

THE NEW RULES OF REALISTIC EVALUATION

This chapter is a brief recapitulation of the argument of the book.

Learning a paradigm, as Kuhn once said, is like learning a language. This summary takes the form of a final bombardment of realist research terminology, and so is presented as a glossary of the 'key terms'. We have faced a temptation, in writing this book, to introduce a whole raft of realist neologisms and to wreak havoc on some of the established evaluation usages. By the end, though, we find that we have been more influenced by our self-imposed injunction to be 'realistic'. Evaluation is, after all, *applied* research. And although we have called throughout for injections of 'theory' into every aspect of evaluation design and analysis, the goal has never been to construct theory *per se*; rather it has been to develop the theories of practitioners, participants and policy makers. The proof of this realist pudding thus lies, not so much with the question of what realism can do for evaluation, but with the issue of what realist evaluation can do for programming and policy making. The upshot is that we have avoided being terminological terrorists and have tried instead to be conceptual captivators in smuggling the realist paradigm ever so gently into the existing language of evaluation. We remain true to this spirit in this final statement of the essence of our ideas. Accordingly, this *realistic evaluation thesaurus* is the ultimate pocket reference, and what we feign as a 'bombardment' is in fact a little coruscation of just eight terms.

Each concept is introduced with an associated *rule* for the conduct of evaluation research. With backgrounds in sociology we are naturally super-sensitive to the expression 'rules of inquiry'. The very idea of laying down the law on social science method has remained controversial ever since the famous attempt by Durkheim (1962). Gallons of methodological blood have been spilled subsequently with regard to the extent that one can pre-specify, in abstract and general terms, pathways of investigation which are, by definition, novel, time tied, and location specific. Our usage of 'rules' is not meant to imply we are involved in the production of some axiomatic truths as set down by the realist thought-police, nor is meant ironically. Methodological rules actually develop in the way we have generated them in this book – by going between principle and practice (Diesing, 1972), by traversing logic-in-use and reconstructed logic (Kaplan,

1964). Methodological rules are not written in stone but are the medium and outcome of research practice (Pawson, 1989, Chapter 1).

Methodological progress can actually be charted in much the same way as we trace the success of a program. The same realist explanatory principles apply. Our baseline argument throughout has been not that programs are 'things' that may (or may not) 'work'; rather they contain certain ideas which work for certain subjects in certain situations. We hardly have to change terminology at all to come to the appreciation that methodological rules contain ideas which work for certain investigations of certain processes. These ideas on the scope and salience of methodological rules do not simply descend from metaphysical heaven but are established by being tried and tested in research practice. In the beginning, the great epistemological and ontological principles act as high totems inspiring the research practitioner to certain broad and prized objectives. Success in a specific investigation may be claimed as an embodiment of these principles – but that particular inquiry will necessarily have a range of distinctive features which allowed it to meet the goals. It is the task of the methodologist to elucidate these features, so that the research community can copy the practice. And that practice, having been imitated across a range of inquiries with mixed success, then enables the research community to reflect back on the principles (and allows the methodologist a little carve of the totem). Over time there develops a picture of the research *contexts* in which the particular investigative *mechanisms* work best to generate the desired methodological *outcomes*. Methodological rules, like program theories, are thus under a process of continual refinement. This book has attempted to distil the process by acting as go-between, connecting up the ideas of a half-dozen philosophers and a handful of evaluators. The following statements will thus only gain currency if many more evaluators adopt the rules by adapting them.

Rule 1: generative causation

Evaluators need to attend to how and why social programs have the potential to cause change. Causation is not to be understood 'externally' and so the basic evaluation task is *not* to hypothesize or demonstrate the constant conjunction whereby programme X produces outcome Y. The change generated by social interventions should be viewed 'internally' and takes the form of the release of underlying causal powers of individuals and communities. Realists do not conceive that programs 'work', rather it is the action of stakeholders that makes them work, and the causal potential of an initiative takes the form of providing reasons and resources to enable program participants to change. The capacity for change of natural and social phenomena is only triggered in conducive circumstances. The evaluator needs to understand the conditions required for the programs' causal potential to be released and whether this has been released in practice.

Rule 2: ontological depth

Evaluators need to penetrate beneath the surface of observable inputs and outputs of a program. Social (and physical) reality is stratified. That which we can observe, including the most manifest and routine regularities, is produced by the operation of underlying generative forces which may not be immediately observable. Interventions are always embedded in a range of attitudinal, individual, institutional, and societal processes, and thus program outcomes are generated by a range of macro and micro social forces. In social life, the choice making behaviour of individuals in their different situations is fundamental to understanding their manifest patterns of behaviour. In social interventions the stakeholders' capacity for choice making is, of course, subject to social constraint and is always limited by the power and resources of their 'stakeholding'. Program evaluations need to grasp how the changes introduced inform and alter the balance of the constrained choices of participants.

Rule 3: Mechanisms

Evaluators need to focus on how the causal mechanisms which generate social and behavioural problems are removed or countered through the alternative causal mechanisms introduced in a social program. Realist evaluators seek to understand 'why' a program works through an understanding of the action of mechanisms. Mechanisms refer to the choices and capacities which lead to regular patterns of social behaviour. Causal mechanisms are at work in generating those patterns of behaviour which are deemed 'social problems' and which are the rationale for a program. Programs are often prolonged social encounters and even the simplest initiative will offer subjects considerable compass for decision making. A key aspect of evaluation research design is thus to anticipate the diversity of potential program mechanisms involved and a key analytic task is to discover whether they have disabled or circumvented the mechanisms responsible for the original problem.

