

Beyond Evaluation Standards?

Petri Virtanen and Ilpo Laitinen

Contact details of the authors:

Petri Virtanen is Senior Partner at Net Effect Ltd., a Helsinki-based consultancy specializing in policy, programme, project and institutional evaluations.

Address: Mikonkatu 13 a 34, FIN-00100 Helsinki, Finland. e-mail: petri.virtanen@neteffect.fi.

Ilpo Laitinen is a division director at Centralised Services of Youth Department, City of Helsinki.

Address: Youth Department, City of Helsinki, Box 5000, City of Helsinki, Finland. E-mail: ilpo.laitinen@nk.hel.fi

Abstract: It has now become a truism to suggest that evaluation is a highly respected, appreciated and venerated enterprise. This article is based on three central claims. First, evaluation standards and ethical principles are useful only to the extent that one recognizes what they can and cannot do. Secondly, they can never be applied in algorithmic fashion, but must always be interpreted in the evaluation 'case' at hand. And thirdly, they are, at least to some extent, shaped by cultural norms and understandings. It appears, as this article concludes, that morally correct action does not become certified on the basis of an order or a norm, because even one counter-example is enough to conclude that dependency between a morally correct action and a norm is not logically valid. Morality should also express an individual's own freedom and the motives of action related to it. Standards do not have any causal consequences as such.

Key words:

Evaluation ethics, standards, guidelines, morally correct evaluation practice.

Introduction

For most people in the evaluation community today the issue of evaluation standards is no longer 'hot' news. It is however the current authors' opinion that the need remains for a critical analysis of current evaluation standards. What then are we talking about when discussing the ethics of evaluation or outlines for evaluation ethics? Who deals with these questions? Whose ethical standards apply? Do these matters concern the academic community, the professional evaluators or do they deal with wider questions of general societal justice. In this article it is argued that questions regarding evaluation ethics simultaneously concern wider societal issues, and academic questions whilst also raising important practical issues for the evaluation profession in particular. As such, the question of 'evaluation ethics' needs to be considered and addressed from a number of different contexts.

The question of evaluation standards relates to the question of evaluation culture and to the norms or ethics of the evaluation profession as a whole (Furubo & Sandahl, 2001: 6-12). In an international comparison, national evaluation societies have played a dominant role not only in making up the evaluation fabric but also in formulating the evaluation standards, ethics and norms. As Derlien and Rist (2001: 452-453) point out, a shift has already taken place, more accurately we can see a decentralization effect, from the evaluation actors at the central government level to the lower levels of government and other public sector domains. Evaluators at the regional and local government level have also emerged in large numbers and they have become new players in the evaluation field.

Today, there are number of evaluation standards available. A whole series of standards and meta-evaluation criteria indeed were developed within the evaluation communities in various countries, national and trans-national organisations, and evaluation associations throughout the 1990s in particular. Probably the best-known set of standards is the one adopted by the American Evaluation Association in 1994 known as the AEA Programme Evaluation Standards (on the background to standards, see e.g. Shadish et al. 1995; Virtanen, 2004). Since then, a number of evaluation standards and ethical codes have been developed, which can mainly be derived from the AEA Programme Evaluation Standards. This is the case for the Canadian Evaluation Society, the African Evaluation Society, the Swiss Evaluation Society, and the German Evaluation Society. AEA Programme Evaluation Standards have been somewhat modified in these cases, though the basic structure and formulations of these case examples are identical to the AEA model. Undoubtedly, as national and trans-national evaluation communities and associations seem ever more keen to draw up set of evaluation standards, that the question of evaluation ethics is becoming increasingly recognised as being of some importance. The question is, then, what is it that the authors assume they are achieving in setting up these standards, and to whom are they addressed? Are they used to validate the quality of evaluations, for training purposes, or are they used for something else? It has been argued (Virtanen, 2004), that the current moral landscape is so fuzzy and uncertain that the applicability of such standards is probably far from self-evident and straightforward.

Perhaps due to the fact that there has been growing interest in the topic – ethical principles, standards, guidelines, and meta-evaluation criteria – the perception of evaluation practice has however changed somewhat in the past few years. In polemical

terms, it is now almost a truism to suggest that evaluation has become a highly respected, appreciated and venerated enterprise.

