Workshop Report: Framing Ethics in Impact Evaluation: Where Are We? Which Route Should We Take?

Dr Chris Barnett and Tamlyn Munslow

October 2014

WORKSHOP REPORT: FRAMING ETHICS IN IMPACT EVALUATION: WHERE ARE WE? WHICH ROUTE SHOULD WE TAKE?

Dr Chris Barnett and Tamlyn Munslow

October 2014

Contents

Acronyms and abbreviations

AEA   American Evaluation Association
CDI   Centre for Development Impact
DAC   Development Assistance Committee
DFID   Department for International Development, UK
EES   European Evaluation Society
ESRC   Economic and Social Research Council
FRE   Framework for Research Ethics
GEF   Global Environment Facility
IDS   Institute of Development Studies
IRB   Institutional Review Board
JCSEE   Joint Committee on Standards for Educational Evaluation
MDG   Millennium Development Goal
ODI   Overseas Development Institute
OECD   Organisation for Economic and Co-operation and Development
OPM   Oxford Policy Management
RCT   randomised control trial
UEA   University of East Anglia
UKES   UK Evaluation Society
UNDP   United Nations Development Programme
USAID   United States Agency for International Development

1 Aims of the workshop

This was an interactive workshop – supported by the Institute of Development Studies' (IDS) Department for International Development (DFID) Accountable Grant and the University of East Anglia (UEA) – with a view to opening up the debate on ethics in impact evaluation to a wider audience in 2015. Our working assumption is that ethics, while well established in social science research (through standards, codes and institutions), is generally undervalued in impact evaluation.

The workshop set out to explore the landscape of ethical practice as it currently exists, with a particular focus on the concerns of practitioners presently involved in implementing impact evaluations. The workshop focused especially on exploring the potential for a broader ethics model, where ethical values and principles inform every part of the evaluation process – both *within* the evaluation itself, as well as *in relation to* society. Furthermore, consideration was given to how evaluation fits within the knowledge system that informs international development – and by implication, the extent to which evaluation could/should perform a role in better understanding and critiquing what constitutes 'good' (ethical) development.

Through presentations and interactive discussions with 15 researchers and practitioners (see Annex 2 for the full list of participants), the workshop raised important questions to address in a future event. Being a preparatory first step, the workshop did not intend to answer all of the questions raised, but rather prepare for a subsequent event.

In summary, the aims of the workshop were:

1. To have a **wide-ranging and insightful discussion** on the key ethical challenges facing impact evaluation;
2. To have reached *some* consensus on the **framing of ethical issues** in impact evaluation;
3. To have made a significant **step towards an outline programme** for an autumn event, in terms of topics, speakers, participants, etc.

The agenda for the workshop is presented in Annex 1.

2 Insights from the workshop

2.1 Plenary session 1: Agenda setting, definitions and views from practice

(Dr Chris Barnett and Tamlyn Munslow)

This session reflected on the current state of ethics in impact evaluation: *How do researchers and practitioners understand ethics in impact evaluation? What is specific about ethics in 'impact evaluation' relative to other areas of research and practice? What are the common problems?* The session drew on two main sources. Firstly, a small-scale study that scanned the current landscape, based on a documentary analysis and interviews with researchers/evaluators. And secondly, two examples of ongoing impact evaluations and their contrasting approaches to ethical dilemmas.

2.1.1 Definitions and key terms

For the purposes of the workshop, we have adopted the Centre for Development Impact's (CDI's) working definition of impact evaluation. This postulates that: 'Impact evaluations are evaluations that assess the contribution of an intervention towards some outcome or goal. The contribution may be intended or unintended, positive or negative, long-term or short-term. Impact evaluations attempt to identify a clear link between causes and effects, and explain how the intervention worked, and for whom'.[1] This definition does not limit impact evaluation to only experimental and quasi-experimental methodologies – although these are important methodologies – but still maintains a focus on understanding impact and causality. It also has a number of distinctive features:

- Firstly, impacts can be unintended, positive or negative, long- or short-term. The importance of longer-term, sustainable and transformational change is important when assessing impact, and is often overlooked in assessing 3–5-year interventions.
- Secondly, the definition focuses on exploring the links between cause and effect, although importantly, this is not limited to a counterfactual framework of causal inference. This characteristic, in particular, sets 'impact evaluation' apart from other types of organisational or process evaluation.
- Thirdly, there is more equal weight given to understanding *how* the 'impacts' came about, as much as measuring the *extent* (size) of the impact.
- And lastly, this definition highlights the importance of power dynamics, including who defines impact, and who is affected by the impact (the winners and losers). This is important as the impact on one socioeconomic group may negatively or disproportionately affect another part of the same population.

In thinking through a suitable entry point for discussing ethics in the context of impact evaluation, we found that the language was diverse and sometimes confusing. For the purposes of the workshop, we drew extensively on the work of Simons (2006) which draws attention to the fact that most evaluation and research texts indicate that there is no single, context-free set of (abstract) principles that can be applied to guide ethical judgements in evaluation. Instead, what researchers/evaluators usually encounter in practice are **ethical dilemmas where complex judgements need to be made, taking into account a range of factors and competing interests in a particular socio-political context** (Simons 2006: 243–44). This characterisation captures the contestations and ethical predicaments faced in

---

[1] Presentation by Chris Barnett introducing the Centre for Development Impact at the IDS Annual Review, July 2013.

the course of an evaluation that aims to provide information to inform decision-making from an impartial and independent stance.

