to be new is a growing demand for effective evaluations of complex policy interventions at the international, national, and local levels. There is now pressure on politicians, stakeholders, and senior management officials to demonstrate that resources invested in policy initiatives have been well spent. They need to justify the overall cost of the policy and the allocation of policy resources to international, regional, and local programs and projects. Increasingly, they are enquired about the value obtained for the money, relative to alternative investment channels. Evaluators are in turn asked to address these questions. (p. 2)

The edited book takes up the challenge of evaluating such overarching policy initiatives (which, I reiterate, epitomize the complicated not the complex) in two ways: (1) Discussing important conceptual distinctions to help make sense of what can be done evaluatively in the face of such multifaceted interventions and (2) examples of actual evaluations of large-scale policy interventions.

Insights About Evaluating Large-Scale Policy Interventions

Putting aside my criticism that the book fails to adequately distinguish the complex from the complicated (or, for that matter, the simple), the editors and authors have thought deeply about the nature of large-scale policy evaluations and, in some cases, about evaluating complexity. I have selected one insight from each chapter to share in this review. These are 12 reflective gems polished through real-world practice.

1. *Introduction* by Kim Forss and Robert Schwartz, commenting on *innovation*:

To succeed, managers as well as evaluators should recognize that successful organizations are driven from where they are to a destination that they create and discover, and this process of creation and discovery is a paradoxical one. It is, first, to promote and protect order and stability in the

day-to-day conduct of the existing business and in the existing strategic direction. But it is also to create an atmosphere of questioning and contention, disorder and chaos, which threatens the bureaucracy, and then manage the boundaries around the instability that has been generated. The challenge for evaluators is thus more to feed into such processes than to provide fixed answers to specific questions. This . . . opens up the possibility of innovation. (p. 26)

2. Patricia Rogers (from Chapter 2 on distinguishing simple, complicated, and complex interventions and the implications for evaluation), *on matching the design to the situation*:

The nature of the intervention is an important consideration when planning, implementing, and reporting evaluation. This does not mean that evaluations should always be complicated or complex. If the intervention can usefully be thought of as simple, it should be. An overly complicated evaluation design and findings that accurately capture unimportant nuances can be less useful than clear, simple messages that are approximately accurate. However, where there are important aspects that are complicated or complex, evaluating it as if it were simple risks producing findings that do not accurately represent what is happening, or that cannot be readily used more widely. (pp. 49–50)

3. John Mayne (Chapter 3) *on contribution analysis*:

Contribution analysis aims at confirming or revising a theory of change, not primarily to discover a theory of change. It is not a panacea for addressing attribution, but can provide useful information on the contribution an intervention is making. Neither is it a cop-out, allowing a limited statement of a theory of change to cover the attribution issue. It requires careful thought, in-depth analysis, evidence gathering, testing, and re-testing. (p. 89)

4. Mita Marra (from Chapter 4 on micro, meso, and macro dimensions of change), *on the need for multiple options and interpretations*:

[O]ur analyses are inevitably based on observations and theories that refer only to a small part of reality. The underlying causes and explanations may involve other factors, and the ‘policy effects’ may be delayed, ripples or possible tidal waves. This consideration suggests that systems may evolve not always toward innovation and change, but also toward decline and impoverishment. Some models of particular situations can help us understand what it is we believe is going on and therefore how we might behave in ways that are most advantageous for economic growth and HD [human development]. Yet, since this knowledge is extremely dubious, it is always better to have multiple options and interpretations available to deal with what we cannot understand and could not anticipate. (p. 113)

5. Jacques Toulemonde, Douglas Carpenter, and Laurent Raffier examine the poverty impact of European support at the country level using Tanzania as a case study (Chapter 5) and discuss *the evaluability barrier*:

The term ‘evaluability barrier’ has been defined as the impossibility to evaluate the achievement of the overarching goal of a complex and complicated intervention, such as a country assistance strategy . . . [T]he evaluability barrier could not be broken in the case of the evaluation of the EC support to Tanzania, in spite of an exceptionally enabling environment. It is strongly suggested that the barrier will never be broken in any other similar instance. Moreover, the new challenges of joint assistance and untargeted budget support programs are rising [sic] rather than lowering the bar. (pp. 142–143)

6. Peter Wilkins reports on evaluation of a multiagency response to homelessness in Australia (Chapter 6) and comments on *transparency*:

