

 Open access • Journal Article • DOI:10.1111/1759-5436.12117

## Prosaic or Profound? The Adoption of Systems Ideas by Impact Evaluation

— [Source link](#) 

Bob Williams

**Published on:** 01 Jan 2015 - IDS Bulletin (2015 Institute of Development Studies)

**Topics:** Impact evaluation

Related papers:

- [Evaluating complex social interventions in a complex world.](#)
- [Complex, but not quite complex enough : The turn to the complexity sciences in evaluation scholarship](#)
- [Evaluation in Complex Policy Systems](#)
- [Breaking\) The Iron Triangle of Evaluation](#)
- [Towards Systemic Evaluation](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/prosaic-or-profound-the-adoption-of-systems-ideas-by-impact-vexsrgyg4e>

# Prosaic or Profound? The Adoption of Systems Ideas by Impact Evaluation

Bob Williams

---

**Abstract** All evaluation approaches have to address questions about their legitimacy, validity, relevance and usefulness. As the complexity of interventions is more widely acknowledged, impact evaluation appears to be especially vulnerable to these challenges. This article explores the potential of the systems field to address these vulnerabilities. The systems field is conceptualised as understanding interrelationships, engaging with multiple perspectives and reflecting on where boundaries are drawn in terms of those interrelationships and perspectives. This article argues that achieving a balance between these three elements is critical. An emphasis on interrelationships is likely to bring only limited (prosaic) benefits to impact evaluation as a whole. On the other hand, a strong emphasis on perspectives and boundaries could result in profound changes to the way in which impact evaluation is conceived and delivered. In particular, it could change the nature of the relationship between the evaluator and key stakeholders, especially funders and managers of interventions.

---

## 1 Introduction

What is the potential contribution of systems and complexity concepts to impact evaluation? What is the potential for it to be game changing and profound and what is the potential for it to be more prosaic and incremental? For *individual* impact evaluations it is difficult to generalise about the use and value of particular systems approaches, since the contribution will relate to the particular circumstances under which the evaluation is taking place. However, for impact evaluation as *an entire endeavour* I will argue that it depends on which aspects of the systems field evaluators collectively adopt and adapt.

The systems field is often understood as characterised entirely by a greater understanding of interrelationships. That is only part of the story and if that characterisation dominates I argue that the contribution of systems ideas in impact evaluation as a whole is somewhat prosaic.

But the systems field has more to offer than just a heightened understanding of interrelationships. It includes an acknowledgement that the engagement in multiple perspectives deepens understanding of interrelationships. It acknowledges that the boundary choices, including value judgements based on particular perspectives

reflected by interventions and their evaluations require careful deliberation for ethical, political and practical reasons.

Engaging in the implications of multiple perspectives and reflecting on boundary choices tends to challenge dominant impact evaluation practices, in particular the relationship between the evaluator, the intervention and evaluation commissioners (see also Reynolds in this *IDS Bulletin*). On the other hand, these aspects of systems thinking could assist impact evaluation and address some of its current challenges. For those reasons adopting these aspects of systems thinking tends to have a more profound consequence for impact evaluation as a whole, potentially making it more relevant to the circumstances of international development in the twenty-first century.

## 2 Impact evaluation today and tomorrow

As an observer of the impact evaluation scene over the past decade or so, three things strike me as significant.

Firstly, it is characterised by a strong descriptive and causal focus that at times has overpowered the evaluative element: 'Did it – the intervention – achieve a previously specified result and how

did it achieve that result?’ rather than ‘Did it – the intervention – have merit, value or worth?’<sup>1</sup> In other words, the normative (value) and instrumental (worth) aspects of the judgement process have often been defined in ways that imply or are close to the intrinsically good (e.g. merit) aspect of the judgement process.

