

Assessing contrasting strategies for ensuring ethical practice within evaluation: institutional review boards (IRBs) and professionalization

Peter O'Flynn, Chris Barnett, Laura Camfield

Author for correspondence: Peter O'Flynn

Affiliations: Centre for Development Impact, Institute of Development Studies

Address: Centre for Development Impact, Institute of Development Studies, Brighton, BN1 9RE

Mobile: 00447507630350

Email: p.o'flynn@ids.ac.uk

Other authors:

Chris Barnett, Centre for Development Impact, Institute of Development Studies, Brighton, BN1 9RE

Email: c.barnett@ids.ac.uk

Laura Camfield, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ

Email: l.camfield@uea.ac.uk

Abstract

This paper explores the application of ethics in two contrasting approaches to evaluation: one that views evaluation as essentially a research project, and the other that sees evaluation as an extension of project management. We argue that the growth in so-called rigorous impact evaluation, characterised by practitioners as evaluation using experimental or quasi-experimental methods, has seen evaluation treated increasingly as a sub-set of research. This has entailed greater use of ethical committees, and specifically institutional review boards (IRBs) as many academics promoting the use of experimental methods are based in the USA. Elsewhere, evaluation is treated more as a management activity, with professionalization initiatives such as membership standards and ethical guidance often used in the place of formal review. In this paper we question whether the simultaneous growth in usage of IRBs and professionalization addresses the ethical issues faced by evaluators.

Keywords: IRBs, professionalization, evaluation, ethics, public-sector management, participatory evaluation

1. Introduction

While there is longstanding recognition of the importance of ethical conduct within evaluation (Newman and Brown, 1996), we argue that in the past decade there has been a greater emphasis by evaluation societies and funders on developing guidelines and establishing formal structures for review (Munslow, 2016). So-called rigorous impact evaluation involving experimental and quasi-experimental methods has become increasingly popular (Bedecarrats et al, 2015) and is often viewed as the 'gold standard' of evaluation methodologies (Cupitt, 2015; ILO, 2014). Evaluation of this sort, which often involves largescale household surveys and ethical challenges around the use of control groups (Prowse and Camfield, 2013) is increasingly being treated as a sub-set of research, and subject to ethical review through IRBsⁱ in the same way as clinical trials.

An alternative view of evaluation characterises it as an extension of the project cycle or public sector managementⁱⁱ, which may therefore not need formal ethical reviewⁱⁱⁱ. Indeed, most development programmes themselves rarely go through any ethical review or screening, and whether the evaluation is commissioned by management or its funder, it rarely goes through formal ethical procedures. This situation is described by Rodgers (2014, para. 5):

'Sometimes they [evaluations] go through our university ethics committee in exactly the same way as a research project, and the commissioning government department is keen for the proposal to have this scrutiny... But sometimes, especially where we are supporting a government agency to re-examine existing data, we argue that it is more like program management activity and an ethics application is not needed, although of course ethical practice always is... But the dividing line between the two is not always clear, and university (or institutional) ethics boards/committees are not always set-up in ways that can provide timely oversight and review of adaptive evaluations.'

Instead, evaluation societies and evaluators engage in self-monitoring^{iv} and argue for greater professionalization using mechanisms such as membership standards, training, ethical guidance, codes of conduct, etc.)

In this paper, we^v argue that while growth in the use of IRBs and increasing professionalization have many benefits for the discipline of evaluation, namely the ability to ensure formal protocols are adhered to (IRBs) and guidance of personal conduct (professionalization), both are primarily focused on the planning/design and data collection phases of evaluations. They also have different strengths: IRBs are typically better for dealing with - or more focussed on - the micro issues of an evaluation,^{vi} ensuring ethical practice within the internal processes of an evaluation (Barnett et al, 2014).

Professionalization approaches may be better at dealing with macro issues such as interests and politics (Green et al, 2006), of which micro-issues such as failing to feedback to communities are often symptomatic.

This paper is structured as follows. After a brief discussion regarding the definition of ethics, we look at the rise in the use of IRBs and professionalization. We consider how ethical concerns are being addressed in these approaches and what this means for the future of ethics in evaluation.