Today, these issues are immanent in evaluation conferences, as a large number of sessions, workshops, and roundtable discussions are normally organized around such topics. For instance, the 5th conference organised by the European Evaluation Society (EES) (10-12 October 2002 in Seville, Spain) had a roundtable entitled, '*Do we need European evaluation standards?*' and this roundtable attracted over 50 evaluators and commissioners of evaluations. Moreover, the theme of evaluation standards will also appear on the agenda of the next EES conference in Berlin in the autumn of 2004.

Keeping in mind the fact that both the European Evaluation Society and many national evaluation societies across Europe and elsewhere are developing or discussing their ethical codes, we aim to situate this on-going development activity within the broader discussion of evaluation ethics. Our point here is that evaluators (trend-conscious or otherwise) should have a clearer sense of the moral and political meaning and implications of evaluation practice and they should be able to recognize the limitations of its current conceptualisations. In our view, the current debate on evaluation standards has been overly focused on the 'technical' aspects of the standards. As such, these types of evaluation standards may be seen as necessary, but not sufficient grounds for any proper evaluation exercise. Everyone should thus act to ensure that all evaluators show a similar level of responsibility. This however may prove to be a rather two-edged sword in terms of 'Ethical Evaluation'.

This article is based on three central claims: the limits of evaluation standards and ethical principles, the need to apply these principles in ways that take into consideration the specific circumstances of the evaluation at hand, and the need to see these standards as necessarily being shaped by cultural norms and understandings. Although the first two claims seem initially to be verging on truisms, they still need to be clearly stated, as we try to attend carefully to different ways of ethical reasoning in evaluation practice. Concerning the third claim, we are aware that it would be of interest to develop an in-depth argument on how cultural norms and understandings are reflected in different codes of professional conduct. Such a task would however require a significant amount of research and is thus not possible within the context of the limits of this article. The article is organised as follows: we start by discussing the ethical roots of evaluation standards trying to shed light on their problematic construction. Thereafter we turn to the trivial side of the evaluation standards by pointing out that problems in standard setting are numerous and obvious. Finally, we offer some tentative conclusions on where to look, as it appears that current evaluation standards have serious limits in the terms of their applicability.

We have identified five objections to the standards of evaluation. They are mentioned as follows as our core arguments (which we aim to defend in this text):

- The justification problem: this relates to the lack of universal acceptability and to every possible situation covering a set of principles.
- The testing problem: if codes cannot guarantee changes in action, the empirical evidence is very limited.
- The applicability problem: there are no consequences for the evaluator if standards are not followed.

- The interpretation problem: standards tend to be treated as ‘add-ons’, as constraints upon ‘the real evaluation’ thus acting at the margins or as an ‘all the best take care of even this’ – type of action.
- The personal accountability problem: standards that specify behaviour in particular situations seek to take judgment out of ethically charged situations and diminish the evaluator’s own responsibility.

The ethical roots of evaluation standards

At the outset it is important to distinguish between ethics and morals. In common speech they are often used interchangeably, but in philosophy they have different connotations. The study of ethics is the study of theories of right and wrong conduct, while the study of morals focus on practice. Throughout the last century, the work of western philosophers has focused more on problems of theory construction than on problems of application. Today, when there is relatively little consensus on moral issues, the problems of moral philosophy seems not to concern the need to affirming moral beliefs that are already known or assumed to be true but, rather, on finding a method for determining *which* beliefs are true. Nevertheless, many critics feel that a serious gap remains between theoretical and applied ethics. As such, when it comes to professional ethics, it is important to investigate ethical theory and moral practice simultaneously. And even then it is a mistake to think of professional fields as being any more concerned with ‘applied’ ethics than ordinary moral situations. Professional areas require knowledge of specialized conditions, but the relation to the moral view is the same.

Next we can ask the question: are the evaluator’s role and the good of society irreconcilable in a discussion on evaluation ethics? How do the values of evaluation bridge the gap between individual and societal ethics? In some instances such a dialogue can be bypassed by pointing out that morals are always, in the final instance, up to the individual. In other cases, the definition of values and norms (de-ontological side) is believed to enable the entire problem to be resolved in a broader context (teleological side). Our claim here is that different moral theories offer different justifications for similar ethical principles (MacIntyre, 1987; Norman, 1991). This means that moral theories must be identified, and must be validated as to how they can offer different justifications for a common set of moral principles (for instance, honesty, beneficence, trust, justice, and so forth). In all, there are a number of historical and ideological roots as to the perceived necessity of establishing universal criteria and ethical standards, and these will be discussed in more detail in what follows.