Furthermore, we drew extensively on the distinction made by Greene (2006) between micro-level issues and more macro-level ones in 'Evaluation, Democracy and Social Change'. This provided a useful focus for discussions around what ethics may (or may not) entail:

- **Micro issues:** These are the *internal processes of an evaluation*, such as how relationships are established with respondents, and the underlying values around consent, anonymity, transparency, etc.
- **Macro issues:** These are about *how an evaluation relates to society*: to which purpose, and whose interests, should the evaluation serve?

2.1.2 Landscaping study
(Tamlyn Munslow)

The preliminary findings of this scoping study were used to explore the landscape of how ethics might be framed within impact evaluation. The aims of the scoping study were twofold. Firstly, to explore the landscape of ethics in impact evaluation practice: *how are ethical principles currently being applied to impact evaluation?* And secondly, to identify key areas to explore further, such as issues around framing ethics within a broader concept of democracy.

This was a rapid review, designed to capture some of the key emerging issues rather than to provide a robust study of ethical issues as they relate to impact evaluation. The study involved a desk review of the most cited evaluation guidelines and ethical standards; these included the Organisation for Economic Co-operation and Development-Development Assistance Committee (OECD-DAC), the World Bank, the Economic and Social Research Council (ESRC), the United Nations Development Programme (UNDP), the United States Agency for International Development (USAID), the Global Environment Facility (GEF) and DFID, as well as the Joint Committee on Standards for Educational Evaluation (JCSEE). Other material on ethics was drawn from over 60 different educational institutes, 'thinktanks' (including the UEA, the Overseas Development Institute (ODI), IDS and Oxford Policy Management (OPM)), and Evaluation Associations and Societies (including African, American and European, as well as the Swiss, German, United Kingdom, Ugandan and Zambian). This was complemented by primary interviews with ten researchers/practitioners describing their individual experiences and challenges in applying ethics in evaluation.

Five preliminary findings emerged from this study: firstly, there is a wide variation in the definitions and language being used by evaluators and researchers, with many different interpretations of what ethics should mean. Secondly, there seems to be a tension between following a clear set of standards and the role of personal ethics – and particularly between more bureaucratic principles/protocols and responding to the dilemmas often faced in reality. Thirdly, there is a trend towards evaluations becoming more open and transparent with their findings,[2] although this tends to focus ethical concerns (such as an ethical duty to stakeholders or society more broadly) much more on the latter stages of an evaluation rather than the ongoing process. Fourthly, there is challenge in applying more formal Institutional Review Board (IRB) procedures to the emergence of new evaluation methodologies. And finally, there was a tentative finding that suggested that there is a disproportionate emphasis on applying ethics in the evaluation field (due to its shared characteristics with the research profession), whereas the same scrutiny is not necessarily applied to the implementation of development interventions. These points are expanded next:

---

[2] While evaluation findings may be more widely communicated, there is still some way to go in terms of making data available for re-analysis and re-use, as well as information on how the findings were generated (e.g. meta-data).

- **Varied definitions and language.** In the interviews, there were many different interpretations about what ethics means in the field of impact evaluation. Some emphasised more formal protocols and guidance, while others used their discretion to refer to personal and professional ethics. The word 'ethics' sometimes hindered conversations – with many of those contacted declining interviews – while the sheer diversity of interpretations suggests that perhaps this is an area where greater debate and better framing of the issues may be helpful in encouraging and supporting a more widespread and open dialogue.
- **Minimum standards and personal ethics.** While guidance on evaluation principles are widely available, the principles that apply specifically to ethics, rather than other governance or methodological matters, are less clear (such as in the OECD-DAC, DFID, ESRC, UNDP and GEF evaluation guidelines).[3] This lack of clarity is found when trying to differentiate between guiding principles for evaluations more broadly and guiding principles that are specific to ethics. This also means that minimum standards are often open to considerable interpretation.

In practice, it seems that the discretion of the individual evaluator/researcher often played a significant role in addressing ethical issues – and this discretion is often guided by an individual's perspective as to whether they viewed evaluation as a standalone function or part of a broader research agenda where ethics may have already been considered. Where evaluation is considered as a form of research, evaluators tended to draw on standards and institutional setups that are consistent with the research world (guidelines, ethical committees, etc.). Other evaluators, however, viewed evaluation as an extension of public/project management, and tended to draw more heavily on personal and professional ethics. As one interviewee expressed it: 'Ethics depends on how the evaluator makes a decision about what is morally right in light of the policy context that they are working within...' (pers. comm. 2014). Both of these perspectives have limitations, as discussed later.

