FOUR

Selecting evaluation criteria

When we say that a policy is working well (or badly) or that a service is getting better (or worse) what do we mean? Do we mean that some or all members of the public think that it is so? Or that managers or politicians think so? Or that 'performance indicators' suggest that it is so? The words 'well', 'badly', 'better' and 'worse' are, to say the least, ambiguous.

For evaluations to be useful we need to reduce such ambiguities as far as possible and to be explicit about what we mean. This is where the notion of a criterion comes in. A criterion is a principle or standard by which something may be judged or decided and derives from the Greek for a 'means for judging' (*New Oxford Dictionary of English*). In evaluating a policy or service it is important to be explicit about the criterion or criteria being used. Otherwise a claim that, for example, a service is 'good' or 'getting better' would be rather meaningless. Would it be getting better in relation to efficiency, effectiveness, equity, reducing costs, or what?

There are a number of criteria which can be used in making such judgements. The criteria are not mutually exclusive, and sometimes the boundaries between them might be blurred; sometimes the criteria might overlap; for example, the 'acceptability' of a service might depend on its 'responsiveness' to users' needs and wishes. We shall examine a range of criteria commonly used in evaluations (in no particular order) and set out what each is usually taken to mean. We also critically comment on the limitations of each criterion, demonstrating that in real-world evaluations judgements will always be problematic and contestable.

It is important to recognise that criteria are not always set or 'pre-ordained'. Sometimes they 'emerge from the specific social contexts of various stakeholders' (Abma, 1997, p 35). For example, the 'acceptability' of a spending programme, as far as the electorate is concerned, is clearly a largely political question.

Common evaluation criteria

Responsiveness

'Responsiveness' often refers to the speed and accuracy with which a service provider reacts to a request for action or for information. It can also refer to the demeanour and social interactive skills of the person delivering the service, and in this sense adds a dimension of 'quality' to the criterion of 'accessibility' (*q.v.*). Speed can, for example, refer to the waiting time before seeing a hospital consultant or the time taken to receive treatment in an Accident and Emergency department, or to the promptness with which a local authority town planning department responds to a request for planning consent.

In the perfectly responsive organisation (which does not exist in the real world), service users are likely to be delighted with the speed, friendliness, empathy and accuracy of all the providers all of the time.

As far as speed is concerned we would add a note of caution. An organisation might require its staff to answer the phone within so many rings to ensure that staff are 'responding' rapidly; but this might be followed by unhelpful, unfriendly or unsympathetic comments from the staff member. A speedy response is not always enough to qualify the response as a 'good' one. For example, in a care home, if a sick elderly resident calls for assistance when feeling distressed and confused but the carer responding to the alarm call does not have a high degree of communications skills or empathy, then the resident is unlikely to feel that the care has been responsive to his or her needs in any meaningful way. Responsiveness thus has many dimensions and the ticking of boxes relating to the *speed* of response is only one, albeit sometimes critical, element.

Equity/justice/fairness

For present purposes, these three concepts (equity, justice and fairness) will be taken to be synonymous. Many people will agree that they are in favour of equity but they are rarely in agreement about what exactly it means. It is commonly defined as meeting equal needs equally (and correspondingly treating unequal needs unequally). So, for example, in a just world people with equal needs would receive equal services. And people with unequal needs would receive unequal services. Thus an individual with a significant disability would, other things being equal, receive more help than someone with no disability.

It can be helpful to distinguish between 'procedural justice' and 'substantive justice'. For example, while our judicial system does not always get things right, it seeks to follow due process (for example, ensuring that those accused of a crime receive a fair trial) with the intention that just verdicts will follow.

Note that although there are many divergent views about what the terms mean, they are in practice commonly used; for example 'it's not fair'; 'unfair dismissal'; 'a fair day's work'. Political parties of all hues claim that their policies will bring about a fairer society (unless party leaders assert that there is no such thing as society). The difficulty is that in practice there will almost always be substantial disagreement about what counts as 'fair'.