Secondly, evaluators are ‘prone to frame their evaluations to meet programme managers’ needs and concerns rather than those of citizens’ (as Robert Picciotto writes in *IDS Bulletin* 45.6, Picciotto 2014). Consequently the language of impact evaluation tends to emphasise the primacy of the donor (who provides the money) or intervention manager (who makes available certain resources and expertise). In terms of the previous paragraph, whatever the programme manager (or donor) thinks is a worthwhile impact *is* unquestionably so. Indeed the concepts of ‘intended’ and ‘unintended’ impacts raise the often unaddressed question of intended by whom? Who’s to say the impacts were ‘unintended’? Interventions do not operate in the realm of magic, things happen for a reason. As I discuss later, it is entirely plausible that in many instances someone somewhere grabbed hold of the resources provided by the intervention mechanisms and took them in a different but entirely intended direction.

Finally, the *evaluative* focus has tended to be on process. Did the collective activities and resources contribute to the impact in a valuable or worthwhile way? Was the programme theory and its assumptions about causality appropriate? (see Hummelbrunner in this *IDS Bulletin*). Thus, debates within impact evaluation have tended to focus on process issues such as attribution, causality, theories of change, models and other mechanisms seeking to link input with result. *The appropriateness of achieving that result is often not part of the evaluation brief.* A key point is that, between evaluating whether an intervention was the right thing to do, or evaluating if it was done in the right way, impact evaluation has evolved into a practice that often leans towards the latter rather than the former.<sup>2</sup> For example, the recent UK DFID Working Paper 38, *Broadening the Range of Designs and Methods for Impact Evaluations*, comprehensively reviewed the impact evaluation scene and drew ten conclusions. Five of these conclusions were specifically focused on causality. Indeed one of those conclusions states that impact evaluation ‘can *best be* understood as a

form of causal analysis that links an intervention with effects for beneficiaries’ (my emphasis) (Stern *et al.* 2012).

These three tendencies didn’t happen accidentally, they are a product of evaluation history and evaluation economics. Evaluation as we know it developed primarily as a branch of applied social science, rather than say planning, policy development, operational research, strategy development, economics or political science. Thus, there tends to be an emphasis on the analytical ‘how’ rather than the more ideological or normative ‘ought’. Additionally, a casual glance of any evaluation discussion group or conference agenda will highlight how keenly evaluators want their work to be used by those who commission their work. While that’s a praiseworthy attitude, those funding or managing interventions understandably often want to know if their resources and efforts are having the intended impact rather than whether their intended result was a good one in the first place. I’ve had programme managers ask me to remove from my report ‘good’ additional impacts of interventions on the basis of those impacts being irrelevant to their own interests. So ‘usefulness’ tends to be oriented towards the use of evaluation by funders and programme managers.

Finally, despite the rhetoric, evaluations generally and impact evaluations especially, have tended to be commissioned and used for purposes of holding those involved in the intervention to account rather than for purposes of intervention improvement and broader learning. The international ‘aid’ scene was dominated by single clearly identified interventions with relatively observable impacts, often provided by single agencies and funded by single donors. Notions of accountability could be based on relatively simple understandings of cause and effect and relatively simple framings of interventions, where the intended impacts were largely uncontested (at least at the time). The institutional form of impact evaluation; its resources, skills, structures, expectations, methods, purpose and means of legitimacy, grew into that space.

The question that this and its companion volume, *IDS Bulletin* 45.6, poses is whether this state of affairs, historically determined and economically supported, is relevant to the kinds

of demands placed on impact evaluation today. And if it isn't then what needs to change?

As described in *IDS Bulletin* 45.6, the world of international development has shifted radically in the past few years (Picciotto 2014; Winckler Andersen 2014).