- **A trend towards greater openness**. From the literature, there appears to have been a more recent shift towards a culture of openness with regard to partnerships, and sharing findings in accessible formats (e.g. e-repositories and portals, open data access, etc.). This has, however, tended to focus far more on the final stages of an evaluation – particularly the 'presentation of findings' phase – and less so in terms of shifting towards ethical ways of working as part of the evaluation process.[4] Furthermore, the trend towards greater openness still tended to focus on the primary power relationship, as findings were most often shared only with the commissioner or key stakeholders, rather than in ways that might engage a broader set of constituents from society.
- **Ethics and emerging methods.** Another theme to emerge from the scoping study was the greater emphasis placed on emerging methodologies, and the corresponding challenge about how to apply traditional (standardised) ethical research standards in such contexts. As one interviewee put it: 'Evaluation designs are increasingly naturalistic and emergent, dependent on the context at hand, and so the ethical issues that arise are different to those currently offered in standard evaluation guidelines' (pers. comm. 2014).
- **Evaluation and implementation.** *Should ethics encompass both interventions and evaluation?* Evaluators/researchers doing evaluations are bound to ethical protocols, and yet implementers are mostly not – despite operating in ethically challenging contexts. While this was not a strong theme to emerge in the scoping study, it does raise some interesting questions and challenges for debate. For one person

---

[3] See OECD (2010), DFID (2011), ESRC (2012), UNDP (2010) and GEF (2007).
[4] The researchers recognise that this finding from the scoping study is skewed by the number of participants interviewed who had worked on DFID open access grants, and/or evidence-based policy grants.

interviewed, it seemed like impact evaluations increasingly had to 'jump through a series of hoops' for ethical approval, while implementers were free to do whatever they want: 'It is one challenge making sure that the evaluator is being ethical, but when those that are implementing the programme do not follow the same set of ethical protocols, this is not under your control…' (pers. comm. 2014).

In summary, the scoping study raised a number of key questions that were discussed at the workshop. These included:

- **Language:** Is it time to find a better way to communicate ethical issues in evaluation (and one that better bridges the more formalised 'research ethics' world and 'evaluation' as practiced by many)?
- **Minimum standards:** Do we have a duty to go beyond what is 'feasible' and 'realistic' as set out in ethical guidelines? How and where do we draw the boundaries? What is the minimum guidance needed?
- **Greater openness:** How might dialogue and public accessibility become key dimensions of ethics during the evaluation process? How might it become better integrated throughout the evaluation process?
- **Emergent methodologies:** How could ethical frameworks respond to new ways of doing impact evaluation?
- **Implementers:** Should ethics encompass both the intervention and the evaluation? What should evaluation have to say about ethical development?

2.1.3 Examples from evaluation practice
(Dr Chris Barnett)

The aim of this session was to provide an evaluation practitioner's perspective, drawing on examples from two contrasting impact evaluations: firstly, a quasi-experimental, mixed method impact evaluation of the Millennium Villages project; and secondly, a case-based impact evaluation of the Tilitonse programme (a governance fund in Malawi). These specific examples were chosen because they are seen as being indicative of a range of approaches and dilemmas that presently exist (see Table 2.1).

Table 2.1   Contrasting examples of impact evaluation designs

Millennium Villages impact evaluation in northern Ghana	Independent impact evaluation agent for Tilitonse, Malawi
Summative evaluation	**Developmental** (formative/summative)
Health, education, livelihoods interventions delivered by single 'project'	**Governance interventions** delivered through a grant mechanism to 50+ grantees
Quasi-experimental (Difference-in-Differences)	**Case-based design**
Counterfactual framework = the difference between a treatment and control group	**Generative framework** = identifying the causal links and mechanisms that explain effects
Large-n quantitative survey to assess attribution (n=2,250 households)… with qualitative methods used to *explain* causal links.	**Qualitative case studies (n=12–15)** to assess strength of causal evidence along the theory of change… triangulated with secondary quantitative/qualitative data.

Source: Presentation by Chris Barnett: 'Two Examples of Impact Evaluations' at the CDI Ethics Workshop, 22 July 2014.

The Millennium Villages is an example of an evaluation that falls under the ethical research procedures typically associated with research studies. All quantitative data collection is approved by an IRB (ethics committee), which requires detailed protocols for data collection, with restrictions on data access to protect 'human subjects'/respondents (e.g. agreements on data security, back-up, storage and sharing). However, there are several limitations to this approach. Firstly, the protocol-driven and somewhat bureaucratic approach has restricted data access for other members of the same evaluation team (e.g. the qualitative researcher's use of household lists from the survey, the evaluation's quality assurer's access to completed questionnaire forms to spot-check quality). Secondly, the IRB approval does not necessarily ensure that ethical standards are met in the field – while in theory, the protocols should ensure this, there is no IRB check on whether they are being followed in the way intended. And lastly, the IRB restrictions around data access give primacy to protecting the respondents (and the research organisation concerned), and this seems out of balance with the broader society interests in data transparency and re-analysis.

The Tilitonse example contrasts this approach 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.