Most national and transnational evaluation societies have defined their values during the last ten years or so. For some, societal goals have been stated, while others have focused on defining the contexts and means of rational action and professional goals. The presumption here being that rationalism brings with it the generalizing means to accomplish such goals. In other words, distinguishing sensible and wise actions from the achievement of such goals is in many cases impossible. At the same time, the chosen study method may limit the evaluator’s ability to act for the good of society. Within the context of any individual evaluator, can their role as an evaluator ethically transcend that of their general role as an individual in a society?

The goals, i.e. values, of action are not identical to the basic principles. The basic principles however are controlling values. Operative rules and procedural directions

are again clearly contextual. In some cases, the definition processes of evaluation values lead directly to the definition of operational rules. This is however not as unproblematic as it first may seem. Ethical standards reflect a world where formal demands presumed by principles are carried out. These demands include, among others, generality and generalization, recognition, finality and the demand for coherence. The demand for coherence in connection with ethical standards therefore means, for example, fairness set as a general standard adjusting to each individual situation. If that standard is not attained, for example because the evaluator did not have all the information needed, has a mistake then been made in applying the standard, leading to some kind of value sanctioning of the evaluator? Or if the general standard is attained without the evaluator having knowledge of the standard in question then, at least in theory, it can be claimed that this general standard of fairness is not a general standard at all. For example, one ethical standard of evaluation can morally *prima facie* obligate the evaluator to destroy the data collected in the process of the evaluation, but according to the law of a certain country or the regulations of a certain organization, that data must be signed over to the subscriber if so demanded.

Again, one can continue by outlining additional challenges drawn from the same example. What if an evaluator acquires an exceptional permit to release this data from the international organization of ethical evaluation? Is s/he now morally autonomous on the basis of an exception permit against ethical standards? Is his/her most important ethical choice then to be obedience? Morally correct action does not become certified on the basis of an order or a norm, because even one counter-example is enough to say that dependency between a morally correct action and a norm is not of logical quality. Morality should also express the individual's own freedom and the motives of action related to it. Standards and their rules do thus not have any causal consequences as such.

The threat of trivial ethics

Because it is impossible to include all of the potentially different individual and communal circumstances into a set of standards, such standards necessarily remain imperfect. To take an example, we can make the following classification inspired by Ilkka Niiniluoto (2003: 176-183). According to this classification, simple instructions can be absolute in form:

1. Do *X*!

Or they can be conditional in form:

2. In case of *B* do *X*!

These instructions may be based on practical experience, they may be based on the prominence of the giver, or they can be of a specific professional character. When moving from experience or prominence towards the scientific testing of these rules, one ends up with instructions demonstrating causal connections in form:

3. Doing *X* in situation *B* produces effect *A*.

In research, technical norms are usually based on the pursuit of truth. Then the norm should be in the form:

4. If you want *A* to happen and you believe yourself to be in situation *B*, then you should do *X*.

If standards are simply taken as rules to be met and the commitment is taken as given, the standard in question diminishes in nature and becomes an instruction as described above in point 2.

Hubert L. Dreyfus had a thesis according to which following rules was the lowest level of a professional hierarchy. He argued that with long-term experience and 'non-calculatory' skills, situations are dealt with as entities without a need to refer to rules. Dreyfus is also critical of the field of ethics. According to him, ethics entail the capacity to make moral choices and therefore the need to learn rules and to follow them belongs to 'the novice'. (Niiniluoto 2003: 182-183.)

The ideal of professional ethics is that there are rational principles or even norms guiding practice. Is an evaluator obliged to follow these norms in such a way that differs from his own everyday morals or the general morals of the surrounding society? From the perspective of applied ethics this questions is much wider than merely that of addressing professional principles.

Standards form a certain type of norm structure, but their normative usage is measured by investigating to what extent they can pursue certain actions or discipline behaviour that is perceived to deviate from this norm. The breaking of such norms is then simply an expression of their weakness as binding codes. The weak validity of such norms does not however discredit them as long as a monitoring system is in place, or as long as breaking them ultimately leads to punishment. Standards may solve some, perhaps even the majority, of ethical problems – such as not knowing what is 'good' and what is 'not good'. They cannot however resolve all ethical issues. They do not reach all ethical questions, such as the question of the social responsibility of different actors. Niiniluoto sums this up by arguing that, the advice related to expertise can never be totally mechanical, or be written in norms taking into account all contextual issues. Rules are still needed to identify experts from the perspectives of usability and ethics. (Niiniluoto, 2003: 183.)

