

Book Reviews

Revisiting Realistic Evaluation

Chris L. S. Coryn

Albeit some might argue that this review is a little late in coming, it is worth revisiting Pawson and Tilley's 1997 book, *Realistic Evaluation* (reprinted in 1998, 2000, 2001, and 2002) as the debate about causation and evidence-based research and evaluation continues to be a topic of debate and concern in the evaluation and research communities (see A Call to Action: The First International Congress of Qualitative Inquiry and The Claremont Debate, in this issue of *JMDE*). *Realistic Evaluation* is rooted in the tradition of scientific realism, which is said to be one of the "dominant axes in modern European thinking" (p. 55). In the most general of terms scientific realism concerns "the nature and operation of causal forces" (p. 55). The essential ingredients for assessing these causal forces are C-M-O configurations—where C represents context, M represents mechanisms, and O represents outcomes. Context refers to the "spatial and institutional locations of social situations, together, crucially, with the norms, values, and interrelationships found in them" (p. 216). Mechanisms are the "choices and capacities which lead to regular patterns of social behavior" and the causal mechanisms which generate these patterns of behavior are "deemed 'social problems' and which are the rationale for a program" (p. 216). Outcomes "provide the key evidence for the

realist evaluator in any recommendation to mount, monitor, modify, or mothball a program” (p. 217). From the C-M-O configuration, the authors argue that the way in which causation in the “social world should be constructed” and that the “basic realist formula” is “mechanism + context = outcome” (p. xv).

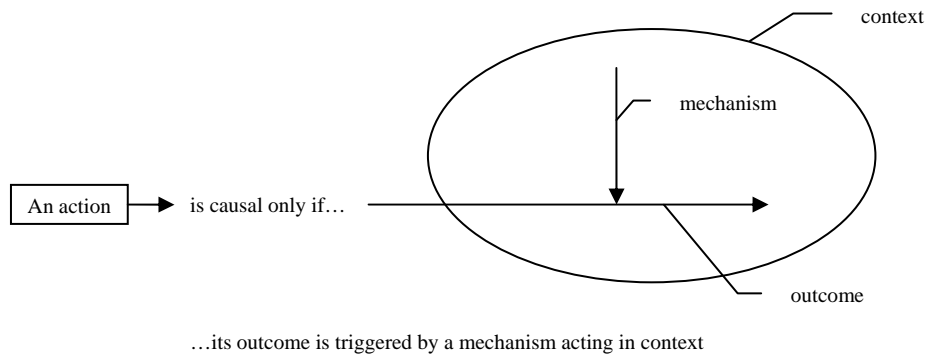
Chapter 1, A History of Evaluation in 28 ½ Pages, presents Pawson and Tilley’s version of the history of evaluation. The authors begin this history with the experimental evaluations of the 1960s of the “great social programs of the ‘great society’ [the U.S.]” (p. 2) brought about by the work of Stanley and Campbell, among others. In short, the experimental paradigm is described as a failure for a variety of reasons, including the lack of external validity (i.e., generalizability) brought about by experiments’ inability to reproduce results in the ‘real world.’ Somewhat out of place, but next in the short history of evaluation are the utilization-focused approaches. These approaches are criticized on the grounds that “he who pays the researcher calls the methodological tune” (p. 14). Finally, the emergence of constructivism in the 1970s is reviewed and also described as a disappointment because of the “inability to grasp those structural and institutional features of society which are in some respects independent of individuals’ reasoning and desires” (p. 23). All in all, the authors paint a bleak picture of evaluation’s past and contend that if the future is to be brighter then evaluators had better take theory seriously, although the authors also find serious flaws in the various theory-driven approaches of Chen, Weiss, and others. These faults are described as the lack of attention given to context and the emphasis on experimental methods, for example. This 28 ½ page history of evaluation is intended to set the stage and substantiate the authors approach to and purposes for evaluation: determining not only ‘if’ a program works, but also ‘how’ and for ‘whom.’

Chapter 2, Weaknesses in Experimental Evaluation, presents Pawson and Tilley's expose on the inherent problems with the experimental tradition; namely, the experimentalists' "epistemological assumptions about causation and their lack of fit with the nature of social programs" (p. 30). Essentially, the authors argue that more often than not that change cannot be captured in *OXO* terminology. All in all, it is asserted that "by its very logic, experimental evaluation either ignores these underlying process [causal mechanisms], or treats them incorrectly as inputs, outputs or confounding variables, or deals with them in a *post hoc* and thus arbitrary fashion" (p. 54).

In Chapter 3, In With the New: Scientific Realism, the authors present the principles and practice of scientific realism. As previously mentioned, the realist view (generative) of causation can be described thusly (as illustrated by the explosion of gunpowder):

Our basic concern is still, of course, the *outcome* (the spark causing the explosion). But what does the explanatory work is first of all the *mechanism* (the chemical composition of the substance which allows the reaction), and secondly the *context* (the physical conditions which allow the mechanism to come into operation). This proposition—causal outcomes follow mechanisms acting in contexts—is the axiomatic base upon which all realist explanations build.