Table 2.2   Contrasting examples of ethical approaches

Millennium Villages impact evaluation	Tilitonse impact agent
Formal: IRB (ethics committee) approves the evaluation	Informal: no formal ethical approval by an external body (standards set by team)
Procedural: an emphasis on protocols, consent, anonymised data, data security and data access	Personal: emphasis on personal and professional conduct in the field, verbal consent, anonymised reporting and feedback to respondents
Limitations: no follow-up in the field (Are protocols followed? Is appropriate consent given?). Can be used to restrict data access and protect the interests (liabilities) of the research organisation.	Limitations: no explicit standard or framework used to guide evaluation practice – a gap between standard guidance and situated practice, with a lot left to the 'professionalism' of the individual.

Source: Presentation by Chris Barnett: 'Two Examples of Impact Evaluations' at the CDI Ethics Workshop, 22 July 2014.

While the particularities of these cases govern the approaches taken, the issues raised do resonate with some of the extreme contrasts that exist in evaluation practice. This diversity of practice is encapsulated in the following quotation:[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.

---

[5] From a blog post comment by Patricia Rogers on the BetterEvaluation website, 2014.

In summary, there is a **great variance** in practice between *highly formalised* (IRB/protocol-driven) standards used in some impact evaluations, and *more informal/personalised* approaches to ethics that may be better situated within a context. The former may draw heavily on the research tradition, but there are also concerns about the extent to which ethical standards have a meaningful influence on practice. In both cases, however, there remains a primary focus on only a few actors: the conduct of the evaluator (or the organisation represented), and the protection of the respondent/subject. There is much less guidance on other potential ethical commitments and how these are worked out in relation to key stakeholders, and society more broadly.

2.2 Plenary session 2: The meaning and value of ethics in evaluation
(Professor Helen Simons)

The first part of this session provided a reflection on ethical guidelines and standards, their role and function. As mentioned earlier, there are many examples of guidelines, principles and standards that have been produced, such as by the DAC Evaluation Network, various agencies (World Bank, UNDP, etc.) and many professional Societies (the UK Evaluation Society (UKES), the American Evaluation Association (AEA), etc.). For the most part, these aim to: promote good practice in evaluation; enhance the status of evaluation as a profession; protect evaluators, participants and the public interest; help build a culture for ethical evaluation; and enhance the management of evaluations.

And yet, standards may also detract from 'good' evaluation. There is after all no single agreed, universal standard; and no single evaluation has yet met all the standards that are specified in these guidelines. So-called ethical guidelines often prescribe too tight a template, or too many indicators, and in essence, set out a false hope which cannot possibly be met. Furthermore, many standards have an absence of ethical statements (Newman and Brown 1996; Fraser 2001a, 2001b), and are easily blurred with the quality standards of an organisation. More often than not, standards are about governance rather than ethics, focusing instead on the quality of the end-product, the audit of a process, or the methodology of the evaluation. Ethics should instead focus on:

- how the evaluation was conducted, such as in fair and just ways, and;
- how the relationships affect the outcomes, taking account of personal and power dynamics.

The argument put forward in this session was that ethics is a relational concept which is primarily about '**situated practice**' (Simons and Usher 2000). In most cases, the ethical principles set out in professional guidance are abstract statements of intent, and it is not always clear to the evaluator/researcher how these should be applied in specific contexts. Instead, there is a need to interpret principles in specific socio/political contexts. In such contexts ethics and politics are often intertwined, with a need to trade-off one principle against another in what is frequently a clash between 'right and right' (Russell 1993).

In order to achieve this, it is often necessary to have criteria and/or an ethical theory to guide ethical decision-making. It is not a case of one individual's perspective against another: the ethical theory guides the choices being made. In short, ethical principles, guidelines, codes, and theories inform and guide ethical behaviour. But ultimately, it is how the evaluator behaves in the field that indicates whether s/he has *acted ethically.* As House (1993: 168) puts it: 'The balancing of such principles in concrete situations is the ultimate ethical act'.

The final part of this session considered the 'democratic evaluation' approach as an example of situated ethics; where democratic ethical theory is central to informing behaviour in the field. Democratic evaluations aim to: (i) democratise evaluation knowledge; (ii) promote dialogue and debate about critical issues important in a democracy; (iii) advance democratic values in how we conduct evaluation; and (iv) inform public decision-making and contribute to the policymaking processes of society. The evaluator's role, while independent, is to negotiate the exchange of evaluation knowledge between different groups and ensure dissemination to a wider public. Democratic evaluation is guided by a set of principles to guide the collection and sharing of information that requires a balance to be struck between a person's 'right to privacy' and the public's 'right to know' (Simons 2010).

As an approach, democratic evaluation gives more weight than is usual in ethical protocols to procedures such as informed consent (e.g. rolling, process consent). It also engages participants in the process, by identifying issues, checking relevance, and giving them a voice in how they are represented. There is also a need to negotiate what data becomes public (on a criteria of accuracy, relevance, fairness), and engage in dialogue with stakeholders on the focus and use of the evaluation.

In summary, this session raised the debate about the need for specific ethical procedures that embody a clear framework of values, recognising that the ultimate ethical decision (in a specific cultural, socio-political context) is about balancing different principles and coming to a judgement that is both professional and political.