Professional ethical oaths emphasize the commitment and motivation of a certain group of people. The main focus then is not only on the power of norms *per se*, but rather on the fact that what is underlined is a particular (ethically accepted) way of doing research. As a consequence of this, problems in standard setting are inevitable and thus numerous. Standard setting does not separate situations and circumstances, but sees them as actions without context, as the free universality of every action. Similarly, such an approach presumes every individual to be basically the same, to understand directions in the same manner, to act correspondingly, and so forth. Unavoidable preconceptions are connected to this interpretation of the morally acting individual. These preconceptions view the individual as a reasonable and rational choice-maker. In addition, it is assumed that s/he is able to make a choice based on values in each given situation, that s/he is a person who has his/her own moral autonomy, with other interests and desires not governing his/her decisions. Other prerequisites in setting standards include the similarity of individual rationality across the board, and of the individual's imaginal capacity when outlining the results of moral decision-making.

The point of view in question here is thus characterized by the idea of the 'objective self'. For that, the inner rationality of the self is typical, the historical self, the universality of self and the division of self into reason and desire. The foundation of this is the objective, absolute moral code as the ethical standard of evaluation modelled. This is precisely the problem faced in evaluation value work, the difficulty, attached to the deductive structure, of applying an acceptable instruction from the general norm to the 'individual, specific situation at hand'. The motive of moral philosophy is the foundation of an individual's action, how ethics are demonstrated in an individual's moral action. The validity of the action should also be mirrored with those total endeavours whose singular manifestations individual actions encompass. An evaluator working according to standards is aware of his/her own actual motives and has a moral perspective on them. Similarly, s/he then has some standard or criteria of what is correct.

This contains the risk of equating the motive with the criteria of evaluation ethics. Then for the justification of a motive it is sufficient to carry out a standard according to which the act is necessarily correct. Thus the standard defines validity, by stating what is correct. Therefore it is assumed that the action of the evaluator is morally correct if, and only if, it is based on moral motives, for which a criterion of validity is simultaneously set. However, in order to act in an ethically correct manner, one must first have an idea of what is valid before one can act according to one's moral motives. That point of view does not automatically follow the standard or the moral motive. Just as with any other norm, the standards of evaluation can be seen as universal, though this still does not say anything about coherence, universality and commitment. On the other hand, supporting 'loose' ethical principles, on which the majority of the evaluation community could agree, lead only to trivial ethics. Even with standards, there are many problems. Universal ethics are basically inapplicable. Ethical standards can generally be explained in terms of rule ethics. Morality is conceived purely as a guide to action, and concerned with what is right (in terms of action) rather than with what is 'good'. Standards promoting morals concern what we ought to do, even though we may not be obliged to do so. It can thus be seen that evaluation ethics are rather doctrinal in nature and indicative more of statements about the world ('is' questions) rather than actual normative solutions ('ought' questions).

The word 'ought', for example, can be used in a purely prudential sense. No moral obligation is implied in such a situation: simply that it would be in the evaluator's interest to do as advised. With some ought-sentences it is not easy to be sure whether the prudential or the moral consideration is uppermost.

The central question of evaluation ethics concerns the total theory from which the evaluation ethic is examined and its possible norms are derived. In order for any moral standards to be effective, each individual must commit to them. Morality is thus not a regulatory system of behaviour in which some intermediate norm or instrumental method is binding. What morality seeks to bind together is larger than the mere performance principles promoting the common good of the community in question (in this case the evaluation community). The question of the evaluator's ethics attached to the role of an evaluator thus emerges. The good intentions alone of an evaluator do not in themselves guarantee the achievement of a good outcome. The technical 'goodness' associated with an evaluator's role is in fact not necessarily associated with ethics. Indeed, von Wright, (1963: 76) asserts that '...when an essential bond

exists between an action and a goal, then technical good is measured at the end with the terms of instrumental good (...) It is relevant to say that technical good here is secondary in comparison to the instrumental good, because the former depends logically on the latter'.