In current jargon, international development is today seeking to address and resolve so-called 'wicked' problems (see Williams and Van 't Hof 2014; Reynolds, this *IDS Bulletin*). The focus is on 'development' rather than 'aid'. The enterprise employs multiple agents (foundations, governments, commercial businesses, migrants) that form webs of partnerships, with access to multiple resource streams (donations, royalties, migrant remittances, bi/multi-lateral trade arrangements), undertaking diverse interventions aimed at somewhat nebulous impacts such as 'capacity' (Morgan 2013). The mechanisms which help generate impacts to occur are variable and unpredictable to such an extent that the notions of 'accountability' become tenuous since accountability necessarily requires identifiable actions carried out by identifiable bodies to be directly linked to identifiable results (Gregory 1995). Furthermore, impacts and the efforts to achieve these results are often highly and, in recent years, violently contested – from dam building to polio reduction. And, of course, impacts can be produced by several causal drivers of which the intervention, in the best of cases, is only one. Inevitably these matters raise questions as to whether the existing practice of impact evaluation continues to be fit for purpose.

It is within this emerging context that the adoption and adaptation of ideas drawn from across the systems field is identified as a means to address the challenges impact evaluation faces. How hopeful can we be that the systems field will successfully address these challenges?

The answer, as is often the case in evaluation matters, is yes, no and maybe. In this article I paint two possible scenarios. If impact evaluations continue to comply with client accountability imperatives, and maintain their focus on causal relationships, then the systems field can certainly provide some powerful tools for making greater sense of interrelationships in complicated and complex situations. However, overall, I suspect that the potential contribution

of those aspects of the systems field is likely to be prosaic... valuable but not profound. On the other hand the contribution could be profound if impact evaluation uses systems approaches that promote the multi-perspectival, ethical, democratic and normatively driven stance that already underpins the domains of democratic evaluation, empowerment evaluation, feminist evaluation and indigenous evaluation.

A short history of the systems field will help further explore these two scenarios.

### **3 A short history of the systems field**

Any history describes disputed territory but systems ideas can be traced back at least a couple of thousand years. According to Gerald Midgley,<sup>3</sup> the modern systems movement can be seen to trace its development through three main phases or 'waves', starting in the 1930s, accelerating during the Second World War, cresting in the 1950s and 1960s, followed by more waves and turbulence during the 1970s and 1980s before lapping periodically against the shores in the years since then.

From the early days until the late 1960s, the focus of the systems discourse was very much on *interrelationships*. This 'first wave' discourse is still influential today, generally reflected by labels such as 'hard systems', systems engineering, first-order cybernetics, system dynamics and complex adaptive systems. Some of the familiar mapping and 'wiring diagram' approaches (e.g. network maps, concept maps, causal loop diagrams) originated during this first phase.

By the early 1970s, some in the systems field (notably Peter Checkland and Sir Geoffrey Vickers) realised that the relative importance that people ascribed to particular interrelationships within a situation often depended on the different perspectives through which they observed that situation (Ramage and Shipp 2009). Thus systemic thinking began to include the implications of applying different *perspectives*, worldviews or framings to the same situation. Broadly speaking, approaches that emphasise this aspect of the systems cannon are known as 'soft systems'.

However, by the mid-1980s, systems thinkers such as C. West Churchman, Werner Ulrich, Michael Jackson and Gerald Midgley had concluded that focusing primarily on

interrelationships and perspectives had its own problems (*ibid.*). Perspectives influence what we consider relevant or irrelevant; they determine what is ‘in’ our framing – the way we understand a situation – and consequently determine what lies ‘outside’ that framing. Thus, whoever defines the dominant perspective controls the boundary of a systemic inquiry or intervention. Thus, the importance of identifying *boundaries* and critiquing boundary decisions (including those who made them) is the third core concept underpinning a systems approach. Methods and methodologies that emerged from this wave are collectively known as critical systems.

Although I’ve described these waves and the methodologies as distinct phenomena, it is important to understand that methods and methodologies contained within the systems field address all of these three orientations to a greater or lesser extent:

- Understanding interrelationships;
- Engaging with multiple perspectives;
- Reflecting on boundaries.

The primary concern raised in this article is when impact evaluation draws more from the interrelationship dimension than from the other two. So let us look now at the implications for impact evaluation of these three dimensions of the systems field.

#### 4 Understanding interrelationships

Only connect...