Mooney (2009) has argued that it is important to view barriers (to achieving equity) and the height of barriers in terms of citizens' perceptions. Barriers need to be seen through the eyes of potential users of services. Such barriers can include the lack of information presented in an appropriate form, inconvenient geography, financial issues and lack of confidence. In these senses there are close links between problems of equity and problems of access (*q.v.*). In using equity as an evaluative criterion it is therefore necessary to examine the context in which potential service users are living in terms of educational levels, geography and financial constraints. Equity is also relevant to how finances are raised in order to provide public services. Are council taxes fairer than the community charges (poll tax) which they replaced? At central government level, is it fairer to rely more on direct or indirect taxes; income tax or value added taxes? These issues are important ones, but a detailed discussion of them is beyond the scope of this chapter.

We shall return to the closely related issue of ethics towards the end of the chapter.

Equality

Again, 'equality' is a term which is used in different ways; for example, does it mean equal outcomes? Or equality of utilisation? Equal opportunities? Equal access? Some people are better placed than others to take advantage of what might appear to be 'equal opportunities'. Over the years there have been numerous debates relating to inequalities; for example, geographical or social inequalities (or 'variations', 'diversities' or 'inconsistencies') in health states, treatments and outcomes.

The question of whether continuing inequalities should be regarded as a matter for serious concern is largely one of value judgements and

politics. Given that complete equality is unattainable, and arguably not desirable, a useful way to think about inequalities is to ask about what *degree* of inequality is acceptable in various circumstances.

In local government boundaries between local authorities commonly run along the middle of a street, especially in large conurbations. It is not unknown for council officials from different authorities to specify requirements in relation, for example, to food hygiene or health and safety at work. Businesses on opposite sides of the boundary road could be forgiven for feeling aggrieved when different standards are being applied on the other side of the road. Or in the case of national (or international) chain stores, surprise may be expressed by business branches in different localities when they are expected to comply with the law in different ways. The businesses might well expect a degree of uniformity or equality to exist across different local authority areas.

Equality rarely comes top of the list of priorities for people when evaluations are being undertaken, but it is often used in conjunction with other criteria. One often hears, for example, of the need for 'equal access' to services. However, in our view equality will normally lose out to equity as a high priority criterion.

Effectiveness

This concerns the extent to which stated objectives have been achieved. To use effectiveness as a criterion, aims and/or objectives need to be made explicit. Without such explicitness it is impossible to say anything convincing about the extent to which they have been achieved. In such a situation no one can know whether the organisation (or any individual member of staff) has been effective.

Explicit objectives also help in the following ways. To begin with they help professionals and their managers to question and abandon inappropriate and/or outdated objectives. The act of making such objectives explicit makes it easier to challenge them. This can make an organisation more flexible than if things are merely taken for granted and where practices are therefore never questioned. Secondly, having explicit objectives can help when decisions about priorities and resource allocation must be made. Thirdly, explicit objectives can help in undertaking staff appraisals or individual performance reviews; staff can be judged on the basis of whether they are achieving objectives, rather than on whether their 'faces fit'. Fourthly, they can help with staff motivation. If staff do not have a clear idea of what it is they are expected to achieve they may become unfocused and demotivated. In addition, it is easier to feel a sense of achievement (and therefore

possibly increased motivation) if clear objectives have been achieved. Fifthly, having explicit objectives can help with coordination. It can reduce the chances of an organisation pulling simultaneously in contradictory directions.

As far as possible and practicable, objectives need to be:

- *explicit:* they should be clearly documented;
- *specific:* they should not be general or vague; for example, 'to achieve a reduction in the infant mortality rate of *x*% in a defined locality by a specific date' is a more specific objective than 'to provide improvements in public health';
- *measurable*, if that is appropriate; it might not be possible satisfactorily to measure, in a quantitative way, all aspects of all services (for example, is it possible to measure all aspects of care for dying patients?); but where measurement is possible, then it can be very useful in preparing to evaluate the effectiveness of a service;
- *scheduled*: it should be clear by what date an objective is to be achieved; otherwise managers would never have to answer 'No' to the question 'Have you achieved the agreed objectives?'; they could simply say 'Not yet';
- *prioritised*: if possible, agreement should be arrived at about which are the most/least important objectives; this is particularly important when there are severe resource constraints on a service;
- *'owned'*: objectives should be seen to be attainable and preferably not imposed in a top-down way;
- *related to each other*: this can help an organisation to be clear about the strategic direction it is going in;
- *communicated* to those who need to know.