Thus, technical good is represented by the ability to act in a certain way that is connected with the 'membership of the species' (i.e. of being an evaluator). What good then does the technical good serve? What kind of value choices does the evaluator make when settling into his/her role? Do standards exist that would affect the evaluator only when s/he is 'on duty', but not when s/he is 'off-duty'? Thus serious problems arise if the *persona* and the role are separated from each other.

Beyond standards – where to look, and where to go?

Theory and practice should be brought closer together, not to transfer power from a theory separately established and designed to dictate practice to a practice that determines theory or chooses perspectives to suit its convenience. The prominence of practical decision is not itself a theoretical answer. Professional ethics and its questions cannot be adequately answered without returning to the nature of ethical theory itself, and the roles that different theories give to the importance of character development in acting rightly.

As far as we conceive it, thinking in terms of moral choice is both essential and unavoidable: where no choice exists, no moral judgment can be made, and where it does exist, it cannot be escaped from. The basic distinction between ethical principles and moral rules is that the former are concerned with the overriding aims of human behaviour, the latter with the application of these aims in day-to-day situations. It is no easy matter to simply state ethical principles however, since as principles they are generally unchallengeable. The problem with principles lies not in reaching agreement with regard to their validity, but in establishing some kind of hierarchy or 'pecking order' when they clash, as they occasionally do. On the whole, agreement on a principle by no means guarantees agreement on action.

Having considered the differences between facts and values, between what *is* demonstrably the case and what *ought to be* the case we will now take this explanatory approach a little further. These two types of statements are made as the result of observation and the use of reason; they describe matters of fact. Normative statements are concerned with rules (in ethics) and the recommendations or proposals that follow on from these. Indeed, it is easy to make critical comments without being constructive. What, then, would be our suggestion of what comes after the standards? We will try to further elaborate upon this in what follows.

The model described in this paper is based on the difference between internal and external instructions divided into components of evaluation. Internal norms are the ones regarding the competence of a certain professional community. This then refers to the evaluator as a part of an evaluation community, a community of evaluators. External norms refer to the connection to various stakeholders in society, commissioners of evaluations, users of evaluation knowledge, the general public, in fact, society as a whole. We have already addressed the issue of changing ethics and changing norms, but we should note that even when change is natural, it seldom happens coinciden-

tally. Norms are justified and legitimised according to values and by extension attached to values and thus are always merely guidelines within a larger context.

Many research norms can be justified by having the truth as their main objective. The notions of academic freedom and inter-subjective testing are the means by which to attach the method to the search for objective truth. This goal requires freedom of action. Fairness has traditionally also been one of the most important academic norms, the third being the avoidance of causing harm. These norms however have often remained on a rather superficial level or have been predominantly confined to the realm of discourse. In addition, there have not always been robust ways of confronting these norms. Such norms are seen in the context of the ethical framework of the Finnish Evaluation Society where the division between internal and external values is made evident and visible and the different components of the evaluation process are described. This means for instance that the value of truth cannot exceed that of fairness in treating the objects of evaluation.

Thus in our view the aim should be that the ethical codes of evaluation is based on the critical aspiration of *truth*, collective *fairness* and responsibility in collecting and analyzing information.

When defining the values and ethics of evaluation one should move from the instrumental use of values into an equilibrium, in which the human, informative and context bound dimension of ethics is connected to its principle and value based dimension.¹ For an evaluator there are no specific or independent evaluation ethics to follow, but evaluation is always steered by the general ethos of a society. This includes the dimensions of the political field, and the state's administrative culture and values as a whole. The ethics of evaluation are not separate from those of its surrounding community, but are rather a part of the societal context where an evaluation is exercised. An evaluation process is never carried out in 'closed' conditions; an evaluator as an individual, and society as a whole, are not ethical counterpoints but instead are intimately connected to one another.