E.M. Forster, *Howard’s End* (1910)

The concept of interrelationships is what comes to most people’s minds when talking about the adoption of systems ideas. How things are connected and with what consequence, stems from the earliest thinking about systems. It is also the concept most strongly embedded in the popular imagination. When we talk about the education system or the health system, we imagine a set of objects and processes that are interconnected in some way. The popularity of system dynamics and complex adaptive systems in many parts of the world cements the notion that interrelationships are an important – in many cases the only – systems concept.

However, it is a mistake to think that the systems field concerns itself with just any

interrelationships. It focuses on particular aspects of relationships:

- Dynamics: how the interrelationships affect the behaviour patterns within a situation over a period of time;
- Non-linearity: how the size of the output or effect of interrelationships appears unrelated to the size of the input to the interrelationship. This is often but not always caused by feedback. The classic example of non-linear relationships is the exponential growth curve familiar in ecology and (perhaps less frequently) your compound interest-bearing bank account;
- Context sensitivity: malaria control methods that work well in Thailand may not work in the Philippines. The same interrelationships in different contexts can produce different results. Or that the interrelationships themselves may change according to the context.

Five interrelationship questions are relevant to impact evaluation (indeed any evaluation):

- What is the structure of the interrelationships within the situation (i.e. how are the components arranged)?
- What are the processes between components of that structure?
- What is the nature of the interrelationships (e.g. strong, weak, fast, slow, conflicted, collaborative, direct, indirect)?
- What are the patterns that emerge from these interrelationships over time, with what consequences and for whom?
- What are ways in which these complicated and complex dynamics can be identified and managed effectively?

#### 4.1 So what for impact evaluation? Interrelationships

Interrelationships is a familiar evaluation concept, especially in impact evaluation, with its focus on causal attribution. Methods and techniques such as multivariate statistical approaches, theories of change, outcome hierarchies, programme logic, results chains and logframes are extensively used. However, these methods tend to be static: freezing interrelationships in time, and on the whole, *map* rather than *model* interrelationships. The systems field has the potential to introduce more dynamic and dynamical modelling approaches, such as system dynamics (see Derwisch and

Löwe, and Grove in this *IDS Bulletin*), agent-based modelling and viable systems models to name just three. These approaches can help address the five interrelationship questions more deeply and in ways arguably more valid to the messy situations that evolve over time in which impact evaluation occurs. Understanding better the dynamics of interrelationships greatly improves individual impact evaluations, as the articles in this *IDS Bulletin* attest. However, this improved understanding of interrelationships at the individual evaluation level doesn't address, in any fundamental way, the larger challenges facing impact evaluation described earlier.

### 5 Engaging with perspectives

The systems approach begins when first you see the world through the eyes of another.

C. West Churchman, *The Systems Approach* (1968)

As stated in the introduction, just focusing on interrelationships does not make an intervention or its evaluation systemic. People will see and interpret those interrelationships in different ways depending on their perspectives. A local café owner might evaluate an intervention attempting to prevent the spread of listeria bacteria quite differently than someone from the health service, even though both may 'see' the same thing. But there is more to it. What a health inspector *does* when he or she 'sees' a café premises will be different from what the café owner does when he or she 'sees' the same thing. Our perceptions promote behaviours that affect the way a situation unfolds. They *affect* the nature and dynamic of interrelationships. Thus, what we describe as unintended or unexpected patterns within a situation often results from our unwillingness to understand deeply or explore other people's perceptions and subsequent behaviours. For us to fully comprehend the dynamics of an intervention, the way in which interrelationships resources and activities transform into meaningful results, we must also identify and understand the range of relevant perspectives that people bring to that intervention. To do so, it is helpful to distinguish between three aspects of perspective: stakeholders, stakes and framings.

Stakeholders are groups of people that possess a common role in an intervention (e.g. teachers, consumers, writers). In contrast, stakes relate to individual values and motivations... the

stakeholders' skin in the game (e.g. wealth, honour, fairness, past history, purpose, ideas of professionalism, safety, security). Individuals frequently take on more than one stakeholder *role* in a situation (e.g. teacher and parent) and these different stakeholder roles may share the same stakes (e.g. student wellbeing). Conversely, any single stakeholder role will contain within it several different, perhaps conflicting, stakes (e.g. commitment to student wellbeing vs. compliance with curricular demands).