2.3 Plenary session 3: Ethics, moral philosophy and international development
(Professor Elliot Stern)

The ethical debate in evaluation has traditionally focused on promoting standards of ethical behaviour among evaluators, such as through standards and guidelines for how evaluators should behave, given the work with the powerful and the powerless. *If we are only concerned with how evaluators conduct themselves – while important – then we will miss major ethical questions that should concern evaluators: what is good development, and how do we recognise it when we see it?* While such moral and ethical dilemmas may not be capable of being resolved, they nevertheless are capable of better explanation.

This session focused on thinking beyond how ethical standards should govern the behaviour of evaluators, to consider what moral philosophy and ethics can offer by way of criteria for judging results-oriented development initiatives. The session argued for a need to rebalance impact evaluation away from empiricism and analytics, and more towards judgement-making. This is because: (i) analytics on its own is incapable of fully answering many evaluation questions; and (ii) a new and still emerging development architecture is shaping a new generation of development policies that are complex, ambiguous and often confusing in their consequences.

The discussion raised the following issues:

- Evaluation ethics is also about the process of knowledge construction: how do we understand what it is we think we are doing? Where do we get our concepts from? How do we relate it to broader bodies of knowledge? What constitutes the right knowledge and the best way of evidencing something?
- Language matters: for example, the current persuasiveness of economic language influences how stakeholders engage in development-related debates.
- It seems that there may be an opportunity to further consider the values related to development and unpick the moral issues at each stage of the process. Based on

this, it could be possible to reconstruct the knowledge and be more explicit about the ethical (and political) basis of this knowledge.

- As evaluators there is scope to ask questions that others may not be able to, and having a language from ethics helps in asking difficult questions, within a bounded space. For example, IDS uses a social evaluation approach in Zambia alongside a much more expensive randomised control trial (RCT) of social protection.
Step 1 – put people at the centre of evaluation. Step 2 – ask them what matters to them and about what the theory of change is (i.e. the RCT has its own theory). The aim is to create a debate between different theories of change, including a deliberative space that allows reflection between the RCT and local people's theory about how change happens.

2.4 Plenary session 4: Reflections on broadening the ethical debate

(Facilitator: Dr Laura Camfield)

This session drew upon the earlier plenary sessions, picking up on the themes about how evaluation might go beyond 'good practice' and ethical duties to society: either in terms of taking on particular values such as democracy, justice, fairness, etc. in the way of working (plenary 2), or as a criteria by which evaluators may better assess and shape development (plenary 3). The aim of the session was to go beyond a more bureaucratic understanding of ethics (protocols, committees, etc.) and consider some of the emerging themes in more depth. What follows is a summary of the group discussions:

- *Is there a 'necessary minimum' we can agree on in relation to ethical practice? What would this look like?* This group discussed the challenges of creating a universal standard, particularly where long lists of principles may be unrealistic and difficult to implement in practice (see plenary 2). Instead, there seemed to be three main points that offered an absolute 'minimum'. Firstly, it seems that the 'minimum' focus for an evaluation should be to **keep people at the centre** – going beyond people as informants or subjects of evaluation research, to focus more on protecting and raising issues related to the people who are meant to benefit from the development intervention whose voices may be marginalised. This builds upon the relational perspective of ethics: although rather than accepting the pre-existing power relations (with evaluation performing a relatively neutral role in delivering empirical findings), it recognises that evaluation can play a more active role in arbitrating different stakeholder perspectives about which findings are important and for whom. Secondly, there was a discussion around **the principle of 'do no harm'**, where both the intervention and the evaluation should adhere to 'minimum rights' not to cause harm – perhaps building on approaches like the Universal Declaration of Human Rights. And thirdly, a key minimum seemed to be to have **an arbitration process for different ethical principles**. The acknowledgement here was that one universal standard/principle would not hold in all contexts – and different ethical principles may be equally important but be in conflict when played out in practice (see plenary 2). Therefore, instead of having an absolute minimum standard, it seemed to be useful to put in place a process whereby standards may be arbitrated in relation to the specific socio-political context.

- *What are the most effective mechanisms to support ethical practice in evaluation (e.g. professionalisation, guidelines, formal review)? How could ethical guidelines/procedures create better evaluations? What are the risks?* This group focused on some of the different aspects about how ethical standards and principles may be put into practice. One major suggestion was to move beyond formalised protocols and IRBs (ethics committees), which tend to disempower the evaluator by

taking a bureaucratic view of ethics. Key points made included: (i) being clear on the professional obligations and standards that need to be adhered to in the course of an evaluation; (ii) communicating the benefits of an ethical process better to all stakeholders; (iii) creating transparent/upfront agreements or memorandums of understanding, so that they provide a reference point when needed during the evaluation; (iv) monitoring the effectiveness of ethical standards on the ground, to check how ethical standards translate into reality; (v) making use of peer reviewing, checking and other forms of auditing, that provide an external prompt around ethical standards; (vi) possibly the use of helplines/information lines where individuals can phone/log/text in feedback and complaints; (vii) using seminars and training events within the evaluation process to better explore ethical dilemmas.