In our view then, the ethical code of conduct for the evaluation community should be constructed so that we can separate out four basic dimensions: *The Evaluator*, *The Evaluation Process*, *The Object of Evaluation*, and *The Evaluation Community*. This kind of framework indicates that the evaluator and the community are interactive and fundamentally engaged with one another. The framework highlights the various aggregates of an evaluation process. The values of evaluation are based on these aggregates. Therefore the core of evaluation ethics is an assumption that the good of another is of equal importance to the good of oneself. (Wright, 1963: 195-205.) The responsibilities of the evaluator, object and the process are also always seen as responsibilities towards a larger community. The community authenticates the different fields of a value framework. Positive values on which all evaluation activities are based can thus be categorized under four headings: values that are good for the

¹ The following ideas have not been developed in an intellectual vacuum. The authors of this article have been involved in the developing process of the ethical guidelines for the Finnish evaluation community and Finnish Evaluation Society (FES) in particular during the period 2001-2003. Please note that the views expressed in this article are solely those of the authors of this article and do not necessarily reflect the views of the FES. We are grateful of the comments received in connection with our ideas during this period of time.

evaluator, values that are good for the object of evaluation, values that are good for the evaluation process and values that are good for society in both the short and long-term perspectives. This framework is, above all, a description of the ethical perspectives for which an evaluator is morally responsible. How this moral responsibility is carried out is however always an evaluator's individual choice.

(i) Evaluator – Truth

According to the idea of the framework, good evaluation practice refers not only to value-based evaluation practices but also to the way of perceiving evaluators' rights and responsibilities. An evaluator must have free access to information and the freedom to seek the truth. Truth is therefore the ultimate arbiter of his actions.

(ii) Object of evaluation – Justness

The fair treatment of an evaluation object means taking into account their rights and treating them in a righteous manner. This may also be illustrated by reference to the concept of reciprocity; with equal rights, equal responsibilities, freedom and well being as the ultimate goals of all action. From this perspective an evaluator might try to place himself in the object's position – 'interrogating himself' – and thus investigating his own values and motives from that angle.

(iii) Evaluation process – Ability

The evaluator is expected to rely on valid evaluation methods and procedures, this being the core of an evaluator's professional ability. The evaluator embodies the responsibilities of a researcher, while being subject to the community that legitimises his/her right to conduct evaluations and set evaluation practices. By giving authorisation, the community also sets up the accepted norms of evaluation. Therefore ethical evaluation means more than just the proper use of evaluation methods. Discussing the values of evaluation methods one refers to the norms implicit in evaluation practice – in other words to the validity of procedures. The external norms of evaluation again are factors that connect the results of the evaluation process to the larger societal context. As such, evaluation is also always a product of co-operation and is thus attached to the surrounding community, at least indirectly. The premise here being that integrity and fairness are realised in the evaluation process and that the process provides socially relevant information.

(iv) Community – Responsibility

Finally it is about how all this is interpreted in a society which enables the evaluation practices and holds responsibility over the results and entitlement of the actions, from the perspective of both an evaluator and of the client commissioning the project. Even when all the aspects of a framework are of equal importance, the starting point should be the community. The collegial evaluation community authorises the evaluation. In addition, the evaluator, the object, and the commissioner of an evaluation are always part of their surrounding community, and thus are neither independent nor self-sufficient.

To sum up, the ethics behind the framework described above do not give concrete indications of what is right and what is wrong in any single case. Rather they supply us with methods and measurements for concluding what is right and good. If the values alluded to here are seen as being in the mainstream of western moral philosophy that is because these values are more than just parcels of words or descriptions. Indeed,

such value systems can actually be seen as ambitions or sets of aims that are valid in practically all circumstances. Therefore it can be said that a person can be equated with the values s/he abides by. Interpreted in this way, values cannot exist as lists of words, but instead only as one's personal choices and as frameworks for action according to the choices made.

Lessons learned: Evaluation ethics, the post-modern version?

We find it is easy to agree with Shadish and others (1991: 47-48) who acknowledge that the distinction between descriptive, prescriptive and meta-theoretical approaches to values is inadequate (see also Shadish et al., 1995). The differences between the above-mentioned value approaches however tells us little about the specifics by which evaluators actually answer questions about values in evaluation, and even less about the ways in which they actually implement their research. Current evaluation standards seem then to have relatively little to offer this broader discussion of values and ethics.

It would be tempting to believe what Shadish and his colleagues (1991: 49) say about the ethical competence of the modern evaluator. According to them, no evaluation theorist engages in descriptive, prescriptive and meta-theorizing activities explicitly and systematically. Evaluation standards as they currently exist, actually express very little with regard to values, and even where they do, the content of these values remain obscure. This means that standards remain as lists of proposed good practice. Related to the previous point, it would be naïve to assume that ethical codes as such could exist in a way that everybody conceives or interprets them in the same manner.

