Ethics in Evaluation

Some of the most intractable ethical problems arise from conflicts among principles and the necessity of trading one off against the other. The balancing of such principles in concrete situations is the ultimate ethical act. (House, 1993, p. 168)

Ethics is a branch of politics. That is to say, it is the duty of the statesman to create for the citizen the best possible opportunity of living the good life. It will seen that the effect of this injunction is not to degrade morality but to moralise politics. (Thomson, 1953, p. 26, commentary on The Ethics of Aristotle)

Introduction

The two quotations above underscore much that will be explored in this chapter as evaluators go about their business of evaluating programmes and policies in the socio-political contexts in which they are commissioned, enacted and disseminated. Drawing attention to the interdependence of politics and ethics and conflicts among principles, highlights the unique nature of the evaluation task and the key responsibilities of the evaluation role. Evaluation involves at least four levels of social-political interaction - with government and other agency policy makers who commission evaluation; with participants in the programmes, policies and institutions evaluated; with the evaluation profession; and with the wider audiences to whom evaluators in a democratic society have a responsibility to report. Evaluation has to operate in this multilayered context of different interests, providing information to inform decisions while remaining independent of the policies and programmes themselves. In such a context it is not surprising that ethical dilemmas arise as to which is the best course of action to take.

Most evaluation and research texts where ethics are mentioned start with a reminder that there is no context-free abstract set of principles that can be applied to guide ethical decision-making in evaluation. What we encounter in the practice of evaluation are
ethical dilemmas, where we have to make a complex judgment, a choice between alternative courses of action, taking into account a myriad of factors - social, personal, political, cultural - that are pertinent in the particular [p. 244 ↓] context (House, 1980; Lincoln, 1990; Mabry, 1999; Pring, 2000; Simons & Usher, 2000).

Take, for example, the following dilemma. A powerful institution, sponsor of a multi-country evaluation study, threatens to hold back payment to colleagues employed by them on projects which have been completed and independently evaluated, until the director of the evaluation agrees to the changes the institution wants made to the final evaluation report. The participants in the projects are dependent on release of the evaluation report for payment and for their livelihood. The institution wants the evaluation to blame certain people in the report. The evaluator argues for the fairness and balance of the report, as he sees it, to be maintained. How should the evaluator respond? Give in to the institution's wishes so that the participants are paid, though it will compromise the integrity and impartiality of the evaluation? Or continue to negotiate for a reasonable outcome that will maintain fairness to individuals and how they are represented and ensure credibility to external audiences and international sponsor, though it will delay payment to participants and have consequences for them and their families? These are not the only options available in resolving the dilemma. They are merely illustrative of how the argument could run.

Or take the dilemma Morris (2004a, p. 236) raises as to whether or not it is appropriate to have a client representative on the steering committee of an implementation evaluation of an agency providing community-based residential services to the chronically mentally ill. The evaluator, who has experience working with community-based programmes for the mentally ill, thinks it is appropriate to have such a client representative. The director of the agency does not, believing that the psychiatric disabilities of the agency's clients would make it extremely difficult for them to be productive and that it might end up doing more harm than good.

Resolving dilemmas such as these is rarely a choice between right and wrong. This would be relatively easy to resolve. More often than not, it is a choice between mutually conflicting principles, what Russell (1993) has called a “clash between right and right” or House (1993) “trading one off against the other.” In such situations, principles are
needed to guide action but they do not determine what is the “right” action to take. How then can the evaluator be assisted to conduct evaluations that are ethically defensible?

Evaluation textbooks on the whole offer little guidance on this issue. Newman & Brown (1996) surveyed 21 evaluation texts from 1972 to 1994 and found that 80 percent did not mention ethics at all and those that did averaged about four pages per book. A survey ten years on may reveal more attention to ethics, although the contemporary context for evaluation has led to procedures and mechanisms that frequently have more to do with control and surveillance than principles for ethical behaviour. In the last ten years we have witnessed an intensification of accountability and managerialism in institutions, an increase in governance and quality assurance of programmes, and a preoccupation with audit, monitoring, regulation, and control.

The rise of evaluation standard setting by evaluation associations and the intensification of ethical committees and institutional review boards (IRBs) within institutions may be one response to this external pressure for regulation by government and other agencies. However, it is important to distinguish institutional mechanisms for oversight of evaluation which have a regulatory function from the trend in the evaluation profession for setting ethical guidelines, principles, or standards. The latter are properly part of the self-regulation and self-accountability demanded of a profession with responsibility to its members, those they evaluate and the public. It is a separate question, however, to what extent ethics is a central feature of such guidelines or a subset of more general standards.

Ethics is about how we behave or should behave as individuals and as part of the society in which we live in interaction with others. It is, at one level, a personal morality, to do with the rights and wrongs of our actions, pervading our everyday life. Guiding precepts are honesty, sincerity, kindness, respect for persons, being honourable. “Do unto others as you would have them do unto you” is a golden rule, a categorical imperative that often guides our action and enables us to “know” we have done the right thing.

As members of a liberal democratic society we are also likely to share certain principles about how we should behave and how we should treat people in the conduct of our
private and public lives. Such principles would include respect for persons, respect for “truth,” equality, justice, freedom and autonomy.

However, as evaluators and as a profession we also share principles about how we should conduct and disseminate evaluation. This is not just a question of our personal predilection or moral stance. It is a question of community - promoting and underpinning what it means to act ethically and promote “right” practice in the professional field of evaluation. This will incorporate public criteria, such as those noted above, and in practice involve recourse to the sources of our personal morality. But there is a professional ethic or ways of thinking ethically concerned with the work that evaluators do. It is this that is the main focus of this chapter.

Characteristics of the Field of Evaluation

As much of the literature on ethics is applicable to all social research, I first outline some identifying features of the field of evaluation to provide a context for discussing ethical evaluation. I am aware that not all evaluators may subscribe to these characteristics. Perspectives on the nature and function of evaluation and the role of evaluation in society are contested and differ according to different world views, preferred methodologies, and in different cultural contexts. The characteristics below represent for me a coherent set of statements for the conduct and dissemination of evaluation in a liberal democratic society.

First, the role of the evaluator is to elicit and make public the essential values and merits of programmes so that individuals and groups can contribute to informed policy-making and debate. In this sense “evaluation is a public decision procedure” (House, 1980), and the role of the evaluator that of the “public scientist” (Cronbach et al., 1980), providing an evidence base for informed judgment and public dialogue.

Secondly, evaluation is about judgment, ascertaining the merit or worth of a programme, distinguishing good from bad. Different methodological approaches consider the role of judgment in different ways but at a fundamental level, the task is to discern quality. In the last four decades in which evaluation has emerged as a separate field of inquiry, the roles evaluators can adopt and the methodologies they can
employ have proliferated enormously. The choice of evaluation roles is a moral choice for society. And the choices evaluators (and sponsors) make about the methodologies they value have political implications.