One of the reasons that effectiveness is an important criterion for evaluating a policy or a service is that it refers to the extent to which a service is achieving what it is intended to achieve. Or to put it another way: is the service in question having the intended 'effects'? The criterion is commonly used to address the issue of goals, objectives or targets. The principal question that is asked when evaluating effectiveness is 'To what extent and in what ways are the goals (or objectives or targets) being achieved?'.

The terms *goals*, *objectives* and *targets* all refer to some intended effect but they are often differentiated in terms of specificity. *Targets* are often seen as very specific and measurable; for example a local authority might ask 'Have we collected 99% of the council tax due by the end of the financial year?'.

Goals, on the other hand, usually refer to more general intentions; for example, 'to improve the health of a local population'. This is a laudable thing for which to aim but it is fairly vague. It is unclear what would count as improved health: perhaps a reduction in death rates, or better quality of health? And how would this be measured?

Objectives are commonly seen as a half-way house between goals and targets, more specific than goals but less specific than targets.

There are no rigid agreed definitions of goals, objectives and targets, and the three are sometimes used with a degree of interchangeability. One person's 'goals' are another person's 'targets'. There are also aims, purposes, mission statements and so on, again with little agreement on specific definitions. Sometimes the terms are used in a hierarchical relationship to each other. For example, at the highest level a *goal* might refer to improving the health of a local population and within this there might be some *objectives* such as providing hip replacements for everyone who needs them within six months of referral to a consultant. This objective might further be broken down into a number of *targets*, such as reducing waiting times to six months (by a particular date).

The more specific a goal (or objective or target) the easier it is to see what kind of data will need to be collected in order to evaluate the effectiveness of a service. But in evaluating the effectiveness of a service there might also be informal or unofficial goals to be considered. It is not uncommon for staff to have personal objectives or 'hidden agendas' which they hope will serve their own self-interests, and these 'political realities' need to be taken into account if effectiveness is to be evaluated thoroughly. If individual managers are achieving their individual goals relating to departmental growth but the organisation as a whole is failing to achieve its overall goals, one might conclude that the managers are effective at one level (that is, in relation to their informal goals) but ineffective at the organisation-wide level. This kind of conclusion should not be seen as a contradiction. It merely illustrates that multiple, and often conflicting, goals co-exist within organisations. Evaluators need to be clear about which goals they intend to use when evaluating effectiveness.

There can, of course, be 'goal-free' evaluations. This might occur when there are no explicit goals/objectives/targets. Here the evaluator seeks information about what is being achieved by the organisation in terms of benefits delivered for the organisation's clients and about the costs of doing so. Strictly speaking, such an evaluation is not assessing effectiveness but is using some other criterion such as efficiency.

One of the other problems about evaluating the effectiveness of a policy or service is that different stakeholders will often have different

objectives and priorities (Karlsson, 1996; Thomas and Palfrey, 1996). The 'clients' (the intended beneficiaries of a service), the clients' families, the front-line staff or professionals, the managers, the politicians, the public at large, pressure groups and the mass media might all have different perspectives. Indeed even within each of these groups of stakeholders there will rarely be identical expectations. Against this background of multiple objectives, which objectives should be used in the evaluation of effectiveness?

One starting point is to look at the 'official' goals and objectives published in a corporate plan or service plan, but it is unlikely that all the stakeholders will agree that these are the most important objectives. Should we just accept the official line or should we also (or instead) look at the objectives of the other stakeholders? If the objectives of various stakeholders are to be taken into account, whose objectives should carry most weight?

There is no simple answer to such political questions. One approach is to be explicit about the objectives and priorities to be used in the evaluation so that people can see what position is being adopted. This might result in an evaluation that concludes something along the lines of 'If we evaluate against the clients' objectives then the service seems to fail in relation to at least some of the expectations, but in relation to the managers' priorities then we have to arrive at a different conclusion …'. One might therefore end up with apparently contradictory conclusions, but this is not necessarily a bad thing. It is often better to provide an evaluation that takes account of multiple perspectives rather than to restrict the evaluation to the objectives of one group with the vested interests that that might entail.