2. Defining Ethics

Many authors writing about evaluation observe ‘a lack of shared understanding of what ethics are’ (Groves, 2016, pp.2) and an unwarranted assumption that ethics do not need definition as there is a common understanding (Groves gives the example of the UN Evaluation Group (UNEG) and DFID Principles (2011), which do not provide a definition of ethics). However, for the evaluation field, Munslow and Hale (2015) building on the work of Simons (2006) suggested that there is no single, context-free set of principles that can be applied to guide ethical judgements in evaluation. Context is key to ethics in evaluation as evaluations occur in situations with different power dynamics, be they informational, financial or political. For these reasons, adaptability and reasoning is required by the evaluator to make ethically valid judgements, requiring knowledge of approaches that are likely to be successful within different situations (Munslow et al, 2015). These include making complex judgements in the context of competing interests, providing sufficient information to funders to make informed decisions, and attempting to maintain impartiality and independence (Green et al, 2006). The challenges posed by these issues explain why the process of evaluation is considered to be where the ‘heaviest assaults to ethics threaten’ (Mabry, 1999, pp. 199).

The confusion over the nature of ethics (for example, whether it only applies in extreme situations such as working with survivors of gender-based violence) and whether ethical practice is context-specific or underpinned by common principles, values, or competencies may explain why ethical approval is not obtained for all evaluations (Groves-Williams, this volume).

3. IRBs in Evaluation

As described earlier, the use of IRBs to review evaluations has grown in the past decade; a growth that is in part due to a rapid rise in quasi and experimental evaluation and particularly Randomised Control Trials (RCTs), which have historically been monitored by IRBs. Bedecarrats et al (2015)

demonstrate this growth in experimental approaches in Figure 1 for evaluations conducted for the World Bank.

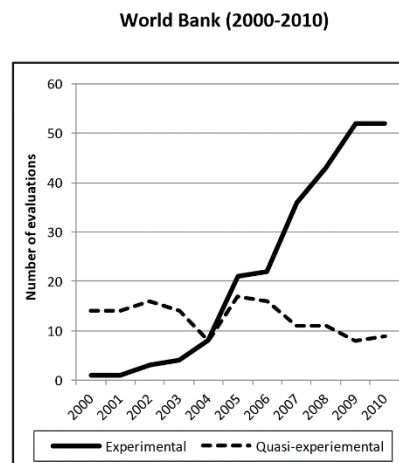


Figure 1: The growth in the use of experimental methods by the World Bank (2000-2010) Source: Bedecarrats et al (2015)

The reason for the increased ethical scrutiny of evaluations using experimental methods is provided by Ravallion (2014, pg. 2) who notes, RCTs “deliberately alter the program’s (known or likely) assignment mechanism”, which has direct effects on the wellbeing of some of the participants. We argue that the rise of IRBs to address these problems has not been universally welcomed. There have been some positive outcomes for beneficiaries, including ethical oversight and the adoption of standards in evaluators’ contracts relating to data protection, independence and conflict of interest. However, possible negative outcomes include a focus on particular kinds of methods to the exclusion of others and a reluctance on the part of IRBs to share data as many evaluation funders now require (see for example, DFID’s Open Data Strategy 2012-14^{vii}).

Ethical review through IRBs tend to cover issues of consent and anonymity (what Camfield and Palmer-Jones (2013) refer to as ‘care of the subject’), but may be less able to see the bigger picture. An example of this includes the power dynamics between evaluator and funder, for instance increasing pressures on evaluators to only report good news, even if this means misrepresenting findings (Morris, 2012). In relation to IRBs, some evaluators lament the fact that they seem to care less about the evaluation *per se*, but are part of the reputation management exercise of the associated institution (Silberman et al, 2012; Lincoln et al, 2004). Nonetheless, we note that IRBs potentially provide the following features that guarantee a minimum standard of ethics:

1. They can act as a screening mechanism, providing the first line of defence against harm from a project being undertaken. This may put the institute's interest ahead of the evaluator's in the sense that an evaluation that appears risky due the use of unfamiliar methods or because it offers greater than usual access to data may not be approved. This caution relates to the IRB's duty to ensure risk mitigation, so that the institution can continue to promote evaluations that do no harm while not suffering negative reputational effects that could affect funding.
2. It can signal to other institutions and the wider public that the evaluator and the institution are working in an upstanding manner, thereby promoting the institution and the methods.
3. The IRB has the ability to suggest amendments to evaluation design or methodology to promote more ethically sound practice.
4. Depending on what the submission process requires, dealing with an IRB can support professionalization, by ensuring that evaluators think about the ethical consequences of their actions. It also produces a standard of conduct that is visible and can be replicated by others.