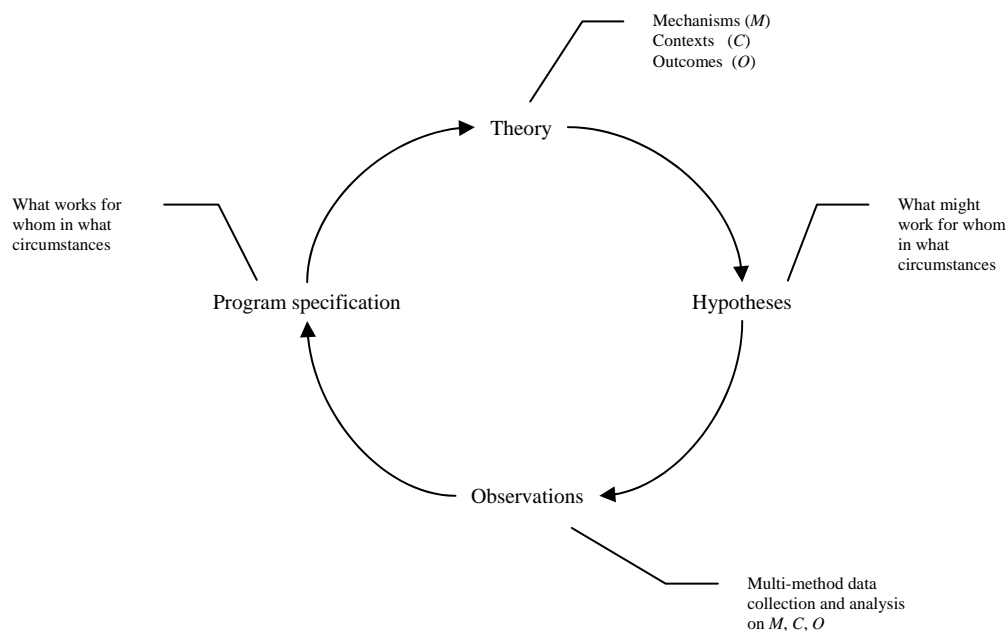
(Pawson & Tilley, 1997, p. 58)



Source: Pawson, R. and Tilley, N. (1997). *Realistic evaluation*. Thousand Oaks, CA: Sage.

Figure 1. Generative Causation

Chapter 4, How to Design a Realistic Evaluation, presents the realist evaluation cycle (see Figure 2) and three case studies which apply realist evaluation principles to varying degrees.



Source: Pawson, R. and Tilley, N. (1997). *Realistic evaluation*. Thousand Oaks, CA: Sage.

Figure 2. The Realist Evaluation Cycle

The first case study presented is an evaluation of property marking and described by the authors as *testing theory*, the second is an evaluation of a housing project and described as *theory formation and development*, and the third is an evaluation of a prison-delivered higher education program also described as *theory formation and development*. These case studies are described in some detail and are intended to represent exemplars of realistic evaluation.

I have opted to exclude a review of the remaining chapters (5-9) as these merely focus on collecting realist data and the methodological procedures involved in conducting evaluation as prescribed by Pawson and Tilley.

Despite the book's title, the true underlying premise of Pawson and Tilley's *Realistic Evaluation* is not merely a proposition of how to conduct evaluation, but rather a treatise on the nature of causation and science. While the author's notion of causation (scientific realism) is compelling, I am not entirely convinced that it is the "final solution" to the causation debate. Neither is it a dramatic improvement over either successionist or other traditions. In their haste to prescribe generative explanations they fail to recognize or acknowledge that numerous experimentalists (and non-experimentalists) give considerable attention to context in their accounts of causation (e.g., moderators, mediators, interaction effects), often to a greater degree than the examples provided throughout the book suggest. Moreover, these causal accounts (i.e., realist accounts) seem little more than explanations of program effectiveness for different groups or consumers, which can be accomplished without the use of realist principles.

Prior reviews (Patton, 1999; Rogers, 1999) of *Realistic Evaluation* have been mixed. For example, Rogers (1999) stated that "this is one of those rare books that has the potential to permanently change one's perspective on program evaluation"

(p. 381). Patton (1999), on the other hand, was not entirely convinced of the credibility of Pawson and Tilley's contribution and responded to their criticisms of utilization-focused evaluation (p. 14) thusly:

I rarely respond to attacks on or distortions of my views, especially when they're based on the twenty-year-old first edition of the book (Patton, 1978) and don't take into account subsequent revisions and elaborations (Patton, 1986, 1997) that I hope have corrected at least some earlier weaknesses, and have benefited from well-deserved and well-meaning critiques. I have learned that responding to a distortion risks reinforcing the very thing I want to correct by calling attention to it. However, the distortions in the opening chapter of Pawson and Tilley, in which they sarcastically and disparagingly review (and bemoan) the history of evaluation, are anything but innocent or trivial. The irony is that, in the introduction, the authors claim the mantle of "detachment," "objectivity," and "scientific evaluation" (p. xiii). Their mocking review of evaluation's history has one primary purpose: positioning themselves as saviors of the profession by redirecting us to be scientists first and foremost.

(Patton, 1999, p. 387)

While *Realistic Evaluation* has spurred serious interest and debate, and even spawned an issue of *New Directions for Evaluation* (Henry, Julnes, & Mark, 1998), the approach has not quite received the attention in North America that it has in the United Kingdom and Europe. A search of the American and Canadian evaluation journals did not turn-up any publications related to the approach (with the exception of Patton and Roger's reviews of the book). While a search of the major European evaluation journal (*Evaluation: The International Journal of Theory, Research, and Practice*) returned 56 articles which focused on, or emphasized, the realistic evaluation approach.

References

- Henry, G. T., Julnes, G. & Mark, M. M. (Eds.) (1998). Realist evaluation: An emerging theory in support of practice. *New Directions for Evaluation*, 78.
- Patton, M. Q. (1999). Realistic evaluation [review of the book *Realistic evaluation*]. *American Journal of Evaluation*, 20(2), 385-388.
- Pawson, R. and Tilley, N. (1997). *Realistic evaluation*. Thousand Oaks, CA: Sage.
- Rogers, P. J. (1999). Realistic evaluation [review of the book *Realistic evaluation*]. *American Journal of Evaluation*, 20(2), 381-383.