- *Being an 'ethical evaluator' – does democratic evaluation offer an appropriate model? What adaptations might need to be made for work in developing countries?* In this group, the discussion was about the appropriateness and difficulties (or modifications needed) for democratic evaluation approaches to work in the developing country context. It was viewed as appropriate in terms of: (i) promoting participation, inclusiveness and negotiation through an ongoing dialogue, and using existing entry points for that dialogue; (ii) providing a transparent process by making the role of the evaluator more explicit (i.e. their prevailing worldview within a developing country context); (iii) by using stakeholder power mapping (to better acknowledge the different interests in the evaluation findings); and (iv) by keeping records about the evaluation process (e.g. rolling consent). Some of the difficulties in the 'democratic evaluation' approach, and which may need modification were:

  o The need for more feedback loops (perhaps a communications strategy), to ensure that ethical challenges are fed back and dealt with, given the often larger geographic and institutional distances between donors and ultimate beneficiaries.
  o Some adjustment in the approach, given the architecture of development and evaluation, for example, pooled funds, where there are added layers of complexity and less frequent interactions with key stakeholders.
  o The practicalities about how to engage with all stakeholders through more regular 'interim' reporting processes. This was seen as particularly challenging in the international development context, where contacting and having face-to-face interactions with stakeholders is often much harder.
  o What is the role of the evaluator in making recommendations?
  o How does one know that the stakeholder representatives actually represent the stakeholders? This is particularly problematic as donor agencies have differences between headquarters and country-level operations, and government stakeholders may not be able to represent citizen views.
  o Democracy may be viewed as a Western-centric concept – and have different interpretations in different country contexts. It therefore seems to be more important to have agreement over particular ethical principles than use the term 'democratic evaluation'.

- *How can evaluation contribute to a better understanding of ethical development?* This group argued that the starting point for such an approach should be the framing of evaluation as part of the broader (global/national) knowledge system. This framing also needs to take account of the rapidly changing, post-Millennium Development Goal (MDG) architecture – with different actors, and different public/private modalities. The challenge is how to find ways of feeding grounded experience and knowledge of development practice into public debate about 'good' development – while also including the media, but not just as 'PR' (public relations) for aid. This requires revisiting the idea of social contracts, the role of commissioners and policymakers and interrogating dilemmas of aid, trade, security, foreign policy, etc.

3   Conclusions and next steps

One of the key purposes of this event was to consider the vast landscape of ethics and how it relates to impact evaluation, as a first step in preparing for a public event. Greene's (2006) distinction between macro and micro issues, in discussing the role of evaluation in democracy and social change, emerged as a useful framework for thinking about different types of ethical issues:

1.   **Micro-ethical issues:** these issues 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. These issues were not discussed in detail at the workshop, with much of the debate more focused on the disconnects between ethical theory, principles and procedures, and the reality of 'situated practice' – where complex and sometimes competing power dynamics make ethical practice far harder in reality.
2.   **Macro-ethical issues:** these issues focus on how evaluation relates to society: to which purpose, and whose interests, should the evaluation serve? Much of our discussion considered how citizens could be better involved in the deliberation of evidence, such as through more democratic approaches to evaluation – where evaluation performs a role not just as an extension of public sector management, but actually contributes to public debate.
3.   **Development ethics:** this focuses on how evaluation could usefully perform a role in contributing evidence and insights about the ethics of 'good' development. As such, rather than seeing ethics as purely something that is applicable to evaluators only, it is also something to which evaluators could make a more important contribution towards more ethical development.

3.1   Emerging themes

In discussing these different levels of ethical issues, the last session raised a number of possible themes for a subsequent event agenda. These are summarised below, and provide the basis for subsequent discussions and planning for this event.

Firstly, the issue of 'why ethics in impact evaluation *per se*?' was raised. Admittedly, many ethical issues (and by extension, the corresponding structures, guidance and procedures) could equally apply to research more generally, or other forms of evaluation. Indeed, this may be so, but the field of impact evaluation has become particularly driven by methodology and analytics, and there is perhaps a greater need to redress the balance towards value, meanings and judgements – particularly given the complexities of the post-2015 (post-MDG) development field.

Other themes to emerge were:

Theme 1: The role of evaluation in challenging ethical values in international development:

*   What is the role of ethics in development, and why does it matter?
*   What values underpin the way we put together knowledge/evidence of development?
*   How might evaluation challenge and provide insights on the values behind development?

Theme 2: Bridging the divide between ethical theory and practice:

- What is ethics and what theories/values underpin this?
- What non-evaluation guidelines can we draw on (e.g. from other professions)?
- What framework is needed to navigate through the many ethical theories and principles?
- How do we break down barriers between theory and practice to bring learning from each?

Theme 3: Making evaluation more democratic:

- How do we make evaluations useful to those being evaluated?
- How do we usefully share democratic evaluation as one set of principles?
- Who does the analysis of findings, and what ethical issues does this raise for society?

Theme 4: Ethics as a form of 'situated practice':

- How do we apply micro/macro and development ethics in practice?
- How do we negotiate the tension between different principles (e.g. where different principles are in conflict)?
- Who is responsible for evaluation ethics (ethical for whom)?
- What are the challenges of meeting ethical decisions in specific contexts?
- How are ethical decisions made in practice? What concrete examples are there?
- How do we adapt principles for local contexts and local projects?