Thirdly, evaluation is inherently political, concerned with the distribution of power and the allocation of resources and opportunities in society. That is its purpose, to assist choices between different courses of action by collecting and presenting evidence relevant to these choices. Evaluation has consequences for who [p. 246 ↓] gets what, whose interests are served in an evaluation, who stands to gain or lose by the findings of an evaluation and so on. Evaluation cannot be value free, nor can it safely be left to the personal values of the evaluator nor confined to the interests of those who have the power to commission it.

Fourthly, evaluation needs to be independent - conducted “without fear or favour” to any one group. He who pays the piper should not call the tune. Evaluation cannot be bought. It should be sponsored as an impartial service to all stakeholders in the matters under review. It should envision accountability as a two-way process that includes reporting to the public about government-sponsored policies and programmes that are conducted in the public interest. In this context “respect for truth,” “fairness,” “justice,” “equality” are all generic ethical principles to underpin an independent evaluation process.

Finally, where evaluation is funded by public money, ethical issues arise not only in relation to contractual obligations but also the nature of the topic being investigated. Are there some evaluation contracts one would refuse to take on ethical grounds, such as when certain groups might be put at risk or disadvantaged?

In the following sections I indicate the various routes and stances evaluators can take in deciding how to conduct evaluation that is ethically defensible. First, I examine the different ways in which ethics has been conceived professionally, including the generation of standards, principles, codes and guidelines, as well as ethical committees and institutional review boards. Second, I briefly outline different ethical theories and offer some examples of how they might provide justification for ethical decision-making. Third, I draw attention to the interrelationship of politics and ethics indicating, through a brief account of democratic approaches to evaluation, how ethics may be procedurally
managed in specific socio-political contexts. Finally, I re-examine certain procedural ethical conventions, focusing on those aspects that present particular ethical issues in evaluation.

Professional Ethics

Evolution of Ethical Standards/Codes/Principles in Professional Societies

Most professional associations in disciplines from which evaluators stem and evaluation societies now have written ethical guidelines/principles/standards/codes to facilitate ethical practice. Such guidelines traditionally embody a normative ethics - concerned with how people ought to behave, in this case in conducting and disseminating evaluation. However, there is little consistency in how these terms are used. According to Newman & Brown (1996, pp. 21–22) the terms - rules, codes, standards, principles, theories - are distinguished by degree of specificity and purpose. Their summary may be useful.

Ethical rules are specific statements about ethical behaviour; ethical codes are compilations of ethical rules. Ethical standards can be synonymous with ethical rules and codes but may go beyond that definition to suggest model behaviour. Ethical principles are broader than rules serving as the foundation for codes. Principles stand as models of behaviour and practice, providing and encompassing not only situational rules but also serving as guides for unspecified practice. Ethical theories provide a justification for how we make ethical decisions and aid us in resolving conflicts among principles or rules. (ibid., p. 22)

In practice some evaluation societies have set standards to judge the quality of the evaluation and product (e.g., Joint Committee Standards). Others prefer more general statements of principle for the conduct of evaluation (e.g., Australasian Evaluation Society, Canadian Evaluation Society), accompanied in some cases by
guidelines for interpreting the principles in practice. Yet others are couched in terms of more regulative rules or codes which promote and protect the profession and the public and to which members of a society must subscribe.¹ In this chapter I shall use the term guidelines to cover all these attempts to set norms for the profession (see also Bustelo, 2004) but make the distinction between standards and codes that are quite specific, often prescriptive, and reflect model behaviour (Newman & Brown, 1996, pp. 21–22) and principles that are more general, often normative, and, most importantly, aspire to good practice.

The evolution of ethical guidelines in evaluation societies and in other professional associations (see e.g., Goldner, 1967; Simons, 1995) has been gradual and for the most part has only occurred in the last ten to fifteen years.² The exceptions are the Joint Committee Standards in the US (1981, revised in 1994) and the standards of the Evaluation Research Society (ERS, 1982), which merged with the Evaluation Network to become the American Evaluation Association (AEA) in 1986. The AEA chose not to adopt the old ERS standards but rather to develop its own guiding principles (Fitzpatrick, 1999, p. 6), drawing attention perhaps to the distinction between standards and principles referred to above. In setting up a Task Force for this purpose in 1992, the AEA Board “specifically instructed the Task Force to develop general guiding principles rather than specific standards of practice” (AEA, 2004, p. 6). The AEA has not formally adopted the Joint Committee Standards, though it does support the Joint Committee’s work and has a representative on its committee (http://www.eval.org/Publications/publications.html).

The resulting AEA’s Guiding Principles for Evaluators were adopted by the association in 1994 (http://www.eval.org/Guiding%20Principles.htm) and address both behavioural (as in the importance of “systematic inquiry”) and normative (as in the importance of “respect for people”) strands of evaluation practice. The most extensive of the current published guidelines remains the Joint Committee on Standards for Educational Evaluation (1994), which also address both standards of practice (e.g., “accuracy” and “feasibility”) and principles of practice (e.g., “propriety”). These standards have been adopted and/or adapted by several evaluation societies (including the Swiss Evaluation Society, and the German Evaluation Society).
The difference between *standards/codes* and *principles* may perhaps best be exemplified in the evolution of ethical guidelines within the Australasian Evaluation Society (AES). The AES started with an interim code of ethics but, failing to gain endorsement on standards of behaviour for society members, developed a set of “Guidelines for the Ethical Conduct of Evaluation” framed around general principles and procedures for the conduct of evaluation. These were endorsed by the AES Board in 1997, accompanied by a statement which indicated that they complemented the Joint Committee Standards. A Code of Ethics incorporating the former Guidelines was eventually endorsed in 2000, reflecting the need to protect the public and the profession in a changing context ([http://www.aes.asn.au/about/ethics.htm](http://www.aes.asn.au/about/ethics.htm)). The development of specific AES quality standards has been on the agenda since the late 1980s but has not yet reached resolution. Fraser (2005) suggests that this is partly due to the growth of quality standards and frameworks in several state government departments and the emerging concern of indigenous communities to develop culturally relevant quality standards (Fraser, 2005, p. 78).

There are a number of reasons why professional associations took time to develop or adopt a code of ethics for their members. First, the field of evaluation practice is complex and any code cannot anticipate all contingencies nor offer guidance for mediating conflicts that may arise. Second, codifying behaviour implies that sanctions might be imposed on those who do not comply. While peer group evaluation is acceptable in some situations - for example, anonymous refereeing of proposals - pronouncements on unethical behaviour that would lead to sanctioning the right to be evaluators is likely to be perceived by colleagues as a step too far, indicating lack of trust in one's personal ethics and competence. It is also the wrong end of the process. The aspiration in such codes, many would argue, should not be to sanction bad practice or ensure that evaluators do nothing wrong through “commission or omission” (Newman & Brown (1996, p. 3), but rather to promote good practice.