Finally, we must also consider the current nature of post-modern morality and ethics in order to understand the limits of the applicability of the evaluation standards. We should perhaps ask how evaluation standards 'fit in' with in post-modern ethics, and how do they enhance evaluation culture and the dissemination of evaluation information? If Zygmunt Bauman (1995; 1997) is right, no logically coherent ethical code can 'fit' the essentially ambivalent condition of post-modernity. Moral phenomena are today inherently 'non-rational' in the sense that they are not regular, repetitive, monotonous and predictable in a way that would allow them to be represented as rule-guided. This kind of reasoning does not leave much room for any codes of ethics in evaluation practice. Bauman argues that morality and ethics are no longer universal. The question thus emerges, how then do we attempt to control and guide something that is inherently neither controllable nor guidable?

This does not mean that morality is absent in evaluation practice, on the contrary. Values are there, as we are moral beings, but ethical codes seem irrational. In today's world we implement evaluation studies within the context of morality without a pre-determined or generally agreed ethical code. It would be comforting to think that most of the people, evaluators amongst them, most of the time, can do very well without a code certifying its propriety. Perhaps people need these kinds of codes and the authorization they provide so seldom that they hardly ever discover their absence. Most people – evaluators included – tend to follow the habitual and the routine; we behave today the way we behaved yesterday and as the people around us go on behaving.

Taking into account all that has been said above, we might conclude, perhaps rather surprisingly, that the best way to appraise various evaluation standards is also the

most straightforward: use them, but be critical, and apply them taking into account the cultural context within which they have been developed. Bauman (1997: 200) warns us: ‘history is fraught with mass murders committed in the name of one and only truth,’ this is perhaps a useful quotation to remember in terms of the application of different evaluation standards.

Evaluation standards cannot provide ‘miracle’ solutions. By following the guidelines set out in various standards we cannot be sure that the evaluation at hand is automatically of good quality and ethically of acceptable standards. Though being quite critical of the current interpretation of evaluation standards, we very much acknowledge the importance of ethics in evaluation practice. The point is, as Schwandt (1997) puts it in the title of his article some years ago: the landscape of values in evaluation is still to a major extent ‘unexplored territory’.

Morals cannot be ordained as laws. Therefore immoral actions cannot be sanctioned in the same manner as illegal actions. Moral action is thus in principle voluntary. The ethics of evaluation should be rewarding and attractive (‘you should strive for the best possible performance) rather than intimidating, imposing and enforcing. Evaluation societies could produce morals by being transparently ethical and promoting the definition process of evaluation ethics. This means that the ethics of evaluation would have a process-nature and could not be reduced to simple and pure outcomes or indicatives. The ethics of evaluation are effective only when being a process. Instead of plain common standards of evaluation different ways should be available to process the evaluator’s inward sense of public reason, the morals of society and moral conduct of evaluation communities.

One needs to be able to trust the evaluator as well as the results presented by him/her. In addition, there also needs to be a shared conviction that in conducting an evaluation process the rights of the surrounding community are respected. Evaluation standards (considered as abstract ‘how-to-do-it-cookbooks’) do not guarantee this, since they actually are not necessarily based on values, and do not guide evaluators towards carrying out the evaluation process in relation to high ethical standards, rather, such handbooks only provide information on technical knowledge (i.e. how to implement certain evaluation procedures according to certain ideas and practical hints). Therefore, today’s evaluation standards can be considered superficial: they do not give us many useful ideas in respect of how to act in accordance with certain ethical principles. When defining the values and ethics of evaluation one should move from the instrumental use of values into realms wider than those dealing with practical means alone. The ethical questions concerning the whole evaluation process – the nature and the management of the process, the use of evaluation data etc – make the ethics of evaluation a challenge at every step of a process, from defining the focus of the evaluation to the drawing of conclusions.

Where then does this lead us? Perhaps to a discussion of the role of the individual values of a single evaluator in the context of the collective values of the evaluation profession? On the other hand, if the creative class (modern evaluators included) favour individuality, meritocracy, openness and diversity, at least to certain degree it is an individuality, meritocracy, openness and diversity *of elites*, limited to highly educated and well-off people in research institutions, ministries, universities and consul-

tancies. To conclude, the lessons learned can be summarized as follows.