(Source: Authors' analyses)

However, although IRBs may have these advantages, by the nature of the process of review they are linked to 'institutional contexts of power' (Weiss, 2005). These may favour particular methodological approaches when ideally it should be up to the skill of the evaluator, through a recursive process, to decide which approach is best suited.

4. Can evaluation be seen as just another stage of project management?

While rigour is often seen in terms of the largest possible separation between the project and its evaluators, an alternative view proposes that a truly ethical evaluation is one that is embedded within the project to maximise the opportunities for learning – and ultimately impact. In this view the process of evaluation can be seen as an extension of project cycle management, or public sector management (as seen in UNODC, 2016; European Commission, 2002; Bedi et al, 2006). This can lead to evaluations of programs 'lack[ing] transparency, public input, and deliberation' (Weiss, 2005, pp.

1). Evaluators in this situation may have little support regarding ethical concerns, especially if their own organisation feels threatened by their findings. Of course, this can also happen with an external evaluation and Barnett et al (2014, pp. 10) highlight this when looking at a case-based evaluation of impact in Malawi:

‘...By having no formal requirement for ethical approval (no IRB/ethics committee), although there is guidance under DFID’s standards for research and evaluation. It is a multi-donor programme, and ethical decisions concerning the evaluation are largely at the discretion of the evaluation team members (such as accepted professional behaviour, verbal consent, anonymised reporting of interviews, etc.). A lot therefore rests with the individual/team and their accepted notions of ‘professional standards’. There is no explicit ethical framework that guides the work.’

Addressing macro issues, such as the welfare of people that are not included in the study, but are still affected by the evaluation, may be beyond the scope of the IRB review process. Support for evaluators can come from evaluation societies (e.g. the Australasian Evaluation Society, 2010; American Evaluation Association, 2004), and associated guidance from development partners (Camfield, 2014), but this may be partial and the extent to which it is used is unclear. This guidance is part of the push for self-monitoring and greater professionalization in the industry which supporters claim will ensure ethical quality. As Picciotto (2011) explains ‘Professional autonomy...[is] justified by the need to control the market disorder that inevitably prevails when quacks and amateurs are allowed to enter the fray with inevitable consequences: distrust, confusion and poor service quality’ (pp. 169). The movement towards professionalization also comes from INGOs/NGOs who are required to demonstrate impact and generate funding. Given both the cost and the separation between evaluation and learning entailed in external evaluations, internal evaluation is more important in these settings. However, INGOs/NGOs need to demonstrate the rigour of their internal processes through issuing guidance and sharing examples of good practice.

Professionalization approaches, through network and membership platforms, can potentially address more of the macro-concerns of the evaluator. This is contrasted with the more micro-perspective of the IRBs, which reflect their origin within research rather than evaluation. For example, network and membership platforms allow the opportunity to build on others’ experiences and talk about some of the wider pressures facing evaluators in their projects, such as the dynamics of power and increasing pressures on evaluators to only report good news^{viii}. By focusing on the

generic skill set of the evaluator, professionalization approaches act beyond the scope of one evaluation and/or project, providing recurrent learning.

Through skill based development and the sharing of learning, ethical questions are being addressed, but concerns still arise. Typically, professional associations and membership platforms also focus on the design/planning phase, but is enough attention being placed on later stages, beyond questions of methodology? Secondly, by focussing on self-evaluation through guidance, we have little knowledge of how ethical concerns are actually being addressed.

5. Are IRBs and professionalization approaches reshaping the evaluation landscape?

Both IRBs and professionalization have addressed some of the micro/macro ethical concerns in evaluations, but neither tends to go much beyond the design/planning and fieldwork phases of projects/evaluations. Greater feedback loops around ethical practice would be one mechanism to alleviate this; these are there in theory, but rarely implemented. For instance, most feedback loops are based on a one way system of data extraction, although a two-way system is currently being piloted by DFID (Groves, 2015). There is also a secondary concern. As IRBs face consistent problems understanding and approving approaches to evaluation that are not RCTs or similar (e.g. qualitative, process tracing, democratic approaches etc.) and take a long time to review and decide (Abbott et al, 2011), this can lead to systematic bias in the types of evaluations taking place worldwide. For instance Brown et al (2010) suggests that:

