Rigour (-mortis) in Evaluation[1]

Martin Reynolds

Evaluation-in-practice can be regarded as a confluence of interactions between three broad idealised sets of stakeholders – the evaluand, evaluators, and commissioners of evaluations. Elsewhere I have suggested two contrasting manifestations in which these interactions might be expressed; one as an 'evaluation-industrial complex' (similar in form to the 'military-industrial complex' originally used by Dwight Eisenhower in 1961), and another as a more benign 'evaluation-adaptive complex' (Reynolds, 2015).

Building on the idea of an iron triangle that empowers the military-industrial complex, I represented the relationships of evaluation-in-practice as a triadic interplay involving six activities that influence the evaluation process (Fig.1). Here I focus on only one of the six activities – commissioning – and I summarise what it might look like for an evaluation-adaptive complex.
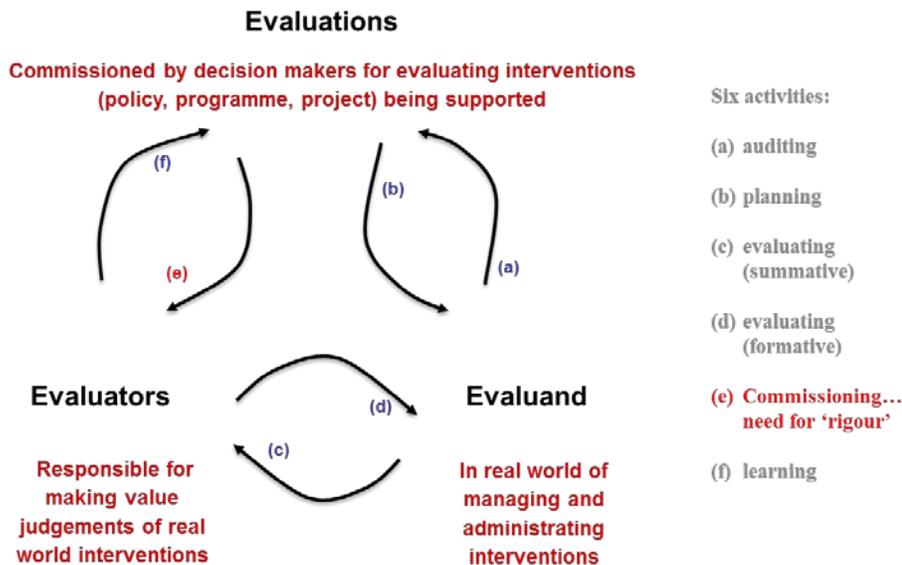

Fig. 1 Evaluation-in-practice involving six key activities
(adapted from Reynolds, 2015 p.75)

One of the key influences illustrated in Figure 1 is the relationship between 'commissioners' and 'evaluators' and the associated requirements for assurance of trustworthiness. Collectively such assurances aim to guarantee rigour. In an evaluation-industrial complex scenario, assurances of rigour are frequently experienced as stifling and rigid leading to an unhelpful and malign confluence that I call *rigour-mortis* resulting in an inability to support radical and

[1] Adapted from 2014 EES conference paper entitled "From rigour (-mortis) to systemic triangulation: ethics and competences in evaluating complexity" - 1st October 2014 Strand 2: Complexity and Evaluation.

transformative interventions. Conversely I contrast this option with more benign forms of rigour, i.e. a set of co-guarantor attributes - or CoGs. In so doing I draw on traditions of American pragmatism (cf. William James and John Dewey), critical social theory (particularly Jürgen Habermas), and critical systems thinking (particularly C. West Churchman and Werner Ulrich).

Each of these traditions is based on Immanuel Kant's fundamental idea that the existence of an absolute guarantee of certainty is a fundamentally flawed notion. Rather rigour lies in the assurance of expert support given to decision-makers (e.g. commissioners/ funders) in evaluating interventions. Specifically evaluators act as guarantors of successful implementation of plans/ interventions (projects, programmes, policy). However this can turn into a source of deception where the guarantee is worthless or false given the inherent limits associated with evaluation inquiry. . In any intervention there is always a built-in risk about the value of the evaluative guarantee.

The search for a more robust albeit provisional set of co-guarantors springs from Habermas' distinction between three knowledge constitutive interests (Habermas, 1972; 2014): (i) *technical* interest in prediction and control of natural and social affairs; (ii) *practical* interest in fostering mutual understanding; and (iii) an *emancipatory* interest in being free from coercion. Figure 2 reconfigures these interests in terms of three sets of CoGs for rigour in evaluation – objectivity in making appropriate representation of the evaluand, complementarity in resonating with alternative representations amongst evaluators, and responsibility in making transparent whose purpose is being served in the evaluation.