Third is the difficulty of actually enforcing a code of ethics and who would do so. Fourth, it may have simply been too soon, not a priority for an association establishing itself within a still-developing professional field, and where there may have been little
consensus over what constituted appropriate behaviour or practice (Fitzpatrick, 1999, p. 12; Fraser, 2005).

The increase in published ethical guidelines by evaluation societies in the past decade may be a reflection of the rapid growth and/or “hallmark of maturity for a profession” (Newman & Brown, 1996, p. 19). Professional associations may have become more confident about the issues which united their members and realised it was time to assert a collective intent to serve the public interest in governance of and access to quality research and evaluation. Institutional norms, as Bustelo (2004, p. 6) points out, also help to create an evaluation culture, share language and terminology, define the field and provide a means of professional socialisation and training.

However, there are also political and economic factors. In the UK, contractual arrangements for funding evaluation became more prescriptive, restricting reporting (written or spoken) stipulating methodologies, and prescribing ownership and copyright of data (Simons, 1995, pp. 438–440). Guidelines were required to protect the role of evaluation in producing impartial evidence to inform public programmes and policies. Increasingly it was possible to cite evidence of evaluation reports being censored by non-publication, delayed until the issue was no longer relevant, or edited by the sponsor. This resonates with Morris & Cohn’s (1993, p. 18) US research on ethical problems identified by evaluators, the most frequent and serious of which was presentation of evaluation findings.

A further factor in the US relates to economic and structural change in the field. Many evaluations are conducted by private research firms that hire evaluators who often have no specific training in evaluation and spend only a few years in the field before moving on. This high turnover of evaluators makes ethical problems possibly more salient and the corresponding need for guidelines on quality more imperative. The more evaluation is privatized the more guidelines are needed to generate trust in the process and quality of the products of evaluations. Fraser (2005, pp. 68–69) also makes the point that evaluation rises and falls in periodic economy drives in government, resulting both in the loss of skilled practitioners and producing a “radical loss of collective memory”.

In recent years, the number of national evaluation societies that have introduced “standards” for their profession has increased and there have been efforts to examine
their universality (Picciotto, 2005; Russon & Russon, 2005). Given that many build on the Joint Committee Standards, it is not surprising to find similarities among them. However, Russon (2005) concludes, in an examination of the development towards international standards, that the utility of standards is more likely to be enhanced when they maintain a connection with the culture in which people live and work. Similarly the European Evaluation Society has chosen not to advocate common standards preferring national societies, at this stage, to be responsive in their development of standards and guidelines to their different cultural contexts (Stame, 2004, p. 506; http://www.europeanevaluation.org/docs/standardspolicyv218022004.pdf).

**Content and Purpose**

A major purpose of ethical codes in professional associations is “to socialize and educate the practitioner about common standards,” a process more likely to be successful through development of “inner controls’ than the ‘external management of conduct”’ (Plant, 1998, p. 165, quoted by Fitzpatrick, 1999, p. 12). This distinction of the locus of control for ensuring ethical practice - and whether codes are educative or regulatory - is often not clear in published guidelines. The choice of the word “principle” or “guideline” rather than “standard” may be a reflection of educative self-regulatory intent. Standards or codes are often assumed to have more of an institutional regulatory function. However, in practice, since few professional societies have instituted enforcement of standards (Fitzpatrick, 1999, p. 12), their use has been educative by default, if not by intent. This may be because the preferred route is development of “inner controls,” though the more likely reason for the absence of enforcement mechanisms in the AEA Guiding Principles Fitzpatrick suggests may be “the need to reach consensus on the meaning and application of the principles, the continuing tensions among the diverse paradigms used in evaluation, and the relative newness of the evaluation profession” (Fitzpatrick, 1999, p. 12).

Whichever terminology is used, ethical guidelines vary on a number of dimensions. These include how they distinguish between ethical-moral issues and scientific-methodological issues (the degree to which methodological competence is regarded as an ethical issue, for instance); whether they address ethical issues for all stakeholders
or solely those of professional evaluators and whether and how they acknowledge difference.

Common to many ethical guidelines are the principles of mutual respect, non-coercion and non-manipulation, justice, equality, and beneficence (see, e.g., Clark, 1997; Fitzpatrick, 1999; House, 1993; National Commission for the Protection of Human Subjects in Biomedical and Behavioral Research, 1979). House (1993) also includes support for democratic values and institutions. The AEA guidelines have five principles - Systematic Inquiry, Competence, Integrity/Honesty, Respect for People and Responsibilities for General and Public Welfare. The AES guidelines also have five statements of principle covering fully informed consent; opportunity to weigh up and reduce anticipated risks; respect for the rights, privacy, dignity and entitlements of those affected by and contributing to the evaluation; judgments based on comprehensive, sound information; and fair and balanced reporting.

These principles on the whole stem from a Western view of the world and, while they embrace respect for diversity and the norms of different cultures and subcultures, few explicitly address the inherent nature of different cultures in their statements of ethical guidelines. The exception is the evolving ethical guidelines stemming from indigenous communities (see, e.g., AIATSIS, 2000; Burchill, 2004; Mataria, 2003; Nga Ara Tohutohu Rangahau, 2004; NHMRC, 2003). These guidelines challenge evaluation practice, as Smith (1997) has argued in relation to a Kaupapa Maori research framework, to move beyond western epistemological understanding. Ethical guidelines from these indigenous groups emphasize the familiar values of respect for persons, equality, non-coercion and non-manipulation. However, they go beyond the individual and his/her rights to privacy, dignity and justice to place greater emphasis on recognizing and respecting indigenous knowledge systems and processes, uniqueness of peoples, social cohesion, cultural property rights and reciprocity, both in the process of evaluation and in the distribution of benefits from research and evaluation.

For example, included in the NHMRC (2003) guidelines are reciprocity, survival and protection, spirit and integrity. It is not just cultural sensitivity evaluators need to be attuned to here. The guidelines that accompany these principles emphasize - beyond the need to become familiar with cultural customs, property rights and ethical norms
of the society - the very manner in which relationships are built and the culture of a people is sustained. In some indigenous cultures, the making of decisions is a collective process involving substantial deliberation among elders and leaders. In such cases time becomes an important ethical consideration in accessing cultural wisdom and, in some instances, determining which methodologies can be used. Take, for example, the Kaupapa Maori framework prescribed by Smith (1997). The second precept *Kanohi kitea* ("the seen face;" a requirement to present yourself "face to face") or the third, *Titiro, whakarongo* ("look. listen … then speak") (cited in Mataria, 2003, p. 6), suggest that only certain methodologies are appropriate when evaluating programmes in this culture.