“Review Boards are generally unfamiliar with [participatory methods], reluctant to oversee community partners, and resistant to ongoing researcher-participant interaction. Institutional Review Boards sometimes unintentionally violate the very principles of beneficence and justice which they are supposed to uphold. For example, some Institutional Review Boards refuse to allow report-back of individual data to participants, which contradicts the [participatory method] principles that guide a growing number of projects.”
(para. 3)

We hypothesise that knowing which evaluations are likely get approval by IRBs can change the types of research that evaluators seek approval for. This may make them more risk-adverse and potentially more willing to adopt an experimental approach, rather than the evaluation methodology most

suitable for the institutional context, in order to get ethical approval (a phenomenon that has been observed in relation to qualitative research and clinical ethics committees in the UK).

Similarly, this concern may be seen, albeit from a different perspective, within professionalization, as evaluation is still a 'fledgling profession: its market is being captured by auditors, economists and management consultants' (Picciotto, 2015, slide. 5). As the market further professionalizes and standardises, we note that there could be significant differences in power dynamics, which may lead to more risk-averse behaviours by evaluators who are not willing to undertake new types of evaluations, or use new approaches, be they democratic or others. This has the ability to shift the balance of power between the efficiency vs. empowerment paradigm in evaluation, moving against more participatory approaches (Cleaver, 1999; Wallerstein, 1999).

Another possible reason for this shift is that once credentials and professional approval are sought, it places more value on the viewpoint of the evaluator ('the expert'), and less on the viewpoint of participants/beneficiaries or the public in general, which may lead to even less inclusion of these voices within the evaluation. This can be seen as an example of 'invisible power' (Gaventa, 2009), whereby 'processes of socialisation, culture and ideology perpetuate exclusion and inequality by defining what is normal [and] acceptable' (Ibid, pp. 29). The effects of invisible power are most apparent if the evaluation institutional framework is captured by narrowly focused interest groups (Picciotto, 2011).

Inevitably both these hypotheses, and their effects, have some bearing on the evaluation landscape and may shift it towards narrower and less participatory approaches. We discuss this possibility in the final section.

6. What next for ethics in evaluation?

While IRBs provide a formal space to check protocols, professionalization generally deals with guidance for personal conduct. However neither really address ethical issues faced by evaluators in taking more participatory, adaptive and learning approaches to evaluation such as the inclusion / exclusion of different individuals. In this way blind spots remain, beyond the view of the approaches discussed. Other gaps include lack of attention to the application of ethical guidelines and an unwillingness to report issues that arise in certain institutional contexts so that the rest of the industry learns. Admitting mistakes requires buy-in from both evaluators and commissioners so that the field may learn more and develop. In line with proponents of democratic evaluation such as Picciotto (2011), House (1978), Munslow et al (2015), and Greene et al, (2006), it may also mean that

the actual ethical dilemmas of evaluators in balancing methodology, inclusion / exclusion and the various interests of stakeholders remain hidden – rather than being unearthed, informed by ethical theory and guidance, and sufficiently deliberated in a democratic manner. We propose that in their initial stages evaluations should carefully consider mechanisms that give close attention to the issues raised by stakeholders and value deliberation, independence and objectivity – rather than defer ethical responsibility solely to the formal consideration of an IRB, or the individual professionalism of an evaluator.

Bibliography

Abbott, L. and C. Grady (2011). "A Systematic Review of the Empirical Literature Evaluating IRBs: What We Know and What We Still Need to Learn." Journal of Empirical Research on Human Research Ethics: An International Journal **6**(1): 3-20.

American Evaluation Association. (2004). "American Evaluation Association Guiding Principles For Evaluators." Retrieved 20/09/2016, 2016, from <http://www.eval.org/p/cm/ld/fid=51>.

Australasian Evaluation Society, (2010). Guidelines for the ethical conduct of evaluations. Retrieved 20/09/2016, from <http://www.aes.asn.au/images/stories/files/About/Documents%20-%20ongoing/AES%20Guidlines10.pdf>.

Barnett, C. and T. Munslow (2014). Workshop Report: Framing Ethics in Impact Evaluation: Where Are We? Which Route Should We Take?, IDS.

Bédécarrats, F., et al. (2015). "The gold standard for randomised evaluations: from discussion of method to political economy." Working Paper - WP/2015-01 **5**(45): 45.