In the often uncomfortable real world, there are some serious limitations and difficulties in using effectiveness as an evaluative criterion. For example, goals and objectives are commonly stated in vague and ambiguous terms. To some extent this is a problem of terminology. When people talk about goals, aims, or mission statements there is often no expectation that these will be specific or measurable. The intention is rather to give the organisation strategic direction. When people talk about objectives or targets, these are usually more specific and, in a managerial sense at least, they are likely to be more useful for evaluating the performance of the organisations or their departments and individual members of staff.

The second problem with the use of effectiveness is that it is easy to fall into the trap of assuming that effective performance is the same as 'good' performance. If the goals and/or objectives being used to evaluate effectiveness are rather unchallenging, or if they are judged

to be morally repugnant (as when health care professionals murder their patients – of which there have been several cases), then there has been 'effectiveness' in that the staff member achieved what he or she set out to achieve, but most people would take some convincing that this amounts to 'good' performance.

A third problem is that if, as is often recommended, objectives are focused on outputs or outcomes, then the use of 'effectiveness' as a criterion will mean that the costs of achieving the objectives are likely to be ignored. It is widely accepted that in evaluating a policy or service the question of resources should be taken into account. Those concerned with the public sector are expected to make good use of the public resources made available, and ignoring cost is not a realistic option. One of local government's 'successes' in some parts of the country, London being a good example, has been the significant reduction of smoke pollution. The Clean Air Act 1956 empowered local authorities to declare 'smoke control areas' in which, with specific exemptions, it has become a criminal offence to burn coal. The work entailed a great deal of time spent by environmental health department officials in visiting huge numbers of premises to arrange the replacement of coal burning appliances with those that burned smokeless fuels (certain authorised solid fuels, gas or oil) or that used electricity. The policy outcomes were successfully achieved (significant reduction in smoke pollution and resultant decrease in the incidence of respiratory diseases). However, this evaluation does not take account of the resources used in implementing the policy. Could the similar outcomes have been achieved by a simpler and cheaper policy option, for example the taxation of coal? Policy options are not that simple but the point serves to illustrate the problem of evaluating outputs and outcomes without regard for the inputs (resources).

Another problem is that work undertaken to achieve goals and objectives might well have unfortunate side-effects. For example the reduction of the time that patients spend in hospital is likely to transfer a cost from the hospital budget but may impose an unacceptable strain on others – such as community services or the patient's family or friends. Evaluators, we would argue, need to take the blinkers off and examine the system-wide effects of changes in policies and services, so that account can be taken of the side-effects.

The fifth problem is that of questionable causal links. Effectiveness is largely a question of identifying and assessing the extent to which a policy or service has had (or is leading to) intended effects. But often it is difficult to be sure that a particular policy has been the cause of what appears to be an 'outcome'. This is particularly the case with

something which is influenced by a wide variety of factors. Health is a good example. Levels of morbidity and mortality are determined by genetic factors, environmental and economic factors, human behaviour patterns, diet and health care interventions. If an objective is to improve standards of health, and resources are put in place to employ more health care professionals, there could be a temptation a few years later to attribute any improved health to the increase in the number of medical and other staff. But the contribution of these increases could well be marginal in bringing about the apparent outcomes. Caution therefore needs to be exercised in concluding that policies have been 'effective'.

It is also worth noting that outcomes can be intended or unintended, and positive or negative. This is illustrated in Table 4.1.

Table 4.1: Possible outcomes of action

	Intended	**Unintended**
Positive	Freedom from pain following a hip replacement.	Reduced stress in the family of a patient whose life has been saved.
Negative	Patients murdered by a health care professional.	Iatrogenesis (for example, use of harmful drugs).

One way of classifying outcomes has been outlined by Helen Roberts (1991):

- Clinical studies (e.g. unplanned re-admissions, incidence and results of cases of MRSA infections)
- Health indices
  - Functional status (e.g. physical, mental, social)
  - Health perceptions (e.g. well-being, pain)
- Measures of disability and distress (e.g. Quality Adjusted Life Years; QALYs)
- Patient satisfaction
- Death (e.g. various kinds of mortality rates and ratios).

