R. Pawson. (2006). *Evidence-based Policy: A Realist Perspective*. Thousand Oaks, CA: Sage. 209 pages.

Reviewed by Greg Mason

Evidence-based policy, evidence-based health care, evidence-based evaluation—the list goes on. Tacking the modifier "evidence-based" in front of any decision-making process conveys the impression of rigour and a break with (bad) past practice. When I first opened Pawson's book, I braced myself against a series of pious homilies on how to assess policies. I was pleasantly surprised. This book belongs on the shelf of all evaluators. Let me explain why.

Chapter 1 starts with a concise and incisive deconstruction of what normally passes for policy evaluation. It is a road map to what ails evaluation in Canada. He identifies *positioning* as the process of ensuring that policy evaluation remains anchored, which is often accomplished by engaging retired senior civil servants to conduct evaluations. This is certain not to question core assumptions and actually increases the likelihood of policy replication. *Portals* refers to the process of using knowledge management to amass data and create databases of outcomes to measure performance. Pawson disavows the process of creating "the latest platform (that) encompasses a set of basic knowledge management functionalities, along with the ability to rapidly develop a collaborative knowledge portal and e-learning solutions within a pull-communications culture." The current flirtation with dashboards and data management systems to track performance is not evidence-based policy in Pawson's terminology. He sees it as an utter waster of time. *Polling* is frequently used to justify programs, and Pawson notes that this is simply the application of politics to evaluation. The fact that 75% of Canadians believe that the federal government should assist evaluation consulting firms to grow is not evidence of value. *Partnership* refers to participatory research. Pawson accepts this as valid at the local level, but argues that this is so situational that it has no generalizability to a wider domain—that is, it generates no evidence across the domains where the policy may be applied. *Partisanship* is the scaled-up version of participatory research. Using the example of disabilities programming, the author notes that this usually results in irresolvable struggles among stakeholders. Finally, *punditry* is the use of experts and wise men (peer

Corresponding author: Greg Mason, Managing Partner, PRA (Inc.), 500-363 Broadway, Winnipeg, MB, Canada R3C 3N9; <gmason@pra.ca>

reviewers?) to bring "independent" advice to an evaluation. Pawson dismisses this, arguing that experts invariably have biases and many either do not disclose them or, worse, are unaware they have them. Experts, especially academics, tend not to recuse themselves when they do not have the requisite knowledge.

The core of his argument starts in Chapter 2 with a discussion of the term *systematic review,* which lies at the heart of evidence-based policy. Pawson defines evidence-based policy as a response to a simple question: "What works?" All the theoretical development, research designs, qualitative assessment, and so on. focuses on that one single question. This chapter does a neat job of presenting the complex causality underlying even a simple program. The diagram that concludes this chapter completely repudiates the simplistic logic models that lie at the centre of all federal evaluations.

Chapter 3 is a gem. Pawson eviscerates meta-analysis, concluding that the meta-analytic process of "standardization, summarization, and aggregation does not produce any enduring truths." In particular, meta-analysis omits program theory and context, and systematically excludes evidence that can answer the question of what works.

Chapter 4 is the heart of the book, where Pawson discusses a "realist synthesis" anchored by the concept of a systematic review. This chapter reveals what it means to combine multiple lines of evidence, something that is stressed but rarely done in Canadian evaluations. A systematic review has six steps:

1. *Identifying the question* is much more than just clarifying terms. It includes theory development and ranking the questions. In most evaluation matrices (rationale/relevance, design/delivery/success/impact, and cost-effectiveness/alternatives), the 20 or 30 questions can be boiled down to two or three that really matter.
2. *Searching for primary studies* involves developing a set of hypotheses to guide the search for material (and these are not key terms to stuff into Google Scholar). These are complex, time-consuming literature reviews, not the superficial scans that now typify most evaluations. Human Resources Development Canada commissioned a Lessons Learned series in the late 1990s that carried budgets of in excess of $120,000 (in 2008 dollars) and required a year to complete. It is a shame that series has not continued.

3. *Quality appraisal* involves the time-consuming assessment of each primary study. This is the tipping point for this approach, as Pawson writes: "If evidence is to have its say it should be based on primary studies carried out with the highest methodological standards" (p. 87).
4. *Extracting the data* is the labourious task of pulling meaning from the primary studies. This section requires close reading, for Pawson reveals what it actually means to review multiple studies to develop insight.
5. *Synthesizing the data* is creating a central message about what all the research actually means. Key issues that emerge from the synthesis are the identification of where the program theory fails, the elimination of weaker theories, and the identification of what may be expected as the program is applied locally.
6. *Dissemination* is a final step in which, Pawson argues, the information must be released widely, not concealed.

Chapters 5, 6, and 7 apply the process of systematic review to three interventions—Megan's Law (the regulation that the relocation of a dangerous sex offender must be revealed to the community), youth mentoring, and naming and shaming (the identification of bad behaviour such as videotaping customers of prostitutes). Each of these chapters explores the application of realist synthesis to reveal the complexity of each program.

The book is not without flaws. A minor issue is Pawson's affection for alliteration and "cute" turns of phrase. These tend to occur mostly in the first part of the book and can be a bit offputting. What is more problematic is that the book does a very good job of undermining evaluation as practiced, especially at the federal level. Pawson enumerates many foibles of current practice, but offers little concrete direction to evaluators operating in the current accountability environment. It would be a great book for a Ph.D candidate about to write a policy thesis, as it energizes its reader with the determination to do better, but then one quickly falls to earth with the realization that no evaluation client could ever allow the time to complete such a realist evaluation of policy.

However, all said, this is a very useful antidote to the formulaic evaluation texts that dominate the field.