Bedi, T., et al. (2006). Impact Evaluation and the Project Cycle. Washington, World Bank. Retrieved 04/08/2016, from http://siteresources.worldbank.org/INTISPMA/Resources/Training-Events-and-Materials/ie_and_projectcycle.pdf

Brown, P., et al. (2010). "Institutional review board challenges related to community-based participatory research on human exposure to environmental toxins: A case study." Environmental Health **9**(1): 1-12. Retrieved 10/08/2016 from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2914003/>.

Camfield, L, Palmer-Jones, R. (2013). Editorial: As Well as the Subject: Additional Dimensions in Development Research Ethics. Progress in Development Studies, vol. **13**(4), 255-65.

Cleaver, F. (1999). "Paradoxes of participation: questioning participatory approaches to development." Journal of international development **11**(4): 597.

Chambers, R. (2012). Participation for development: a good time to be alive. DevPolicyBlog, 2016. Retrieved 11/08/2016 from <http://devpolicy.org/participation-for-development-a-good-time-to-be-alive-20121207-2/>.

Cupitt, S. (2015). Randomised controlled trials – gold standard or fool’s gold? The role of experimental methods in voluntary sector impact assessment. Charities Evaluation Services, NCVO.

Davies, I. and J. Brümmer (2015). Ethics In Evaluation, UNEG Working Group

European Commission. (2014). How to complete your ethics Self-Assessment European Commission - DG Research and Innovation. Version 1.0.

European Commission. (2002). Project Cycle Management Handbook, European Commission – EuropeAid. Retrieved 04/08/2016, from https://www.sle-berlin.de/files/sle/TRAINING/PCM_Train_Handbook_EN-March2002.pdf

ESRC (2015). ESRC Framework for research ethics ESRC, ESRC: 1-51.

Eyben, R., Guijt, I., Roche, C. and Shutt, C. (2015) The Politics of Evidence and Results in International Development: Playing the Game to Change the Rules? Practical Action Publishing, UK.

Gaventa, J. (2006). "Finding the spaces for change: A power analysis." IDS bulletin 37(6): 23-33.

Greene, J.C. (2006) 'Evaluation, Democracy and Social Change', in I.G.R. Shaw, J.C. Greene and M. Mark (eds), The SAGE Handbook of Evaluation, London: Sage

Groves, L. (2016). "Review of Ethics Principles and Guidance in Evaluation and Research", for DFID.

Groves, L. (2015). "Beneficiary Feedback in Evaluation", for DFID. Retrieved 04/08/2016, from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/428382/Beneficiary-Feedback-Feb15a.pdf.

House, E. R. (1978). "Assumptions underlying evaluation models." Educational Researcher 7(3): 4-12.

ILO (2014) Fooled by randomisation: why RCTs might be the real 'gold standard' for private sector development The LAB. Retrieved 04/08/2016, from http://ilo.org/wcmstp5/groups/public/---ed_emp/---emp_ent/---ifp_seed/documents/briefingnote/wcms_335698.pdf

Lincoln, Y. S. and W. G. Tierney (2004). "Qualitative research and institutional review boards." Qualitative inquiry **10**(2): 219-234.

Mabry (1999) in Morris, M. (2011). "The Good, the Bad, and the Evaluator: 25 Years of AJE Ethics." American Journal of Evaluation **32**(1): 134-151.

Morris, M. and B. Clark (2013). "You want me to do WHAT? Evaluators and the pressure to misrepresent findings." American Journal of Evaluation **34**(1): 57-70.

Munslow, T. and K. Hale (2015). Background Report: Right or Wrong? What Values Inform Modern Impact Evaluation?, IDS.

Munslow, T. and C. Barnett (2015). Event Report: Right or Wrong? What Values Inform Modern Impact Evaluation?, IDS.

National Research Council. Committee on Professionalizing the Nation's Cybersecurity Workforce (2013). Professionalizing the nation's cybersecurity workforce. US.

Newman, D. L. and R. D. Brown (1996). Applied ethics for program evaluation, Sage.

Oakes, J. M. (2002). "Risks and Wrongs in Social Science Research An Evaluator's Guide to the IRB." Evaluation Review **26**(5): 443-479.

Picciotto, R. (2011). "The logic of evaluation professionalism." Evaluation **17**(2): 165-180.

Picciotto, R. (2007). "The value of evaluation standards: a comparative assessment." Journal of multidisciplinary evaluation **2**(3): 30-59.