Deliberating on the dynamics of how stakeholders and stakes interrelate allows us to frame issues. Framing is a means to express what an intervention is – or could be – *about*, what purposes it might promote. Framing helps you identify how people understand an intervention and thus how they behave purposefully towards an outcome. Framing is the lens through which you (or others) view an intervention.

Let us assume your task in playing at a rock concert concerns assembling the setlist of songs. What are the ways in which that concert could be framed and the implications for that setlist? Here are some possible framings; ways of understanding what the concert is about:

- A fun night out;
- Income generation;
- Cultural expression;
- Marketing product;
- Nostalgia for lost youth.

Note that these framings are not outcomes. A range of possible (but different) outcomes could fit within each of these framings. Nor are they purposes, although each of the framings suggests different ways of constructing the setlist that imply different *purposes* for the concert. This is a crucial point for approaches to impact evaluation that commonly assume a single purpose for the intervention. If you considered the concert only through a 'fun night out' framing, you might come up with lots of dance-oriented music whereas if you used an 'income generation' framing the setlist might be more oriented towards promoting the latest iTunes release. A nostalgia framing implies a greatest hits setlist. And so on.

Exploring the concert through these different framings allows you to construct a setlist that satisfies most attendees, towards a set of

worthwhile outcomes. The population of ageing 1970s and 1980s rock stars are very skilled at working within multiple framings of their performances, tipping a nod at each of them. Many programme managers are similarly skilled at corralling and accommodating the many different ‘framings’ that participants bring to an intervention, irrespective of the particular ‘framing’ that the funder of that intervention operates within.

Thus, a perspectives orientation raises four important questions for evaluators:

- Who or what are the key stakeholders within the situation?
- What are the key stakes?
- What are the different ways in which the intervention can be understood or framed?
- What plausible intervention *purposes* do these different framings imply?

### 5.1 So what for impact evaluation? Perspectives

Many would claim that evaluation is all about perspectives, since it collects opinions on interventions in order to judge their merit, value or worth. But consider how the notion of perspectives has been used in this article – it is less about opinions *per se* but about the ways in which people *understand* an intervention; the way they frame it; what they believe the intervention is all about. To draw on the rock concert example, you haven’t a clue what the Rolling Stones’ latest album is called, all you wanted to hear as an audience member was ‘Brown Sugar’. And they didn’t play it. So you may think the concert was lousy even though the rest of the crowd went wild. That’s because you understood – and thus evaluated the concert from a nostalgia framing using nostalgia-related criteria. Much of the discourse around impact evaluation is the contribution of an intervention to so-called intended impacts. A perspectives orientation begs the question: ‘Whose intended impacts are being used from what framing?’.

Without explicit discussions about framing, impact evaluation conversations tend to focus on a single ‘theory of change’ (even if it has many branches). Single-perspective approaches don’t acknowledge the possibility that ‘unintended consequences’ may have been intended by someone who perceives the intervention through a different lens; women’s empowerment say,

rather than economic development. And they could have been part of a deliberate strategy to undermine the ‘intended’ impacts’; for example, an HIV/AIDS strategy aiming to improve the financial independence of street workers in an African mining area was largely undermined by failing to take account of the interests of local money lenders.

There are three implications for those conducting impact evaluations who want to engage more deeply with perspectives, and they tend to move towards more profound rather than prosaic impact evaluation practice.

Firstly, there are implications for impact evaluations as tools to help understand the dynamics of an intervention. For instance, a multi-framing approach can help explain much more clearly why a programme succeeded and how to improve it than the single framing based on the ‘intended’ impact. Put simply, if an intervention ‘succeeds’ for many reasons, then unearthing those reasons helps us evaluate that intervention more reliably and accurately. For instance, I once evaluated a pilot programme that was framed by the funders in terms of ‘management skill development’, yet you could not have understood the impact sufficiently to mainstream it successfully without understanding it also from a ‘management trust building’ framing.