Theme 5: Ethics and new technology:

- What are the ethics of information technology for evaluation?
- Does this exclude certain groups, or have the potential to include groups not previously reached?

Panel 1 (orange notes)

- Keynote: WHAT IS ETHICAL DEV + WHY DOES IT MATTER?
- HOW DO WE PUT TOGETHER KNOWLEDGE of development?
- IS AN ETHICAL DEVELOPMENT POSSIBLE
- Who is responsible for evaluation ethics? (e ethical for whom)
- ADDRESSING DEVELOPMENT'S Democratic Deficit
- What is ethics & what theories/values underpin this?
- HOW CAN EVALUATION PRACTICE CONTRIBUTE TO MORE ETHICAL DEVELOPMENT
- How to judge an ethical evaluation's intervention

Panel 2 (orange notes)

- Micro/macro + devel ethics → How to apply in practice
- ETHICS 3 topics MICRO/MACRO development good
- Working towards development of guidelines, but only useful if ultimately accompanied by discussion para / with concrete ex's of dilemmas in e/3 country contexts
- CREATE A DIALOGUE: - How are ethical decisions made in PRACTICE?
- How to negotiate the tension between different principles ('right+right' claims)
- Adapting principles for local contexts/projects
- Challenges of making ethical decisions in specific context. IMPLEMENTATION
- HOW TO MAKE EVALUATIONS USEFUL TO THOSE BEING EVALUATED
- Democratic evaluation as 1 set of principles — useful to share widely. Could be in case study format (ie telling through actual application)

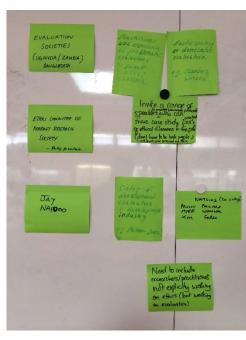**Panel 3 (orange notes)**

- HOW TO BREAK DOWN BARRIER THEORY/PRACTICE BRING LEARNING FROM EACH
- Clearer framework to navigate through many ethical theories + principles
- What guidelines can be drawn on? (not just evaluation related)
- ETHICS OF IT for evaluation or evaluating IT type development
- A focus on the ethical issue around who does the analysis of findings

Panel 4 (green notes)

- EVALUATION SOCIETIES (UGANDA/ZAMBIA) BANGLADESH
- Practitioners with experience of problematic evaluation
- Participatory or democratic evaluation e.g. Clarus Simitus
- ETHICS COMMITTEE OF MARKET RESEARCH SOCIETY - Polly for contact
- Invite a range of speakers who can share case study of ethical dilemmas in the field
- JAY NAIDOO
- Critic of development evaluation + development industry e.g. Pedram Jaus
- Natsios (ex USAID) Alison Mike Kim Pralee Wallace Feres
- Need to include researchers/practitioners not explicitly working on ethics (but working on evaluation)

Annex 1 Workshop agenda

Morning session
The interactive sessions will offer rich opportunity for interactive discussion with **20–30 minutes' presentation and 30–40 minutes' discussion**.

9.30am–10.00am

1. **Opening: Agenda setting**
   What questions do we need to ask in relation to ethics in evaluation? Chris Barnett will outline the structure of the day and our task of laying the foundations for a larger event in October/November 2014.

9.45am–10.45am

2. **Chris Barnett, with Tamlyn Munslow: What is happening in the world of impact evaluation ethics?**
   Landscape scanning and reporting on primary interviews: how do researchers and practitioners understand ethics in impact evaluation (IE)? What is specific about ethics in IE relative to other areas of research and practice? What procedures/guidelines exist? What is the view of commissioners? What are common problems? This information is being collected through a small-scale research project entailing interviews and documentary analysis. Chris Barnett will also highlight issues from practice, framed around contrasting Millennium Villages and another evaluation.

   Plenary discussion – Chris Barnett

10.45am–11.15am
Break

11.15am–12.15pm

3. **Helen Simons: Ethics in democratic evaluation**
   This presentation will be in two parts. First, it will address the meaning and value of ethics in evaluation, moving beyond the claims of many standards, guidelines and principles that state that they advance ethical practice. More often than not, these have a different function, that of governance or protection of institutions or professional fields (witness the growth of IRBs and evaluation society guidelines over the past ten years). At best some have an ethical intent; many are stated as principles. Important as these are to guide practice, principles are abstract statements of intent. Ethics itself is a *situated, relational and political practice.* Principles need to be interpreted in different socio-political contexts and it is sometimes necessary to trade one off the other in making an actual ethical decision. In the second part, from her experience conducting democratic evaluation in programme and policy evaluation, Helen will explore the values and principles of democratic evaluation and the ethical procedures underlying this approach for the conduct of ethical evaluation.