Co-Guarantor attributes (CoGs) of Rigour	Features of each set of CoGs
CoG1 Objectivity *(technical interest)*	1. intra/multidisciplinary 2. based on criteria of reliability 3. inviting disciplinary responsibility in *representing* the 'real world'
CoG2 Complementarity *(practical interest)*	1. interdisciplinary 2. based on criteria of resonance 3. inviting general academic critique in valuing different representations
CoG3 (social and ecological) Responsibility *(emancipatory interest)*	1. transdisciplinary 2. based on criteria of relevance 3. inviting social and ecological critique in making transparent the wider purpose of support

Fig. 2 Co-guarantor attributes of rigour
(adapted from Reynolds, 2001 and 2003)

False guarantors exist where either the criteria of rigour are not appropriately fulfilled or where one set of CoGs is privileged over the other two. So for example, a false guarantor of objectivity may be manifest when, say, a randomised control trial (RCT) might be inappropriately used according to the *scientific* disciplinary guidelines of use, such as in circumstances where control experiments are not feasible and/or ethical. A false guarantor might also be apparent when an RCT is used as sole guarantor of an evaluation with associated claims of abiding by 'best practice', through being 'scientifically' objective and neutral. Such arguments and critiques against the dominant use of evidence-based evaluations such as RCTs for evaluating social interventions have an increasingly impressive and effective tradition in the field of evaluation (cf. Patton, 2010; Pawson et al., 2011; Rogers, 2008).

Arising from this critique of *best practice* emerges an alternative notion of *best fit* based on the contingency approach to evaluation. An expression of contingency is the demarcation between simple, complicated, and complex interventions (Glouberman and Zimmerman, 2002); an approach recently critiqued by Mowles (2014) and Reynolds (2015). The simple-complicated-complex idea is a very helpful heuristic for understanding systemic failure of interventions where inevitable complex situations of an evaluand are misconceived as either complicated or simple. The idea is less useful as a heuristic for rigour in prescribing evaluation 'tools' for predefined situations.

Table 1 sketches a few expressions of CoGs in relation to the two archetypal forms of evaluation-in-practice. Some features of the contingency approach are used here to illustrate some features of rigour in the evaluation-industrial complex. They are intended to invite conversation on other features of rigour that may also inhibit a shift to an evaluation-adaptive complex.

Table 1 Some co-guarantor attributes (CoGs) of rigour compared

CoGs	Evaluation-industrial complex (for example, false guarantors associated with contingency approach)	Evaluation-adaptive complex (towards critical systems thinking approach towards developing CoGs)
1. Objectivity (and reliability)	Reality can be regarded objectively as either simple, complicated, or complex.	Reality regarded as 'unknowable'; an integral mixture of complicated, complex, and conflictual… need for systemic inquiry
2. Complementarity (and resonance)	Pursuit of an ever-increasing 'toolbox' driven by multiple-method ethos, where discrete tools are deemed 'fit' for discrete situations.	Cultivation and adaptation of existing methods and approaches through social learning using others' expertise and experiences… need for pragmatism
3. Responsibility (and relevance)	Legitimacy given by terms of reference for evaluation and perceived fixed bias of different tools regarded as being relevant for different purposes.	Legitimacy given by evaluators' 'political' role in sustaining or challenging purposes of evaluation…need for radical constructivism

The shift towards a more radical type of rigour involves (i) humility about the built-in fallibility of any evaluation tools, (ii) empathy and openness with alternative forms of evaluation, and (iii) innovative political practice in regarding evaluation as purposeful systems design.

References

Glouberman, S. and B. Zimmerman (2002). *Complicated and complex systems: what would successful reform of Medicare look like?,* Commission on the Future of Health Care in Canada Toronto.

Habermas, J. (1972). *Knowledge and Human Interests.* London, Heinemann.

Habermas, J. (2014). *Truth and justification.* London. John Wiley & Sons.

Mowles, C. (2014). "Complex, but not quite complex enough: The turn to the complexity sciences in evaluation scholarship." *Evaluation* **20**(2): 160-175.

Patton, M. Q. (2010). *Developmental evaluation: Applying complexity concepts to enhance innovation and use.* New York, Guilford Press.

Pawson, R., et al. (2011). "Known knowns, known unknowns, unknown unknowns: the predicament of evidence-based policy." *American Journal of Evaluation* 32(4): 518-546

Reynolds, M. (2001). Co-Guarantor Attributes: A Systemic Approach to Evaluating Expert Support. *Eighth European Conference on Information Technology Evaluation: Conference Proceedings*, Oriel College, Oxford.

Reynolds, M. (2003). "Towards Systemic Evaluation: A Framework of Co-guarantor Attributes". Evaluating Regional Sustainable Development submission to Workshop of the EU Thematic Network project REGIONET *Evaluation methods and tools for regional sustainable development*, University of Manchester (UK). 11-13 June 2003.

Reynolds, M. (2015). (Breaking) The Iron Triangle of Evaluation. *IDS Bulletin*,*46*(1), 71-86. http://oro.open.ac.uk/41879/

Rogers, P. J. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, *14*(1), 29-48.