Rule 4: Contexts

Evaluators need to understand the contexts within which problem mechanisms are activated and in which program mechanisms can be successfully fired. Realist evaluators seek to understand 'for whom and in what circumstances' a program works through the study of contextual conditioning. The operation of mechanisms is always contingent on context; subjects will only act upon the resources and choices offered by a program if they are in conducive settings. Context refers to the spatial and institutional locations of social situations together, crucially, with the norms, values, and interrelationships found in them. Just as programs involve multiple mechanisms, they will, characteristically, also include multiple contexts. Another key act

of design and analysis is thus to try to identify the people and situations for whom the initiative will be beneficial by drawing on success and failure rates of different subgroups of subjects within and between interventions.

Rule 5: Outcomes

Evaluators need to understand what are the outcomes of an initiative and how they are produced. Outcomes provide the key evidence for the realist evaluator in any recommendation to mount, monitor, modify or mothball a program. Programs cannot be understood as undifferentiated wholes, as 'things' with some simple brute facticity. They fire multiple mechanisms having different effects on different subjects in different situations, and so produce multiple outcomes. Realist evaluators thus examine outcome patterns in a theory testing role. Outcomes are not inspected simply in order to see if programs work, but are analysed to discover if the conjectured mechanism/context theories are confirmed.

Rule 6: CMO configurations

In order to develop transferable and cumulative lessons from research, evaluators need to orient their thinking to context-mechanism-outcome pattern configurations (CMO configurations). A CMO configuration is a proposition stating what it is about a program which works for whom in what circumstances. The conjectured CMO configuration is the starting point for an evaluation, and the refined CMO configuration is the finding of an evaluation. Whilst realists know that the same program will often work in different ways in different circumstances, they appreciate that sequences of evaluations oriented to one another can improve the understanding of the configurations. Rather than replicate interventions in anticipation of the same results, the realist evaluator sees subsequent trials as an opportunity for CMO configuration focusing, a process in which the relatively well-known action of a program mechanism is fine-tuned to adapt it to local circumstances. Rather than anticipating the cumulation of program wisdom in the form of discovering representative programs which work universally, the realist evaluator seeks to generalize about programs through a process of CMO configuration abstraction, the creation of middle range theories which provide analytic frameworks to interpret similarities and differences between families of programs.

Rule 7: Teacher-learner processes

In order to construct and test context-mechanism-outcome pattern explanations, evaluators need to engage in a teacher-learner relationship with program policy makers, practitioners and participants. These stakeholders clearly have an insider understanding of the programs in which they are implicated and so constitute key informants in the research process. Programs, however, are

embedded in a diversity of individual and institutional forces. Accordingly, there will be limitations to the understanding of any particular group of stakeholders, and the evaluator needs to be attentive to the unintended consequences and unacknowledged conditions of their decisions. Realist evaluators neither assume that stakeholders should act as 'respondents' providing answers to the predetermined questions of the researcher, nor assume that their task is the faithful 'reproduction' of the privileged views of stakeholders. This division of expertise requires a teacher-learner relationship to be developed between researcher and informant in which the medium of exchange is the CMO theory and the function of that relationship is to refine CMO theories. The research act thus involves 'learning' the stakeholder's theories, formalizing them, 'teaching' them back to the informant, who is then in a position to comment upon, clarify and further refine the key ideas. Such a process, repeated over many evaluations, feeds into the wider cycle of 'enlightenment' between the research and policy fields.

Rule 8: Open systems

*Evaluators need to acknowledge that programs are implemented in a changing and permeable social world, and that program effectiveness may thus be subverted or enhanced through the unanticipated intrusion of new contexts and new causal powers. Unlike the physics laboratory – where empirical systems can approximate theoretical models fairly well, and specified mechanisms can be triggered in well-controlled conditions to produce specifiable regularities – in the changing environment of social programs, empirical closure is chronically compromised. Stakeholders always learn from their experience of interventions and emergent causal forces threaten even well-established CMO configurations. The sudden failure of a hitherto successful program is, of course, subject to realist explanation by 'reconstructing' the action of the hitherto unconsidered mechanisms and contexts. But this open system character of social life means that all CMO configurations are *ceteris paribus* where *ceteris* are, and can never be, *paribus*.*

Evaluation is a craft and in this work we have sought to apply a touch of *intellectual craft* to the endeavour. This phrase, of course, belongs to C. Wright Mills and, as our finale, we can apply one of his favourite tricks to our own work. Mills worried over the tendency for social scientists to tread a well-worn pathway to eminence which went from bright ideas to rigorous scholarship to massive erudition to over-blown pomposity. His antidote was to show how the 300,000 word monographs of some key thinkers could be reduced to the odd pithy paragraph (Mills, 1959). We close with the ritual humiliation of reducing our own ideas to a single sentence. In Chapter 1, we used the same diagrammatic matrix in an attempt to summarize the guiding themes of experimental, pragmatic, and

naturalistic evaluation (Figures 1.2, 1.3, 1.4). We rely on the same grid in order to make our final summary statement of the rules of realistic evaluation as in Figure 9.1.

This diagram does, of course, have a more prosaic purpose as well. Comparing it with the other evaluation paradigms, one can see at a glance that the directional flow of our ideas is foundational in that the fountain-head is with methodological fundamentals. In this respect, we have sought to recapture some of the spirit of the evaluation pioneers, whilst swapping the idealism of their motto of the 'experimental method for the experimenting society' for the harsher realism of a 'realist-method-for-a-realistic-world'. On the other hand, as compared with the constructivists and pragmatists, we remain wantonly idealist in refusing to accept that the consequence of programs being peopled and political is that researchers become mere palliators and pundits.

We believe, in short, that *the strength of evaluation research depends on the perspicacity of its view of explanation*. This is perhaps the one methodological rule from which we will not be budged.