The underlying principles of evaluation standards cannot simply be derived from the values of science. Whilst there is a multi-faceted path from the following of evaluation standards, information gathering, making analyses, and deriving findings, to making conclusions and recommendations, an evaluation needs to follow scientific reasoning but cannot be guided by it alone (on this discussion, see for example, Valovirta, 2002).

According to the view presented here, the practical experience gained from using evaluation standards remains somewhat ambivalent, mainly suggesting that they can contribute to the spreading of knowledge about professional conduct in the field of evaluation. In some cases this knowledge has been used for educational purposes in training and as a benchmark for quality in carrying out evaluation studies.

Standards have also fostered a common language between evaluators and the people commissioning evaluations. All of these examples of using evaluation standards are quite positive. However, nobody knows precisely how these standards and guidelines have been utilised. It seems that in Europe at least, the use of these standards and guidelines is limited to the members of national evaluation societies. There are a number of difficulties involved in developing evaluation standards, as the ethical background of current evaluation standards seems rather vague. At the best of times standards need to be specified to be applicable and in order to be effective, they also need to be based on discussion and consensus-building. To understand the nature of evaluation standards, guidelines, or ethics, one has to resort to philosophy and logic. We should however make a clear distinction here between definitive rules (evaluation standards, if you like) and the rules of strategy (actual decisions made during an evaluation process). The rules of definition indicate what is allowed, what is possible, and what is not possible. The problem with the rules of definition is that they do not make explicit *which* decisions are good, bad or better than others. Rules of definition do not then necessarily guarantee anything. Though, it is necessary to become acquainted with these rules of definition in order to understand the logic of the rules of evaluation strategy. Evaluation standards and evaluation criteria do not provide 'miracle' solutions, and it would be naive to assume that they do. By following the guidelines set out in various standards we cannot be sure that an evaluation is automatically of a good quality and that it is ethically 'good' or 'acceptable'.

To conclude, it seems that the doorway to the promised land of an accepted ethics of evaluation remains very much ajar. This situation challenges us to come up with better theoretical analyses of evaluation ethics, and to formulate new theories as well as to indulge in the critical analysis of current practical applications.

References

Bauman, Z. (1995) *Postmodern Ethics*. Oxford: Basil Blackwell.

Bauman, Z. (1997) *Postmodernity and its Discontents*. Cambridge: Polity Press.

Derlien, H.-U., Rist, R.C. (2001) 'Conclusion: Policy Evaluation in International Comparison.' Pp. 439-455 in J.-E. Furubo et al. (eds.) *International Atlas of Evaluation*. New Brunswick & London: Transaction Publishers.

Furubo, J.-E., Sandahl, R. (2001) 'Introduction: A Diffusion Perspective on Global Developments in Evaluation.' Pp. 1-23 in J.-E. Furubo et al. (eds.) *International Atlas of Evaluation*. New Brunswick & London: Transaction Publishers.

MacIntyre, A. (1987) *A Short History of Ethics*. London & New York: Routledge and Kegan Paul.

Niiniluoto, I. (2003) *Totuuden rakastaminen* ('Loving the Truth', available only in Finnish), Helsinki: Otava.

Norman, R. (1991) *The Moral Philosophers. An Introduction to Ethics*. Oxford: Clarendon Press.

Shadish, W.R. Jr., Cook, T.D., Leviton, L.C. (1991) *Foundations of Program Evaluation*. Newbury Park: Sage.

Shadish, W.R., Newman, D.L., Scheirer, M.A., & Wye, C. (Eds.) (1995) *Guiding Principles for Evaluators*. San Francisco: Jossey-Bass.

Schwandt, T.A. (1997) 'The Landscape of Values in Evaluation.' Pp. 25-39 in D.J. Rog et al. (eds.) *Progress and Future Directions in Evaluation: Perspectives on Theory, Practice, and Methods. New Directions in Evaluation Number 76*, Winter 1997. San Francisco: Jossey-Bass Publishers.

Valovirta, V. (2002) *Evaluation Utilization as Argumentation*. *Evaluation* 8(1): 60-80.

Virtanen, P (2004) *The Rise and Fall of Evaluation Standards*. Forthcoming in *Hallinnon tutkimus* 23(2). Accepted for publication.

Von Wright, G.H. (1963) *The Varieties of Goodness*. London: Routledge & Kegan Paul.