Another way of thinking about outcomes is that they can relate to individuals or groups in the populations. It might often be easier to attribute causality in the former (for example, the case of a hip replacement) than in the latter.

The task of evaluating organisational structures and processes can be particularly difficult in establishing causal links. Several attempts have

been made to do this in relation to the NHS. For example, Sheaff (Department of Health, 2006) has analysed data collected from senior stakeholders in the NHS by interviews, telephone interviews and focus group work relating to the apparent impact and influence of organisational arrangements. The key messages from this work were as follows:

- Highly centralised organisations are not associated with optimal performance.
- Organisational change needs to be developed from within, not imposed from outside, since professional engagement and leadership are crucial.
- Frequent reforms have made the NHS unstable.
- Mergers often miss the point (they should not take place merely to achieve a particular organisational size).
- Occupational 'silos' hamper change and innovation.
- Publishing clinical performance information does not influence consumer choice.
- The government should be cautious about promoting for-profit hospitals.
- There is no 'one right size' for each kind of NHS body.

These conclusions are inevitably contestable but they represent a laudable attempt to tackle the intractable problem of interpreting causality in a complex area.

So it can be seen that although effectiveness is commonly judged to be an important criterion – a judgement with which we concur – there are a number of difficulties and limitations inherent in evaluating effectiveness in practice. Nevertheless, it remains a key criterion. Without it people cannot know whether what they are trying to achieve is in fact being achieved.

Efficiency

Efficiency is the ratio between benefits (outputs or outcomes) and costs/resources (inputs). Efficiency can be increased, for example, by:

- increasing the benefits while holding costs constant;
- reducing the costs while holding benefits constant;
- increasing both benefits and costs, but increasing the former more than the latter;

- reducing both benefits and costs, but reducing the former less than the latter;
- reducing costs while increasing benefits.

It is also important to examine the *distribution* of the costs and benefits; for example a reduction in state provision might require an increase in costs/effort/resources by others (such as informal carers). Thus what might appear to be an increase in efficiency might actually be a mere redistribution of the costs (from the state to families).

Well-known examples of efficiency studies are 'cost benefit analysis' and 'cost utility analysis', for example the use of QALYs in the context of health policy and services. But these approaches are not without their critics and one of the difficulties about using efficiency as a criterion in evaluation is the problem of quantification – of costs, and especially of benefits.

Another difficulty is that because efficiency is essentially a ratio (of benefits to costs), how can one decide when a particular ratio represents an acceptable level of efficiency? In the real world the ratio will need to be compared with something to make the ratio mean something useful. One possibility is to compare it with the corresponding ratio in previous years so that trends can be assessed over time: is the efficiency of the service increasing or decreasing? An alternative is to compare the ratio with that being achieved in comparable organisations.

In the world of politics the notion of efficiency is often misused. When people argue that a change in the way a service is delivered will make it more efficient, it seems they mean that the amount of money that the service will cost will be reduced. This is often asserted without any reference to the changes that might come about in the level or quality of service, and so people are talking about efficiency when they really mean economy. Aspects of economic evaluation will be discussed in greater detail in Chapter Five.

Economy

We have noted that efficiency refers to the ratio between outputs (goods and services provided) and inputs (resources used to produce and deliver the outputs), and that effectiveness commonly refers to outputs and/or outcomes. The third 'E' in this triumvirate of 'value for money' elements is economy, which focuses mainly on the input (resources) side. We find this the least useful criterion of the three if it is used alone. It is of course helpful to know what resources are being

used but if it is not used in conjunction with its partner Es then it is of limited value in evaluation work.

It seems that when some politicians call for greater 'efficiency' in public services, they are actually calling for greater 'economy', that is, they wish to see a reduction in expenditure. As we have seen, efficiency is a ratio and it would be perfectly possible to have an improvement in economy with a simultaneous reduction in efficiency. If expenditure is reduced by *x*% and outputs or outcomes are reduced by more than *x*%, then efficiency will have worsened, not improved. Thus 'improvements' in economy do not necessarily lead to improvements in efficiency.