Secondly, there are also implications for impact evaluations as tools for learnings that can be applied beyond the specific intervention. Impact evaluation’s focus on single framings can, inadvertently perhaps, promote an ontological orientation, whereby the theory of change is regarded as a simplified version of what actually happens on the ground. Multiple framings don’t deny an ontological reality, but as Reynolds states in this *IDS Bulletin*, ‘Let us be clear and unambiguous, systems thinking is an explicit endeavour of abstracting from reality’. The real power of systems ideas lies in promoting an epistemological learning orientation... ‘What might we learn if we saw this intervention not as an economic development endeavour but as a women’s empowerment endeavour, and how might that learning improve the economic development impact of projects like this one?’.

Few evaluators would disagree with these arguments. Indeed, there are entire branches of

the evaluation tree (e.g. democratic, realist, feminist, empowerment evaluation) that, to abuse this metaphor, are rooted in such arguments. However, faced with the ontological deterministic reductionism of results-based management and the Millennium Development Goals, diverse accountability pressures, a strong evaluation commissioner dependency and a commitment to their evaluations being used, how many evaluators have the confidence and economic independence to challenge the single-perspective orientation of much impact evaluation?

Which raises the third, final and most profound implication. What does it say about the current state of evaluation that evaluators do not in general have the confidence or economic independence to challenge the single-perspective orientation of much impact evaluation? What would it mean for evaluation if evaluators *did* have the confidence and economic independence to challenge the single-perspective, funder-chosen, orientation? What would it imply for the social role, responsibilities and indeed professional status of evaluation if donors, providers or evaluation commissioners were not primarily responsible for the perspectives that impact evaluation often privileges? Would a true engagement with multiple perspectives help rewrite the book of evaluation? These questions become even more acute when considering issues of boundary.

## 6 Boundaries

‘Tut, tut, child!’ said the Duchess. ‘Everything’s got a moral, if only you can find it.’

Lewis Carroll, *Alice’s Adventures in Wonderland* (1865)

Notions of ‘holism’ and ‘big picture’ can be misleading. In reality, an evaluation cannot consider everything, it cannot take everything into account. Every endeavour has to make decisions about boundaries, about deciding what is ‘in’ and what is ‘out’.

These are not just decisions about physical size or scale but more epistemological concerns about what is deemed relevant and what is deemed to be irrelevant, what is important and what is unimportant, what is worthwhile and what is not, what suits the one in a position of power and what doesn’t, who benefits and who is disadvantaged by an intervention – and of

critical importance to evaluation and the establishment of evaluation criteria, what is valued and what is not valued. So while it is important to engage and acknowledge multiple framings, ultimately decisions have to be made about which ones are more and less important.

This raises one of the criticisms of multi-perspectival approaches; that they can promote a contingency orientation – the idea that there is no universal ‘right’ way so any way will do depending on the circumstances.<sup>4</sup> But decisions about ‘rightness’ have to be taken; the challenge is the basis upon which this should be done. And that leads us to considerations of boundary choice and above all, values.

Relevant questions for evaluators that flow from boundary choices are:

- What does each framing imply in terms of what is *valued* and *marginalised*?
- What scale decision implies what is *valued* and *marginalised*?
- What measurement decision implies what is *valued* and *marginalised*?
- Whose involvement in the evaluation implies what is *valued* and *marginalised*?
- Whose views about the intervention and the evaluation are *valued* and whose are *marginalised*?
- What does this mean for the *criteria* by which the worth of the intervention is judged according to its plausible purposes?
- What does this mean for the *criteria* by which the worth of the evaluation is judged?

## 7 So what – boundaries?

A range of considerations spring from an exploration and critique of boundary decisions. Some of them could have a profound impact on impact evaluation.