**Limitations of Ethical Codes**

Antagonists of ethical codes point to their self-serving nature designed to protect the interests of the profession rather than the researched, their inherent conservatism when built on consensus, and the regulation that may accompany their use. Some authors go further (see, for instance, Punch, 1986) and suggest that such codes threaten to restrict research, create barriers where none exist, and inadvertently protect the powerful rather than the weak.

As a guide for ethical practice in the field, published guidelines have three further difficulties. The first is that ethics do not feature as largely as codes/standards on other dimensions of practice. In the informal analyses of codes conducted by Russon & Russon (2005), standards related to ethics were far fewer than other dimensions such as methodology, quality, utility, and contractual relationships between sponsors and evaluators (see also Fraser, 2005, p. 77, in the same volume). And Newman & Brown observe that in the Joint Committee Standards, "the use of the word ethics is assiduously avoided" (Newman & Brown, 1996, p. 21), though embedded within them, in the section on propriety for instance, are standards that might be considered ethical in nature.

Secondly, in focusing on what to do in order to avoid harm, guidelines frequently do not offer positive guidance to promote "good," and so an opportunity is lost to advance ethical practice in evaluation. Where guidelines are generated from actual
case examples and where the intent is educative rather than regulatory, this is less of a problem. As noted above, however, the intent of a given set of guidelines is not always clear.

Thirdly, there is the obvious nature of codes themselves. Codes cannot capture, nor could they hope to, the “messy” reality of the field, the unanticipated effects, conflicting political priorities and all the personal factors - fears, anxieties and differences involved in the social/political engagements of evaluation on the ground (Mabry, 2004; Worthen & Sanders, 1987).

Ethical Committees/Institutional Review Boards

Ethical Committees or Institutional Review Boards (IRBs) have long been established in the field of medicine to protect participants from harm in the conduct of medical research, especially in clinical settings. Increasingly, they are being set up in the social sciences and other professional fields. While set up by institutions, not professional associations, they are another element in the armoury of activities designed to improve ethical practice.

Ethical committees or IRBs exist to ensure that any research or evaluation proposal has considered the ethical issues that are likely to arise in an evaluation or research study and has developed protocols to protect participants. With the best intent, they contribute along with other practices to a growing sophistication of sound ethical research practice.

However, many of the issues that worry people in relation to codes of evaluation practice also apply to ethical committees. In their formulation and enactment - usually as a lengthy set of questions - and the time they take to process, IRB forms can often seem unduly restrictive and regulatory, offering power to the gatekeeper to control the evaluation even though their overarching intent is to govern good practice.

Five particular concerns have been raised over such committees or IRBs. The first is that this vigilance is at the front end of the research or evaluation. It is not possible
either to anticipate all the risks that might arise nor develop protocols to manage all the interpersonal relationships and situations that may affect ethical decision-making. Secondly, as Nelson (2004, p. 210) has pointed out, one of the most consistently cited principles adopted by these committees, “respect for persons,” cannot entail respect for every human action, but the critical point, he says, is that IRBs are “ill equipped to negotiate the difference”.

Thirdly, and partly because ethical committees have their origin in the field of medicine, to protect participants from harm in the conduct of medical research, such committees are frequently more familiar with quantitative methods and randomized controlled trial designs. Qualitative researchers and evaluators often have to explicate their methodologies and respond to questions that do not easily transfer to qualitative evaluation (Ritchie & Lewis, 2003, p. 63).

Fourthly, increasingly such committees also act as the guardians of research design and of what is to count as research methodology. Lincoln & Tierney (2004) point out that in the US as a result of pressure from the political right to discredit certain methodologies and a heightened sense of legal issues surrounding medical protection, “the stances of IRBs have shifted from assuring that human subjects’ rights are protected towards monitoring and censoring, and outright disapproval of projects that use qualitative research, phenomenological approaches and other alternative frameworks for knowing and knowledge” (Lincoln & Tierney, 2004, p. 220). This inhibits freedom to research, especially topics that may be sensitive and, in curtailing access to certain kinds of knowledge, is anti-democratic.

Fifthly, some have claimed (Furedi, 2002) that ethical committees are in practice not merely concerned with moral dilemmas, but are acting as gatekeepers, defenders of reputations, and to prevent litigation. Some IRBs are quite open about acknowledging that their main concern is to protect the institution from damage. This is a fundamental shift, Lincoln & Tierney (2004, p. 220) state, from [p. 252 ↓] the original purposes of such committees to ascertain risks and assure that informed consent was adequate to prepare human subjects for associated risks. Christians (2000, p. 141) has also noted that while IRBs ostensibly exist to protect the subjects who fall under approved protocols, in reality they protect their own institutions. Where such gatekeeping is the norm, the function of such committees has become one of research regulation, part of
the culture of audit and managerialism, and is not necessarily concerned with ethics at all.

As evaluators and supervisors of evaluations, we have a role to play in trying to ensure that ethical committees and IRBs act with integrity in relation to their original purpose and do not disadvantage methodologies and forms of knowing that are not currently privileged by government dictate or custom. Not only should we continue to articulate and argue for methodological inclusion, but we should also advocate for consumer representation on such committees. The former will further understanding of different methodologies, the latter demonstrate an ethical commitment to representation of different voices. Both can counter a prevailing dominance of a specific methodology and the self-interest of such committees in safeguarding their institutions.

Role of Ethical Theories in Ethical Decision-Making

It may not be obvious which ethical theory guides our judgment in the field, as theories tend to come into play at a distance. In practice there are several levels at which our ethical behaviour is influenced. Firstly, we are guided by those moral and professional dispositions - qualities of character and virtues (Schwandt, 2001) - which have been cultivated over time and which we have internalised in concepts such as integrity, trust and reliability. Secondly, there are the broad principles we accept as guiding behaviour, some of which have been built into the ethical codes mentioned previously and which we adopt before entering the field. Thirdly, there are the particular decisions we actually make in the field taking into account the exigencies of the case guided by principles and by ultimate recourse to preferred ethical theories. In exploring possible choices of ethical theories to justify decision-making in evaluation, I broadly follow the categories of ethical theories cited by Newman & Brown (1996, p. 24). These categories are based on key criteria for deciding whether the behaviour is right or wrong - consequences, duty, rights, social justice, and the ethics of care, which has also been called relational ethics (Mabry, 2004).
Different Ethical Theories and Examples

The ethics of consequences stress utility and outcomes. The decision is right based upon a consideration of its consequences, although one has to ask consequences for whom and in whose interests. Utilitarian ethics - stressing the greatest good for the greatest number - can help to address this question in giving equal consideration to all. However, in certain situations the end could come to justify the means, for instance, where advancement of knowledge is the rationale for deception to gain access. A utilitarian response to the ethical dilemma posed by Morris (2004b, p. 383), in which the new evaluator recruit uncovers deficiencies in the design of the evaluation being conducted by her agency, might take the view that it is better to protect the integrity of the evaluation profession or a larger number of stakeholders than protect those in the immediate programme team from embarrassment through exposure. A relational ethics justification would lead to a different outcome, one that would focus on care for the immediate information sharers and preserve working relationships with colleagues rather than insist on methodological rigour for the overall reputation of the profession (Mabry, 2004, p. 386).