And if local authority (LA) A employs one environmental health officer (EHO) per 4,000 population and LA B employs one per 6,000 population, does it necessarily mean that B is performing more efficiently than A? It might mean that B has decided as a matter of policy to provide a lower level of service than A. In the case of social services, it might be that LA A employs one social worker per 1,000 population and LA B employs one per 1,500 population. Again, this does not necessarily mean that LA B is the more efficient; it might simply be that LA B has decided to behave more economically and offer a lower level of service than LA A.

We do not wish to suggest that economy is an unimportant criterion. A consideration of it can help public bodies in relation to good housekeeping, to identify and reduce unnecessary and wasteful expenditure. However, the examples illustrate the point that it is important for commissioners and evaluators to be clear about the distinction between efficiency, economy and policy decisions when considering level of service.

Accountability

At first sight, 'accountability' might seem a strange concept to include in a list of evaluation criteria. It is a criterion which focuses on organisational structure and processes rather than on inputs, outputs or outcomes but it can have a significant impact on these variables. If there are no clear lines and processes of accountability within an organisation, it is unlikely that the service it is providing will maintain effective and efficient services in the longer term. It is therefore a legitimate and useful issue to examine when evaluating a public sector organisation. Accountability is a particularly important issue in public sector management for, as Hudson points out (in Hill (ed), 1997, p 398), 'accountability is the link between bureaucracy and democracy'.

Elcock and Haywood (1980) have set out what they see as the four dimensions of accountability. They are:

- the location of accountability;
- the direction of accountability;
- the content of accountability; and
- the mechanisms of control.

In other words, *who* is accountable *to whom*, *for what* and *how*?

Location of accountability

Traditionally government ministers were held accountable for everything done by their departments. If things went seriously awry the minister was expected to resign. But in the 1970s this expectation was modified because it was thought unreasonable to expect ministers to be responsible for *all* errors made by civil servants. Furthermore, select committees have established the right to cross-examine civil servants as well as ministers. Accountability has thus become more 'shared'. It is therefore more difficult than it used to be to identify who is responsible or accountable for a particular decision or action.

Direction of accountability

The simplest answer to the question 'to whom are we accountable?' is commonly 'our line manager' or 'the body who has commissioned us'. But the question has become more complex over the years. For example, there is the question of professional status. One of the traits of a 'profession' is that its members expect a relatively high degree of autonomy in the way they plan and implement (and indeed evaluate) their work. Hospital consultants will often prefer to see themselves as being accountable to their peer group or to their patients (a form of 'consumerism'?) rather than to their 'line manager' or hospital chief executive.

A second complexity is the role of trades unions. Members of a profession (for example medics, nurses, social workers, environmental health officers) will sometimes find themselves in a situation where they are receiving conflicting instructions from their employer and their trades union, particularly in a situation where industrial action is being planned. Staff members are likely to feel a conflict of loyalty – to work or not to work. To whom is one mainly accountable?

And to whom is a local authority mainly accountable? To the local electorate? Or the council tax payers? The community at large? Central government? The courts? Service users? The local ombudsman? The mass media? Intended beneficiaries of services? Auditors and inspectorates? Or some combination of these depending on circumstances?

Sometimes, staff believe that they have a wider responsibility than to their direct 'managers'; for example, to what extent do civil servants have a wider responsibility to Parliament and the public if they suspect that a minister is misleading Parliament?

Content of accountability

Traditionally public sector managers were mainly held accountable for probity – which concerns uprightness, integrity and incorruptibility – and for not doing things which were *ultra vires* (that is, outside their legal powers). But as from the 1980s they have been increasingly held accountable for performance in terms of value for money; this is the world of efficiency, effectiveness, economy, centrally imposed targets, performance indicators, league tables and individual performance reviews. These changes have been seen as a part of the 'new public management' (NPM) (Ferlie et al, 1996; McLaughlin et al (eds), 2002) and have not always been welcomed by public sector managers. The measures have had the effect of concentrating the minds of managers, but not always on helpful targets.

Mechanisms of control

Control is the other side of the accountability coin (Peckham et al, 2005). If A is accountable to B, then it is supposed that B has some control over A. The control may be exercised through various mechanisms. An obvious mechanism is money. If B has control over a budget on which A is dependent, then A is likely to be cautious about ignoring guidance, requests or instructions from B.