The end result of the monitoring and evaluation processes was increased public transparency, which in itself makes them worthwhile. (p. 142)

7. Markus Spinatsch reports on evaluating a smoking prevention policy in Switzerland (Chapter 7) and concludes about the *necessary conditions for effectiveness*:

If you cannot treat with attribution or contribution, information on the implementation of the program, specially [sic] about good governance and management, become essential to judge the program, based on the argument that good governance and professional management may not be sufficient, but at least be necessary for a program to be effective. (p. 182)

8. Robert Schwartz and John Garcia examine Ontario's *Smoke-Free* campaign (Chapter 8) and reflect on the emergence of *comprehensive strategies*:

There is widespread agreement in the public health policy and research communities that comprehensive strategies involving concurrent implementation of a range of interventions provide the greatest opportunity for changing behaviors associated with preventable disease. (pp. 190–191)

9. Kim Forss presents an analysis of the global emergency response to Human immunodeficiency virus infection/acquired immune deficiency syndrome (Chapter 9) and reflects on *managing complexity*:

The evaluation teams did not have any strategy to handle complexity per se. At the time, the teams were well aware that we had a difficult and challenging task, but we are in business as evaluation consultants and did not perceive these assignments as anything else than interesting and challenging tasks to be completed according to the terms of reference and within a budget. We had ourselves to blame for ending up with complex evaluation assignments Thousands of evaluations are managed—one way or the other. The evaluators do the job and handle issues—complex or not, irrespective of whether they as evaluators understand complexity in formal and theoretical terms or whether they just see it as yet another difficult task. (pp. 219–220)

10. Burt Perrin and Peter Wichmand report on the challenges of evaluating strategic development interventions aimed at child labor (Chapter 10) and advocate a standard of *reasonable attribution*:

The theory of change . . . provides a basis for plausible, or 'reasonable' attribution. Given the nature of child labor interventions . . . , absolute proof of cause-and-effect causality may be difficult to provide. Indeed, a simple, unidimensional model of cause and effect makes little conceptual sense for child labor interventions because . . . impact will come about through the combination of actions or a series of actions taking place in a given situation. Change will always be a result of the interaction between the activities of a given project or program with contextual variables that are beyond its direct control.

So, how should one deal with attribution in such a situation? . . . [O]ne should aim to demonstrate a reasonable attribution or credible association between child labor interventions and the impact that has occurred One should seek to make a convincing case that the intervention in question has contributed to a reduction in child labor, rather than to pretend to be able to provide absolute proof of this. (pp. 262–263)

11. *Jos Vaessen* on the challenges of impact evaluation using REs (Chapter 11):

“Randomized experiments are potentially strong on internal validity, yet miss out on providing (strong) evidence on other aspects such as, for example, unanticipated and long-term effects, and the external and construct validity of findings. This brings us to the important realization that impact evaluations by default cannot and should not exclusively rely on *one methodology* alone.” (p. 300)

12. *Mita Marra* in the concluding chapter on insights from complexity science (Chapter 12) reflects on *evaluators as part of the action*:

Evaluation of complex policies does not merely result in a set of findings through which policy-makers seek to understand and take advantage of environmental changes. The evaluation of complex policies results in a process by which different players in the environment actively seek to change the ‘rules of the game’ that shape the environment. In this view, the environment is not the antecedent or mediator of evaluation processes . . . , but rather the ‘ambience’ that spawns a given evaluative response. Evaluators become part of such environment that forges evaluation as well. (p. 327)

Speaking Truth to Power

In opening this review, I suggested a truth-in-advertising book title change from *Evaluating the Complex* to *Evaluating the Complicated*. But that change only goes part way. Let’s go all the way and do a *full monty* on the book’s title:

Pretending to Evaluate Comprehensive Policy Interventions Without Telling Policy Makers That What They’re Asking for Can’t Be Done, But We’ll Take a Shot At It Anyway, and Blame the Inadequate Results on Complexity

These evaluators describe managing and negotiating tensions between the ambiguities and the uncertainties of real-world complexity against the demands of policy makers for bottom line accountability, summative judgments of whether an intervention worked, conclusive impact evaluation, value-for-money analysis, and simple attribution—all to be delivered with limited resources on short timelines. There are subtle shifts in emphasis that evaluators make, trusting that most policy makers won’t really notice the switch from assessing outcomes to documenting implementation, or from providing answers to asking clarifying questions for future reflection. All in a good day’s work.