Another ethical theory is rooted in duties and obligations. For example, a duty to tell the truth may be revoked by a higher duty to do no harm, although the issue of how to prioritize is difficult in the absence of standards for doing so (Clark, 1997, p. 158). What constitutes harm may also vary in different contexts. A theory of duties and obligations will also not necessarily focus on the consequences. If deficiencies are encountered in the process or in a policy itself, the “right” thing to do may be to expose this inadequacy, irrespective of the consequences.

Evaluators who adopt rights theories to justify their actions believe that every person must be treated with dignity and respect and that there may be cases where rights to privacy take precedence over advancement of knowledge, programme objectives or the evaluator's theory of the programme. The ethical guidelines on privacy of the Canadian Social Sciences and Humanities Research Council, for example, advocate this precept in the statement “the question of invasion of privacy should be looked at from the point of view of those being studied rather than from that of the researcher” (1981, p. 5). In other words if there is ever any doubt that harm may be done to individuals through the
generation of knowledge, then respect for the individual should take precedence over search for knowledge.

Those who advocate social justice in evaluation take this rights-based philosophy to a group and society level. One of the main advocates of a social justice approach to evaluation is House (1980, 1993) who argues, following Rawls (1971), for a principle of fairness in evaluation in which all persons have an equal right to basic liberties. To realise such a principle it may be necessary in certain cases to redress inequalities by giving precedence, or at least equal weight, to the voice of the least advantaged groups in society who may not be represented, or powerfully enough represented, in primary stake holding groups in an evaluation. Evaluators working from this ethical base would also not take on an evaluation that violated a participant's rights to fairness and justice, where, for example, participants were involved in an experiment that suppressed their autonomy or which might render them worse off than before.

If we return to the first dilemma in the introduction, where participants in a programme evaluation were denied payment that affected their livelihood, while the evaluation report was being negotiated with the managers of the programme, a decision to aspire to maintain the integrity of the evaluation could be justified by an appeal to fairness and justice to those beyond the case and the evaluation profession. A decision to compromise might be justified by an appeal to relational ethics - a consideration of how it would affect the livelihood of those who remained unpaid. But these are not the only choices. A further choice - to bypass the institution and release the report to the overall sponsor (beyond the managers) could be justified by an appeal to utilitarian ethics - the participants get paid, all the stakeholders get the report, the evaluation maintains its methodological integrity and the wider audience gets to debate the findings.

In the last ten to fifteen years the ethics of care (Gilligan, 1982; Noddings, 1984) has expanded our choices of ethical theories to justify our actions. This emphasizes greater attention to the primacy of relationships in the programme (Christians, 2000, p. 143), rather than individual human rights and rules, and to specific contexts, not universal laws or principles (Newman & Brown, 1996, p. 33). Rational thinking is not rejected in this approach but rather must serve something higher - caring for others (Noddings, 1984). Into the frame of ethical thinking here is a concern for how an individual's personal life circumstances affects the decision and how he/she
would be affected by the available alternative choices. Often the preferred choice in feminist ethics (Denzin, 1997; Usher, P., 2000), the ethics of care also has much in common with participatory and democratic forms of evaluation (House & Howe, 1999; MacDonald, 1976) and with the ethical theory of feminist communitarism proposed by Denzin (1997).

Take, for example, the scenario where a local administration, ten months into the life of a project that is providing bilingual education to young children to help raise their self-esteem and adjust to schooling, realise that they have appointed the wrong director to the project. The administration wish to persuade the director to leave and, when this is resisted, seek to sack him. The project director seeks union advice. These relationships constitute data that partly explain the slow development of the project. The ethical dilemma the evaluator faces is whether to record the history of the project and its negative effects on the development and outcomes for the children concerned, or to protect the director's anxiety about his future career, as requested by him, by omitting certain details and transactions. An evaluator adopting an ethics of care position might opt for protecting the privacy and future career prospects of the director; one operating from an ethics of consequence might be more concerned about getting the programme on track so positive outcomes for the children would have a better chance of being realised.

“The Ultimate Ethical Act”

I have briefly outlined key differences in ethical theories with some illustrations to indicate how a theory or combination of theories may ultimately guide our ethical decision-making. However, as the first introductory quotation to this chapter suggests, actual ethical decision-making is balancing mutually conflicting principles in concrete socio-political situations. This is “the ultimate ethical act” to which House (1993, p. 168) refers. For further examples of ethical dilemmas and indications of how they might be resolved see the Ethical Challenges section in the American Journal of Evaluation and for examples of actual concrete ethical decisions in action see Lee-Treweek & Linkogle (2000) and Simons & Usher (2000).
In reaching a decision evaluators have a number of resources to draw upon to inform ethical thinking. Soltis (1999) suggests that the issue/situation be considered from three different perspectives: of the person (the evaluator in this case), the profession, and the public, noting the different dilemmas that occur for each. Newman & Brown (1996) offer a framework that includes intuition, rules and codes, principles, theory, personal value and beliefs and action. While this may appear overly rationalistic, given the uncertainty, complexity and finely tuned professional judgment we have to make in the “ethical moment” (Usher, R., 2000), it draws our attention to a range of sources that we may need to integrate to inform the ethical decisions we make.

Furthermore, as several authors have pointed out (House, 1993; Mabry, 2004; Simons, 2000; Thomas, 2000), we need to keep in mind that interpersonal and political factors also impact on decisions we take in the field. Many of the responses offered to the ethical challenges outlined by Morris (2000, 2001, 2004a, 2004b) are as much common-sense, relational and political as ethical, as Thomas (2000, p. 269) points out in his commentary on “The Case of the Sensitive Survey”. This brings our attention back to the broader political space within which evaluation is located, where politics is often intertwined with ethics and methodology and where findings often result in the allocation and reallocation of resources (Thomas, 2000, p. 269).