Firstly, from an evaluation design point of view, as stated earlier, setting boundaries is not optional. You cannot do everything, consider everything, see everything, record everything. Treating boundaries systemically means that you set boundaries consciously and consider the implications rather than taking (someone else’s) boundaries for granted. In particular:

- Boundaries delineate between what is judged as having merit or worth, and what is not.



Thus, when evaluators and their sponsors decide the criteria for judging the intrinsic merit, normative value, relative significance or instrumental worth of an intervention they are taking a boundary decision.

- **Scale.** Boundary decisions about the scale of that impact evaluation are extremely important because something that might be seen ‘valuable’ at one scale may not be valuable at another scale (see also Garcia and Zazueta’s article in this *IDS Bulletin*) (e.g. the use of certain pesticides may benefit sweet potato but devastate biological diversity and affect the overall food production economy).
- **Measurement.** Evaluators often seem puzzled that something as mundane and apparently technocratic as decisions of measurement can be boundary issues. However, the fundamental task of evaluators is to identify and record the benefits of an intervention to beneficiaries (and the harm to those not benefiting). Yet how do evaluators know whether or not an intervention has had an impact and on whom? They do so by observing and measuring. And as the enduring debates around measurement practices demonstrate, decisions around measurement are not just technical but reflect what kinds of measurement approaches we regard as valuable and those we do not. We are not inclined to include in our evaluations measurement approaches that we (or our sponsors) do not rate so highly. In doing so we make a measurement boundary choice.
- **Decision-making.** Just who is and is not included in key evaluation decisions are important boundary choices. Given that not everyone can be involved in evaluation decisions, who ought to be included? Perhaps surprisingly, who ought to be excluded from the decision-making needs consideration. For instance, a good argument can be made that some kind of limit to what decision-makers can decide is necessary to make them answerable for their decisions.
- **Expertise.** Which expertise is honoured and which expertise is marginalised is a key evaluation design and analysis. What are the consequences of including and excluding certain expertise and knowledge from the evaluation? How do you manage the risk of negative consequences of including or excluding certain expertise?
- **Legitimacy.** The evaluation purpose, criteria

and questions are expressions of particular worldviews. Those who do not share those worldviews may undermine the legitimacy of the evaluation. Ought those alternative worldviews be accommodated by the evaluation, and if so how?

These issues are at the core of critical systems approaches and pose both ethical and pragmatic questions.

From an *ethical* point of view, you hold certain values and those values reflect your ethical stance on things. If you believe that women have a right to determine what happens to their bodies, then your intervention is likely to ensure that their voices are heard and acknowledged. Fair enough, but you might also be tempted to prevent or marginalise the voice of those who don’t believe so. But what are the ethical implications of excluding those voices? On what ethical basis do you claim to exclude them? If impact evaluations are ever to represent and reflect broader social perspectives rather than specific client perspectives, then finding ethical ways of managing such conflicting value systems will require close inspection of the boundary choices made by those involved in and affected by the intervention and inevitably by its evaluation.

From a *pragmatic* point of view, those who are marginalised (or those who represent marginalised interests) by an intervention or its evaluation are not likely to take things lying down. Some people may not like a strategy to consider the interests of loan sharks in an attempt to address housing foreclosures, but if they are marginalised there is a risk that they will oppose the intervention and hinder its execution. Thus, programme managers and evaluators will need to work out a way of managing that possibility. So there is a pragmatic reason to explore who or what is marginalised and see how those marginalised interests can be accommodated in your intervention.

Inevitably, power issues get wrapped up in boundaries. We hear much about evaluators and evaluations speaking truth to power and the deliberation on boundary choices made during the intervention or the evaluation provides a means of determining what ‘truths’ need to be articulated.