There is some speaking truth to power, but most of it comes in the foreword by Elliot Stern who edits the journal *Evaluation: The International Journal of Theory, Research and Practice*, and has served as president of both the United Kingdom and European Evaluation Societies. Ironically, his bio is not listed among the contributors to the book, though his contribution is the book’s most astute and forthright when he opines:

The recognition of complexity, indeterminacy, and the systemic qualities of much of what is evaluated, confronts policy makers with difficult questions about the promises they habitually make to “deliver”; “drive up quality”; “invest for impact”; and “manage for results.” Implicit in these slogans are assumptions about how much control policy makers actually exercise, what are the reasonable time-scales for managing change, the (non) mechanistic nature of policy interventions, and the extent to which and in what circumstances outcomes can be taken-for-granted or anticipated

[T]he most optimistic outcome of this debate would be to open up space in the way policy making is structured and represented. Not everything can be attributed, anticipated, and measured. That also suggests the need to change the collusive relationship in democracies between citizens and their

“leaders”—in which we will only elect them if they promise us the impossible and they seek to “prove” that the impossible has indeed happened! (pp. ix–x)

What Complex Interventions and Strategies Really Look Like

What, we are left to wonder, would it look like for political leaders to actually engage in a strategy based on complexity theory? It turns out that we have a recent exemplar.

Throughout 2014 as the forces of the Islamic State (also known as the Islamic State in Iraq and the Levant [ISIL]) grew in strength and territory, the phenomena of nonlinear emergence and dynamic adaptation were on full display. Political leaders were called on to delineate their counterstrategy. The sudden rise of ISIL was unexpected and unpredicted, and the outcomes of its emergence were, as of this writing, unknown. Strategies for defeating ISIL were interrelated with the political, humanitarian, and military situation in Iraq, Turkey, and Syria, as well as other Middle Eastern countries. On August 28, 2014, in a White House press conference, President Obama replied to journalists’ questions about the U.S. approach as follows: “We don’t have a strategy yet.” He talked about ongoing intelligence gathering, consulting with allies, consulting with Congress, deliberations with his advisors, and discussion of options with the military leadership. The strategy was still emergent and uncertain.

Obama was skewered for being weak and indecisive (e.g., Bruni, 2014). Presidents and other policy makers, at all times, in all places, and on all issues are supposed to have clear, specific, and measurable outcomes to be achieved by a clear, specific, and explicit strategy. The strategy can be simple or complicated but it cannot be complex—uncertain and still emergent—at least not honestly so.

But Obama had already articulated a complex strategic approach to foreign policy. He had articulated his strategy with the White House press corps traveling with him on a trip to a global leadership summit meeting in Asia in June of 2014. Major news outlets reported his foreign policy strategy as follows:

According to multiple reliable sources, he was defensive and, by one account, “fuming.” He felt that the criticism of his approach was unfair. He had clear ideas about how to manage America’s global interests. In his own words, they centered on a single concept: “Don’t do stupid shit.”

In fact, after making this point, he reportedly stood up, headed forward toward his own cabin on the plane, and then stopped. He turned back to the gathered reporters, and, much like an elementary school teacher hammering rote learning into students, he said, “So what is my foreign policy?” The reporters, in unison, then said, “Don’t do stupid shit.” (Rothkopf, 2014, p. 1)

That is articulation of a complex strategy. Evaluate that complexity, if you will. For my part, I’d love a shot at it.

Conclusion

The international orientation of the editors is a strength. Editor Kim Forss has a doctorate from the Stockholm School of Economics and conducts evaluations through the private company he founded. Editor Mita Marra is a public policy professor at the University of Salerno and senior researcher at the Italian National Research Council in Naples. Editor Robert Schwartz is professor in the Dalla Lana School of Public Health at the University of Toronto and editor-in-chief of the *Canadian Journal of Program Evaluation*. The book is part of an important series on *Comparative Policy Evaluation* edited by Ray Rist. This is Volume 18.