**Inter-relationship of Politics and Ethics**

The second quotation with which this chapter began highlighted this interdependence of politics and ethics. This section explores how these inter-relationships may be procedurally managed in an evaluation given the political values and interests of the different actors - commissioners, sponsors, participants, managers and evaluators - with specific reference to two democratic practices. During the past thirty years, as the field of evaluation acknowledged its inherent political dimension (Apple, 1974; House, 1972; MacDonald, 1976), various approaches to evaluation have developed which attempt to address the democratic obligation to conduct defensible, just and fair evaluation in political contexts where the values and interests of different stakeholders often conflict. These approaches include stakeholder evaluation (Greene, 1988), democratic evaluation (MacDonald, 1976), deliberative democratic evaluation (House & Howe,
and various forms of participatory evaluation (Cousins & Earl, 1992).

In their different ways these practices evolved to honour the different values and interests represented in projects and programmes and offered procedures to mediate conflicts of values and interests. They all acknowledge the role the evaluator plays in generating knowledge in society and that creating and maintaining relationships in an evaluation is central to the generation and dissemination of evaluation knowledge to maximize its utility.

This emphasis on politics and relationships in the process of evaluation subsequently came to be embedded in ethical codes and standards. Fitzpatrick (1999) points out that a comparison of ethical codes produced in the early 1980s with those in the early 1990s reveals a major move to focus on non-methodological issues. This is apparent both in the AEA guiding principles which “are more concerned with qualities or principles that permeate the evaluation process” and changes in evaluation training where there is “a greater examination of the many political factors and personal judgements entailed in conducting evaluations” (Fitzpatrick, 1999, p. 6). (See also Greene, 1994.)

The essence of what it means to evaluate ethically within a democratic approach may be demonstrated briefly through two approaches, the first from the UK, and the second from the US. Stemming from liberal democratic theory, democratic evaluation (MacDonald, 1976), aspires to find an appropriate balance in the conduct of evaluation between the rights of privacy of individuals and the public's right to know. In a particular case, this may mean that the public's right to know has to give way to an individual's right to retain some information in confidence and not allow public release. Alternatively, a person's “right to privacy” may need to be balanced in the public interest and, in certain instances, to be breached as, for example, where an illegal act is revealed in the course of the evaluation.

This central aspiration was translated into a set of ethical principles and procedures for the conduct of evaluation that accords equal treatment to people and ideas; that allows a flow of information which is independent of powerful, hierarchical interests; that does not allow any one person or group to veto what has already been cleared by others as public knowledge. These ethical procedures (for examples in practice see Simons,
1987, 1989; [p. 256 ↓] Bridges, 1989) are based upon principles of equality, justice, fairness, and respect for persons and for public knowledge. They are embodied in a contract or agreement that is endorsed by all parties to the evaluation at the outset and to which reference can be made in the event of disagreement.

In the US, House (1980) has long been an advocate of democratising evaluation and of justice in evaluation. He also draws from liberal democratic theory in advocating a concept of liberal democratic evaluation to aid in making choices about public programmes and policies (House, 1980, p. 141). Essential concepts in this approach are respect for the interests of all relevant stakeholders, equality of choice and the “public interest”. However, House accords more weight to distributive principles, like equality, justice and fairness, in judging the procedures by which decisions are made (ibid., p. 151), not the decisions themselves. Drawing on the work of Barry (1965) and Care (1978), who advanced the concept of “procedural moral acceptability,” House (1980) proposes that evaluators draw up a fair evaluation agreement to guide their evaluation practice based on twelve conditions under which policy agreements among persons are morally acceptable. (For further explication of these conditions see House, 1980, pp. 162–170.)

In more recent work, House & Howe (1999, p. 94) have advocated deliberative democratic evaluation, which has the three general criteria of inclusiveness - to represent all interest and values, dialogue with relevant groups - so the views are properly and authentically represented, and deliberation - to arrive at proper findings. While the authors do not specifically mention ethical procedures in their account of this approach, it is clear that respect and autonomy for persons, beneficence, openness and transparency of process are key ethical concerns.

**Ethical Conventions/Procedures in Evaluation Studies**

When it comes to the actual conduct of evaluation, evaluation shares with other forms of social research many of the usual conventions for the protection of participants, such as informed consent, privacy and confidentiality and pre-publication access. However,
evaluation has specific demands and dilemmas given its function and inherent political nature. These include, for example, how we arrive at evaluation questions; how we determine which values to promote (given legitimate concerns among stakeholders); how we determine which evaluation theory or principles to follow; how we design and implement evaluations, including choices of methodology; how we determine costs and benefits; and how we report to legitimate stakeholders in different cultural contexts. In this section I draw attention to the ways in which some of these issues are manifest in evaluation, starting with a comment on differences in methodology and then exploring how certain common ethical precepts may be reinterpreted in conducting and reporting evaluations.

Some Differences in Methodology

Several authors (see, e.g., Lincoln, 1990) have made the point that the ethical issues facing evaluators adopting qualitative methods are different from those encountered when using quantitative methods. However, they are not type-specific (Lincoln, 1990, p. 278). Differences relate to broader issues in ethical decision-making including ethical theories about the right to know, the right to privacy, respect for persons, honesty, justice, respect for knowledge and how it is generated and disseminated and to whom. But whichever methodology is chosen, it has consequences for participants and beneficiaries, directly or indirectly.

[p. 257 ↓ ]

There are certain ethical principles that apply whatever the methodology chosen - one should not lie to respondents, one should not deliberately deceive in order to gain access and one should not falsify results of an evaluation. However, the choice of methodology and the procedures adopted in using that methodology reflect a particular ethical and political stance which will affect how the evaluator resolves dilemmas such as the following: Is deception ever warranted? Are some interventions inherently unethical whichever methods are adopted? What counts as doing no harm? How do you protect the rights of children or cultural minorities affected by misuse of evaluation results? Do contractual obligations have ethical implications or simply legal responsibilities? Who decides who tells the truth?
It is sometimes inferred that ethical issues are not so present when using quantitative methods, as people are not so directly involved and it is easier to anonymize respondents and institutions. Conversely, it is often assumed that ethical dilemmas are faced more acutely when qualitative methods involving direct interpersonal interaction are adopted. Neither is the case. Ethical issues may be more hidden in the use of quantitative measures but they are not absent (see, e.g., Jones, 2000; Lincoln, 1990; Sammons, 1989).

However, there is little, if any reference to ethical issues in popular texts on statistics in the social sciences (Jones, 2000, p. 48). This is either an oversight, Jones suggests, or a significant omission given the harm that can ensue, for example, for students’ life chances from the bias that may be built into psychometric tests, or from the manipulation of statistical data. The manipulation of some of the international comparative data on mathematics teaching, for example, could be harming the educational chances of the very students it seeks to benefit. (Jones, 2000, p. 147). Randomized controlled design studies also raise serious ethical issues for groups who are not given treatments that are demonstrated to have positive effects. The question of ownership of data is also less easy to attribute. For example, who owns the test results of a test that is taken by a child in school hours - the child, the school (in locus parentis), or the local education authority/ district?