Of course, there are evaluators and evaluation approaches, such as democratic evaluation, empowerment evaluation, feminist evaluation and indigenous evaluation that share this consciously ethical orientation. But overall, the pressures that work against these kinds of approaches are currently stronger than the recognised benefits. The big issue here is who determines what is evaluated and what criteria are used in that evaluation. Evaluation practice over the past 50 years has in many areas developed an expectation in funders and programme managers that evaluators act as craftspeople taking instructions from them, rather than as professionals giving independent advice to them. Furthermore, many claims relating to independence are based on technical and structural rather than ethical grounds. It takes a very strong evaluator to push against these forces when they are dependent on fee income or internal promotion and unprotected by academic tenure.

## 8 Concluding comments

The world that impact evaluation occupies in international development is changing. There is a shift from individual interventions at a project level to a programme level and to a focus on broad concepts such as ‘capacity building’. Increasingly, interventions range across private sector, internal public sector, external donor sector, traditional foundations and quasi-private

sector foundations interests. The focus is big ticket items such as agricultural intensification, food security, multinational infectious disease control and climate change. These are controversial and ‘wicked’ issues with interventions and impacts that are often contested. Judgements of worth of these interventions raise fundamental issues about the way in which impact evaluation *understands* interrelationships, *engages* with multiple perspectives, and *reflects* on boundary judgements (see Reynolds, this *IDS Bulletin*).

I have argued in this article that for the systems approach to have a profound influence on impact evaluation *as a whole*, all three dimensions of the systems approach, interrelationships, perspectives and boundaries, are necessary. An emphasis on interrelationships alone can benefit individual evaluations but is likely to have only a marginal, prosaic impact on the big issues that confront impact evaluation. While the perspective and boundary tasks described in this article may be unusual, indeed challenging, to traditional impact evaluation practice, their impact could be profound.

So to return to the original statements in this article, whether the impact of systems ideas on impact evaluation is profound or merely prosaic depends on which aspects of the systems endeavour are adopted.

## Notes

1 I am grateful to Barbara Befani for pointing out that is not just the ‘who’ thinks something has value, but also the ‘when’ they think so... goals are rarely changed as knowledge is acquired about (an emergent) programme, and learning about the programme does not result in goal change, or adopting more relevant goals.

- 2 For a more detailed discussion of this distinction and its implications see Hummelbrunner in this *IDS Bulletin*.
- 3 This section is abstracted from Midgely (2007).
- 4 See Reynolds in this *IDS Bulletin* for a more detailed critique of contingency approaches.

## References

- Carroll, Lewis (1993) (first published 1865) *Alice’s Adventures in Wonderland*, Ware, Herts: Wordsworth Editions Ltd
- Churchman, C.W. (1968) *The Systems Approach*, New York NY: Delacorte Press
- Forster, E.M. (2000) (first published 1910) *Howards End*, London: Penguin Classics
- Gregory, R. (1995) ‘Accountability, Responsibility and Corruption: Managing the Public Production Process’, in J. Boston, *The State Under Contract*, Wellington, New Zealand: Bridget Williams Books
- Midgely, G. (2007) ‘Systems Thinking for Evaluation’, in B. Williams and I. Imam (eds), *Systems Concepts in Evaluation: An Expert Anthology*, Point Reyes CA: EdgePress/AEA
- Morgan, P. (2013) ‘Evaluating Capacity Development’, in S. Donaldson, T. Azzam and R. Conner (eds), *Emerging Practices in International Development Evaluation*, Charlotte NC: Information Age Publishing

- Picciotto, R. (2014) 'Have Development Evaluators Been Fighting the Last War... And If So, What is to be Done?', *IDS Bulletin* 45.6: 6–16
- Ramage, M. and Shipp, K. (2009) *Systems Thinkers*, London: Springer
- Stern, Elliot; Stame, Nicoletta; Mayne, John; Forss, Kim; Davies, Rick and Befani, Barbara (2012) *Broadening the Range of Designs and Methods for Impact Evaluations*, DFID Working Paper 38, London: Department for International Development
- Williams, B. and Van 't Hof, S. (2014) *Wicked Solutions: A Systems Approach to Complex Problems*, E-book
- Winckler Andersen, O.W. (2014) 'Some Thoughts on Development Evaluation Processes', *IDS Bulletin* 45.6: 77–85