I offer this review of *Evaluating the Complex* in the spirit expressed by the editors who stated their desire to foment debate: “an important purpose of ours is to encourage a debate on complexity and to point to new questions for the future as well as emerging methodologies and new thinking around evaluation” (p. 30). As this review demonstrates, in my judgment, the book fulfills that

worthy purpose. While in much of what I've presented here and elsewhere (Patton, 2011, 2015), I have emphasized a different perspective about evaluating complexity, I want to close in partial agreement with the editors' overall conclusion:

... [W]e would like to emphasize that those who accept to undertake a complex evaluation are on their own. There are no handbooks or manuals and that is not what we are offering either. We have come far beyond the use of checklists There are some fairly general lessons to learn and these lessons rather emphasize the loneliness of the complex policy evaluator. The main lessons are negative: "do not use the most common approach"; "do not apply standardized tools and methods"; etc. The evaluator may find some inspiration in the past, but by and large each complex evaluation assignment is unique and needs to be treated as such. That is an insight in itself and an important message coming out of this book. (p. 14)

An evaluation team might seek some inspiration from how complex evaluation tasks have been solved in the past, but they must develop their own responses. Nobody ever solved a problem like theirs in the past and nobody ever will after them either. (p. 30)

I agree that each complexity-focused evaluation is unique. However, I take issue with the image of "the loneliness of the complex policy evaluator." There has emerged a self-organized community of practice among evaluators interested in systems thinking and complexity theory. This book is further evidence of such a community. Facing uncertainty together is one of the bonding elements of this growing community. So let me end where I began, with the book's wise foreword from Elliot Stern:

The debate this volume embodies is important, but the outcomes of the debate remain uncertain It is reasonable that the future of evaluation practice remains uncertain. In a world characterized by interdependence, emergent properties, unpredictable change, and indeterminate outcomes, how could evaluation be immune? (p. xi)

The uncertainty of our individual and collective futures is all the more reason to join together in embracing and celebrating complexity (rather than fearing it), and learn together about *Evaluating the Complex*.

Reviewer's Postscript

As this review was being finalized (December 2014), Brenda Zimmerman was killed in a tragic automobile accident in Toronto. She pioneered the conceptualization and application of the simple-complicated-complex distinctions featured in this book's second chapter by Patricia Rogers entitled "Implications of Complicated and Complex for Key Tasks in Evaluation." Those distinctions are also the foundation for this review. Distinguishing the complicated from the complex was central to *Getting to Maybe: How the World Is Changed* that Brenda Zimmerman coauthored (Westley, Zimmerman, & Patton, 2006). Brenda's work on complexity and her contributions to *Developmental Evaluation* are featured in Chapter 4 of the book by that name (Patton, 2011). Without Brenda Zimmerman there would have been no *Getting to Maybe* and *Developmental Evaluation* (Patton, 2011), had it been written at all, would have been a very different book. I dedicate this review to her contributions and upholding the distinctions she first articulated in her book 1998 book *Edgeware: Insights from complexity ideas for health care leaders* (Zimmerman, Lindberg, & Plsek, 1998).

References

- Bruni, F. (2004, September 2). Obama's messy words. *The New York Times*, p. A23. Retrieved from www.nytimes.com/2014/09/02/opinion/frank-bruni-obamas-messy-words.html
- Funnell, S. C., & Rogers, P. J. (2011). *Purposeful program theory: Effective use of logic models and theories of change*. San Francisco, CA: Jossey-Bass.
- Kuhn, T. S. (1970). *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press.

- Patton, M. Q. (2011). *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. New York, NY: Guilford.
- Patton, M. Q. (2012). *Essentials of utilization-focused evaluation*. Thousand Oaks, CA: Sage.
- Patton, M. Q. (2015). *Qualitative research & evaluation methods* (4th ed.). Thousand Oaks, CA: Sage.
- Rothkopf, D. (2014, June 6). Obama's 'Don't do stupid shit' foreign policy. *Foreign Policy*. Retrieved from http://www.foreignpolicy.com/articles/2014/06/04/obama_dont_do_stupid_shit_foreign_policy_bowe_bergdahl
- Snowden, D. J., & Boone, M. E. (2007). A leader's framework for decision making. *Harvard Business Review*, 85, 68–77.
- Westley, F., Zimmerman, B., & Patton, M. Q. (2006). *Getting to maybe: How the world is changed*. Toronto, Canada: Random House.
- Zimmerman, B., Lindberg, C., & Plsek, P. (1998). *Edgework: Insights from complexity ideas for health care leaders*. Irving, TX: VHA.