By contrast, reference is made to ethics in many qualitative inquiry texts, especially those which emphasize naturalistic, phenomenological, and narrative inquiry, where much of the data are gathered by direct observation and interviewing of people. One cannot conduct a personal narrative or study a person's life without seeking personal informed consent and clearance of interpretation or observe a child who is severely disabled without seeking parent's consent. The above differences in applied ethics and methodology refer to research in general. In the next section I re-examine common ethical procedures in the context of evaluation.
Reframing Ethical Procedures in Evaluation
Designing and Planning the Evaluation

Design. In discussing a cost-benefit approach to weighing the risks and benefits of an evaluation, Mark, Eysell & Campbell (1999) indicate that the debate on whether there is an ethical dimension to the quality of a design has a further complexity in evaluation, given the importance of the criterion of use. They outline a range of positions an evaluator could take, drawing on authors discussing psychological research, namely Pomerantz (1994) who argues that design quality and use are not ethical issues and Rosenthal (1994) who argues that bad evaluation design is actually unethical. In studies of low risk, Mark et al., (1999, p. 50) conclude that design issues “should not be viewed primarily from the lens of ethicality…. When risks for participants are high, a utilization focus (Patton, 1997) becomes an ethical mandate, necessary to justify the high risk” (ibid., p. 51). Once again we have a judgment to make in a specific case, in this instance, whether the quality of design is a technical or ethical issue.

Quality of design for Mark et al., (1999, p. 52) includes a whole range of questions related to whether evaluation methods are adequate to achieve the evaluation's purpose and whether the evidence provides a reasonable warrant for the claims and conclusions made. Take the example of a participatory evaluation design of a community project, which aims to provide services to multi-ethnic groups, and where people have chosen to participate in deciding what questions should be asked in order to inform community action. A trade-off might be that partial feedback is more useful to the community to engage their participation than a more comprehensive design of pre-post surveys and in-depth interviews of all potential recipients of the services.

Planning. In planning an evaluation the ethical issues proposed by Perloff & Perloff (1980) also help set ethical limits and formulate specific procedures for action. While the four these authors mention can all be said to stem from “respect for persons,” each focuses on a different aspect of this principle. The first is withholding the nature of evaluation research from participants or involving them without their knowledge.
The imperative here is clearly not to withhold information, though there are some who claim some practices would not be evaluated unless covert access was gained. How would one evaluate a corrupt police force, for instance, or a sect that was manipulating adolescents? The second is exposing participants to physical stress, anxiety [potentially the case in physical education or medical evaluations], or acts that would diminish their self-respect. The third is invading the privacy of participants, an issue that is embedded in relationships as much as in procedures (see later section on privacy) and the fourth, withholding benefits, is clearly not something evaluators should do. These are all aspects that need to be considered prior to seeking informed consent.

*Informed Consent.* In gaining informed consent there are four particular issues beyond those just mentioned that have ethical implications for evaluation. The first relates to the uncertain political nature of the enterprise. Gaining informed consent is an ethical imperative in any research study and signing an informed consent form is a familiar formal procedure in many contexts. Yet for many this is not consent at all, as it is never possible to know what will transpire as the programme unfolds in a precise socio-political context. Whatever forms are signed, “free and fully informed consent” needs to be realized through the additional different procedure of “rolling consent”- renegotiating consent with each person and/or site once a greater awareness of the context and structure of the study is known, or what Ramcharan & Cutcliffe (2001) have called process consent - renegotiating consent at different stages in the process of the evaluation. Such “rolling consent” is a particularly important procedure to maintain in evaluation studies that have emergent designs, stakeholder involvement, and a democratic or participatory intent.

The second informed consent issue concerns children. Morris (1999, p. 20) draws attention to the ethical dilemma that has been raised over active versus passive consent by parents, querying whether it is ethically appropriate to rely on passive consent from parents knowing that active consent generally produces lower participation rates (Morris, 1999, p. 20). A further issue, not often addressed, is active consent from children themselves. At what age and in what contexts can children be considered to give active consent? Even quite small children have ways of demonstrating their agreement to be involved in an evaluation study and to continue to participate.
The third concerns payment - whether it is ethically appropriate to offer payment as part of informed consent procedures. In some contexts, where respondents are poor, for example, or the cultural context necessitates, it may not be possible to evaluate the programme without offering payment. The AIATSIS (2000) principles, for instance, expressly state that researchers and evaluators should “Recognise that certain cultural information is owned and may need to be paid for” (Principle 7, point 5, p. 12). In other contexts payment may be seen as coercion to participate.

The fourth issue is trust. Where indigenous communities mistrust the language and intention of government and those evaluating government initiatives, guidelines on reciprocal benefits, collective rights of ownership, and recognition of cultural understanding need to be thoroughly negotiated at the outset, maintained throughout and renegotiated if necessary in the light of the findings.

Conducting the Evaluation and Safeguarding Privacy and Confidentiality

Mark et al., (1999, p. 48) succinctly summarize the usual guidelines on ethics in data collection and analysis as minimizing risk; only conducting evaluation when benefits outweigh the risks; obtaining fully informed consent; upholding principles of non-coercion, respect for participants’ privacy and decency, preservation of confidentiality to the fullest extent where data cannot be collected anonymously, avoidance of fraud; and following technical requirements. Yet they also recognize that “the devil is in the details” (ibid., p. 48). Decisions still need to be made in specific contexts as to when cash payments for completing questionnaires, for example, can cross into coercion (Mariner, 1990) or where particular trade-offs need to be made.

The traditional means of protecting individuals’ privacy and confidentiality in evaluation studies is to anonymize participants. This is relatively easy to guarantee in evaluations that use quantitative methods, more difficult in those which adopt qualitative methods and participatory approaches that have educative or emancipatory intent. Numerous procedures have developed to acknowledge this difficulty. These include offering participants control over what they say and how they are represented, seeking
clearance of comments in contexts when it is possible to identify individuals, even when anonymized, and engaging participants at every stage of the evaluation so that even if they are identified they have come to trust the process and accept the findings.

However, even if it is possible to anonymize, this does not mean that privacy is protected or confidentiality respected. Confidentiality is easily offered but difficult to assure, and Morris (1999, p. 19) reports that promising confidentiality when it cannot be guaranteed is one of the most serious ethical violations in evaluation. Respecting privacy or non-intrusiveness is partly methodological - what questions get asked and how as well as what is reported. However, important protection and respect for privacy also lie in the relationships established in the conduct of the evaluation.

Torres & Preskill (1999, p. 64), in considering ethical dimensions of stakeholder participation and evaluation, note that relationships (and the credibility, trust, rapport and mutual understanding they create) are central to successful evaluations but question whether the pragmatic agenda of evaluation will compromise its relational dimensions. In considering this issue they appeal to Schwandt's concept of “knowing through relationships ‘as a possible solution to the question Are we doing the right thing, and are we doing it well?’” (p. 11, italics in original, quoted by Torres & Preskill, 1999, p. 65).