Picciotto, R. (2015). Valuation professionalization and the Voluntary Evaluator Peer Review Initiative (VEPR) [PowerPoint slides]. Retrieved 11/08/2016, from

<http://www.europeanevaluation.org/sites/default/files/NORDIC%20SEMINAR%20PROFESSIONALIZATION.pdf>.

Prowse, M., Camfield, L. (2013). What role for qualitative methods in randomized experiments? *Progress in International Development*, 13:1, 51-61.

The Office of Human Research Protection (2011). Institutional Review Board Guidebook: 1-10. Retrieved 03/08/2016, from <http://www.saylor.org/site/wp-content/uploads/2011/08/PSYCH202A-3.1.4-Institutional-Review-Board.pdf>.

Ravallion, M. (2014). Taking Ethical Validity Seriously, World Bank. 2016. Retrieved 20/09/2016, from <http://blogs.worldbank.org/impactevaluations/taking-ethical-validity-seriously>.

Resnik, D. B. (2015). "What is Ethics in Research & Why is it Important?". Retrieved 04/08/2016, from <http://www.niehs.nih.gov/research/resources/bioethics/whatis/index.cfm?ref=driverlayer.com>.

Rodgers, P in Barnett (2014). "CDI: Placing values back into impact evaluation?". Retrieved 04/08/16, from <http://www.itad.com/cdi-placing-values-back-into-impact-evaluation/>.

Royse, D., et al. (2015). Program evaluation: An introduction to an evidence-based approach, Cengage Learning.

Silberman, G. and K. L. Kahn (2012). "Burdens on Research Imposed by Institutional Review Boards." The Milbank Quarterly 89(4): 599-627.

Simons, H. (2006) 'Ethics in Evaluation', in I.G.R. Shaw, J.C. Greene and M.M. Mark (eds), *The SAGE Handbook of Evaluation*, London: Sage.

UNODC (2016). "Evaluation in the Project/Programme Cycle." Retrieved 07/08/2016, from <http://www.unodc.org/unodc/en/evaluation/evaluation-and-the-project-programme-cycle.html>

Weiss, H. (2005). From the Directors Desk. The Evaluation Exchange, Harvard Graduate School of Education. XI.

Wallerstein, N. (1999). "Power between evaluator and community: research relationships within New Mexico's healthier communities." *Social Science & Medicine* 49(1): 39-53.

Acknowledgements

The authors are indebted to the previous work on ethics conducted by CDI, specifically by Tamlyn Munslow and other contributors, which acted as a base for this article.

ⁱ IRBs are a form of committee, mainly found in the US and/or in relation to clinical trials, which are formally designated for the approval of research and evaluations involving humans and focused on the balance of risk and benefit (Office of Human Research Protection, 2012). We separate these from Research Ethics Committees (REC) and University Research Ethics Committees (URECs), which tend to be more flexible in their approach and encompass a plurality of methods. For further discussion on these institutions please see the editorial of this publication.

ⁱⁱ According to EC (2002, pp.2), project cycle management 'defines different phases in the project life with well-defined management activities and decision making procedures', while the project cycle 'provides a structure to ensure that stakeholders are consulted and relevant information is available'.

ⁱⁱⁱ For example, the ESRC framework for research ethics has historically classed evaluation as audit and therefore they have not considered it something which needs evaluation (ESRC, 2015).

^{iv} For example, the European Commission research and innovation platform produced an ethics self-assessment form for researchers to prepare for funding (European Commission, 2014).

^v For this paper a framing of the authors' positioning is appropriate. One is an experienced evaluator and academic who has conducted evaluations in over 20 countries since 1998. The second author is also an academic, with two decades of international research experience and a focus on research ethics. The final author has experience working in monitoring and evaluation in sub-Saharan Africa, aiding with the stakeholder feedback to multi-stakeholder programmes and working in collaboration with national and international partners.

^{vi} Micro ethical issues are defined as those that focus on the internal processes of an evaluation, such as how relationships are established with respondents, and the underlying values around consent, anonymity, transparency, etc. Macro ethical issues are defined as those that focus on how evaluation relates to society: to which purpose, and whose interests, should the evaluation serve? The definitions are based on the work of Greene (2006).

^{vii} <https://data.gov.uk/sites/default/files/DFID%20Open%20Data%20Strategy.pdf>, downloaded 21/0/16.

^{viii} Kate Hale, pers. comm., based on interviews conducted with evaluators in September 2014.