Plenary discussion – Laura Camfield

12.15pm–1.15pm
Lunch

Afternoon session

The interactive sessions will offer rich opportunity for interactive discussion with **20–30 minutes' presentation** and **30–40 minutes' discussion.**

1.15pm–2.15pm
4.   Elliot Stern – Ethics in impact evaluation
As this workshop is intended to scope a future larger event I have tried to identify a number of possible avenues that a future event could explore in greater depth. Firstly: most discussions in the evaluation community are concerned with the behaviour of evaluators. How they should behave in a world of uneven power, what should be the standards and ethical codes that should govern their relationship with those being evaluated and those commissioning evaluations. I would like us to also consider what moral philosophy and ethics can offer by way of criteria when judging results-oriented development initiatives. Secondly: what are the implications of the new/emerging development architecture and development goals? I would argue that this has implications both for the ethics of development and the ethics of the evaluator. I will illustrate with a number of scenarios and vignettes.

Plenary discussion – Kathi Welle

2.15pm–2.45pm
Break

Interactive afternoon session

2.45pm–3.45pm
5.   Interactive session: Deliberative methods and ethics as our duty to society
In thinking about ethics, do we need to go beyond 'good practice', which can seem like a bureaucratic understanding, to include broader principles of participation, ownership, informed debate, etc. as reflected in the focus of the forthcoming European Evaluation Society (EES) conference on independence, participation, and partnership? How does one do this in practice?

Group work and discussion – Laura Camfield

3.45pm–4.30pm
6.   Final discussion: Ways to take this forward
Next steps in terms of planning a larger event in the autumn, looking at how to incorporate ethical concerns in the process of designing, planning, conducting and presenting an evaluation (for example, what of the questions discussed today should we focus on, who should be invited to speak and participate?). The aim would be to complement and extend the aforementioned debate at the EES conference. How is ethics framed in relation to society and accountability? How do we verify what happens in practice? How do we broaden out beyond internal procedures?

Facilitated session – Chris Barnett
4.30pm
Close

Where: IDS Convening Space
                    Institute of Development Studies
                    University of Sussex

How to get there: Attached

Annex 2   Workshop participants

Dr Chris Barnett	Director of the Centre for Development Impact (CDI)
Kate Bingley	Monitoring and Evaluation Advisor, IDS
Dr Laura Camfield	Senior Researcher, University of East Anglia
Prof James Copestake	Professor of International Development, University of Bath
Dr Philly Desai	Director, Turnstone Research
Maria Kuss	Research Officer, IDS
Prof Allister McGregor	Research Fellow, IDS
Tamlyn Munslow	Research Officer, IDS
Emily Richardson	Consultant, Itad
Prof Helen Simons	Professor of Education and Evaluation, University of Southampton
Gavin Stedman-Bryce	Head of Research and Evaluation, Pamoja
Prof Elliot Stern	Emeritus Professor of Evaluation Research, Lancaster University; Visiting Professor, University of Bristol
Dr James Sumberg	Research Fellow, IDS; Director of the DFID Accountable Grant
Dr Patta Scott-Villiers	Research Fellow and Ethics Convenor, IDS
Dr Katharina Welle	Senior Consultant, Itad

Bibliography

DFID (2011) *Ethics Principles for Research and Evaluation*,
www.gov.uk/government/uploads/system/uploads/attachment_data/file/67483/dfid-ethics-
prcpls-rsrch-eval.pdf (accessed 30 September 2014)

ESRC (2012) *ESRC Framework for Research Ethics (FRE) 2010*,
www.esrc.ac.uk/_images/framework-for-research-ethics-09-12_tcm8-4586.pdf
(accessed 30 September 2014)

Fraser, D. (2001a) 'Beyond Ethics: Why We Need Evaluation Standards', *Evaluation Journal of Australasia* 1.1: 53–8

Fraser, D. (2001b) 'Development of AES Standards', *Evaluation Journal of Australasia* 1.1: 59

GEF (2007) *GEF Evaluation Office Ethical Guidelines*, Evaluation Document No. 2,
Washington DC: Global Environment Facility Evaluation Office,
www.thegef.org/gef/sites/thegef.org/files/documents/EthicalGuidelines.pdf
(accessed 30 September 2014)

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

House, E.R. (1993) *Professional Evaluation: Social Impact and Political Consequences*, London: Sage

Newman, D.L. and Brown, R.D. (1996) *Applied Ethics for Program Evaluation*, Thousand Oaks CA and London: Sage

OECD (2010) *DAC Guidelines and Reference Series: Quality Standards for Development Evaluation*, Development Assistance Committee of the Organisation for Economic Co-operation and Development, www.oecd.org/dac/evaluation/qualitystandards.pdf
(accessed 30 September 2014)

Russell, C. (1993) *Academic Freedom*, London: Routledge

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

Simons, H. and Usher, R. (eds) (2000) *Situated Ethics in Educational Research*, London: Routledge

USAID (2013) *Technical Note: Impact Evaluations*,
www.usaid.gov/sites/default/files/documents/1870/IE_Technical_Note_2013_0903_Final.pdf
(accessed 30 September 2014)

Institute of Development Studies

Brighton BN1 9RE

T +44 (0)1273 606261
F +44 (0)1273 621202
E ids@ids.ac.uk
www.ids.ac.uk