Similarly, respecting confidentiality is more than mere anonymization. Anonymization is a technical procedure which may offer some protection for an individual from public gaze, but the ethical promise in confidentiality is met in how one negotiates the use of the data. It actually means in certain cases not reporting information that a person has given in confidence.

Furthermore, there are situations where anonymization is not the most appropriate procedure to employ. Clearly this is the case where public figures are concerned, where there is only one lead figure in an organization or where there is a need to acknowledge the participation and contribution to knowledge of those involved in the evaluation. It may also be important not to anonymize where the aspiration is to build a collaborative reflexivity about the process of evaluation; discussing moral dilemmas and how to resolve them collectively will generate further trust. Finally, there may be situations
where there is a need in the public interest to reveal dubious or corrupt practices or those which potentially may harm.

**Reporting the Evaluation Fairly**

One of the most frequently encountered problems in evaluation, which was also deemed the most serious by evaluators in a study by Morris & Cohn (1993), involved the presentation of findings. The most specific problem was pressure to slant the findings to suit the interest of a particular stakeholder, usually the main client.

The ethical procedures to ensure fairness and justice rooted in a democratic ethic help to avoid such partiality in reporting. Yet the politics of evaluation practice in a customer-contracting culture often mean that the dominant stakeholder threatens this ethical stance. This can be avoided to some extent by ongoing engagement of stakeholders in the process of the evaluation, involving them in identification of issues, checking with them whether their perspectives and actions have been interpreted and/or represented accurately and fairly and, perhaps most importantly, making interim results public. In this way knowledge is shared by a range of stakeholders in the evaluation throughout the process making it more difficult to modify or censor findings at the final report stage.

Further respect for participants in forms of participatory evaluation can be achieved by acknowledging them as co-authors of reports and, in certain cases, where it is unquestionably their story which is told, offering them a share of the profits if the report is sold. See McKeever (2000) for an example of this principle in action. The issue of profits for participants introduces an arguably problematic notion into evaluation studies, but in certain contexts it may facilitate the educational aim of evaluation and the development of community. In the case of indigenous communities, reporting will need to take account of the reciprocal obligations agreed at the outset, the different culture, language and ways of understanding and forms of reporting (which may include oral reporting) that acknowledge cultural difference. In early negotiations, it is important to establish what financial remuneration and intellectual property rights or portion of them will be accorded to the community (AIATSIS, 2000, principle 6, point 4, 5, principle 9, point 7, 10).
Summary and Conclusion

As the field of evaluation has developed over the past four decades, ethics has come to assume more visibility. Many of the alternative evaluation approaches that evolved feature people’s experiences, perspectives, and values, giving rise to the need to establish procedures to protect individuals’ privacy. Recognition of the inherent political nature of evaluation led to the emergence of principles, agreements, and contracts that aspire to treat stakeholders equally and fairly in the evaluation process and ensure that findings contribute to public discourse and debate. Evaluation has also intensified at the professional level with the rise of professional standards and at the institutional level in the setting up of ethical committees or Institutional Review Boards (IRBs). While the latter developments are often linked to discussion on ethics, it is noteworthy that these practices in some situations are more concerned with defending institutions and controlling what evaluation gets done and which methodologies are appropriate.

In such a changing context evaluators need to be vigilant on a number of fronts. First, we need to distinguish those practices that facilitate and those that inhibit ethical evaluation and challenge the latter. Controlling which methodologies receive funding, for example, is not beneficial for programme participants, evaluators or clients and is certainly not conducive to informed public debate or social justice. Sanctioning evaluation practice according to universal standards that are not sensitive to different cultural contexts and practices is similarly unlikely to lead to better ethical practice.

Secondly, it is clear that ethical committees and IRBs, while potentially useful to consider what requires ethical consideration in proposed evaluations, cannot control or determine that the evaluation subsequently conducted will be ethical. IRBs are only one element in a number of processes that are needed to ensure ethical practice (Mark et al., 1999). Some would argue, furthermore, that ethical practice can only be achieved through relationships in the field (Schwandt, 1998; Torres & Preskill, 1999). Too much attention to pre-procedures in such committees may be a distraction from the ethics of evaluation, as evaluators assume that procedures endorsed by these committees can simply be carried out as methodological dictate or moral rules.
Thirdly, it is clear that some changes to field procedures for ensuring ethical practice in the conduct and reporting of evaluation are required. Anonymization, for instance, needs to be decoupled from concerns with confidentiality and privacy. Not only are there methodologies where anonymization is inappropriate, it does not guarantee confidentiality or protection of privacy. Ensuring privacy in certain qualitative methodologies is an intimate, responsive, field decision linked to, for example, asking nonintrusive questions and making strategic political decisions that would not harm a person's career or position in an organization if information were released. Furthermore if democratic evaluation approaches gain greater credibility, it may not be possible and/or it may be counterproductive to anonymize a range of individuals with whom one has had extensive deliberations during the conduct of evaluations and the findings.

Fourthly, it will be important to document specific ethical dilemmas we encounter in our practice and how these were resolved so that increasingly we can demonstrate which principles and ethical theories influenced our decision-making in the "ethical moment" (Morris, 1999; Usher, 2000). In this way a strong evidential base for ethical decision-making in situ can be developed, which offers integrated examples of principles in practice.

Fifthly, many evaluators (see Fitzpatrick & Morris, 1999) stress two further points. One, that we still have a long way to go in finding the best ways to resolve ethical issues (Morris, 1999, p. 16); and two, that dialogue between evaluators and with evaluators, clients, sponsors and participants may be the most effective way to resolve them (Mabry, 2004; Torres & Preskill, 1999). Similarly, Mark et al., (1999, p. 54) conclude that “it is perhaps better not to dance with the devil alone” in advocating engaging in many forms of deliberation (e.g., review panels, advisory panels, evaluator networks and discussions) to advance ethical practice.

Finally, we need to remember that whatever procedures are set in place, it is only in the field - the political ethical space (Kushner, 2000, pp. 151–152) - in which we negotiate relationships and meanings in full cognizance of all the relevant factors in a specific sociopolitical context, and the ethical theories and principles which have informed our actions, that we can know if we have acted ethically in an evaluation. This is the complex integrity Glen (2000) speaks of - the constant reflexive re-examination of values (ours and others) in relationship with those we evaluate.
Acknowledgements

I am grateful to the editors of the Handbook and to David Bridges, Ernest House, Robert Picciotto, and Patricia Rogers for their critical, generous, and helpful comments.

Helen Simons

http://dx.doi.org/10.4135/9781848608078.n11