A second kind of mechanism is democratic accountability. For example, it is one of the principles of democracy that, in theory at least, public sector managers are – directly or indirectly – controlled by voters. This kind of control may sometimes be assisted by reports from inspectorates and the mass media which can draw the electorate's attention to what the managers have been doing and not doing. There is also managerial accountability and control by way of performance review, fixed-term contracts and performance-related pay.

The law can also be an effective control mechanism. Professional staff are accountable to the public through the criminal law as in cases where professionals have been found guilty of assaulting their clients, through the civil law as when patients sue them for negligence, through contracts of employment and through professional conduct committees which can strike members of the profession off statutory registers.

Structures and processes of accountability are not commonly seen as criteria for evaluating public policies and services but we believe they are legitimate and useful yardsticks in evaluation work.

Accessibility

Accessibility is a commonly used criterion for evaluating public services, sometimes in conjunction with another criterion such as equality – equality of access often being seen as a desirable objective in relation to geography, so that some aspects of postcode lottery may be reduced. Examples of this kind of accessibility include distances that people live from the service in question, and these are sometimes measured in terms of journey times.

Another example is that of waiting times – the length of time that people have to wait for a service to be delivered, or for information to be given or for decisions to be made. In this sense the criterion comes close to being the same as 'responsiveness', which we referred to earlier. Other examples include physical access to buildings in relation to wheelchair access and signage in appropriate languages.

Some scholars measure 'use' of services as an indicator of degree of access (Mooney, 2009; Gulliford, 2009). It could be argued that if the potential to use a service exists, then access is present even when an individual makes the choice *not* to use a service (Gulliford, 2009). If one interprets the notion of access as 'freedom (or opportunity) to use' (Mooney, 2009, p 218) a service, then access can be further analysed into the elements of availability, affordability and acceptability (Mooney, 2009). But it is important to recognise that 'freedom to use' is different from use *per se*. For example, some citizens might not see themselves as suitable 'candidates' for a particular service; and services that require people to keep appointments may be more difficult for those with limited resources. Some people will have the resources to address their problems through private means, such as buying a house and moving to the catchment area of a 'good school'; others might struggle to access services because of poor public transport facilities, especially in rural areas.

Policies aimed at increasing access/permeability (Goddard, 2009) include:

- having waiting list targets; but research suggests that disadvantaged groups, such as the elderly, lower income groups and those from ethnic minorities, tend to be given lower priority by professionals when they are trying to improve waiting times; thus policies aimed at improving access might not reduce inequities;
- organisational re-design; for example NHS walk-in centres, NHS Direct, out-of hours primary care provision; but awareness of NHS Direct tends to be lower among lower income groups, and out-of-hours services often require travel at times when public transport is not available;
- targeting 'under-supply'; for example more resources to tackle the shortage of GPs in some parts of the country; but families who fail to benefit from the extra resources may have not defined themselves as in need despite the efforts of professionals; lower-income families tend to view their health in terms of events/crises as opposed to a process requiring maintenance/attention.

Despite the difficulties of defining and achieving accessibility, it remains an important declared aim in public services and it is therefore a useful criterion in evaluations.

Appropriateness

If a service is to be judged as 'appropriate' it should be relevant to the needs of intended beneficiaries. An important issue here is the way in which 'need' is defined because the way in which needs are assessed will often influence, or even determine, the shape of services to be provided. For example if the main need within the NHS is defined to be 'more hospital beds' and if that judgement is widely accepted, then the policy response might well be to plan the provision of more hospital beds. If however the need is defined as 'better health services' then the policy response is likely to be very different. Thus those with the ability – possibly because of their professional status and autonomy – to have their assessment of need 'accepted' in relation to a particular service are likely to be powerful players.

However, 'need' is a slippery concept. A useful taxonomy of need was developed by Bradshaw (1972) as follows.

Normative need is need as determined by a third party, normally a professional or 'expert'. It is the academic who tells students what

they *need* to read, the social worker who assesses the *needs* of a family who have a variety of complex problems or the medical practitioner who advises patients on changes they need to make to their life styles.

Felt need is, in effect, the same as 'want'. It is a need that is perceived or felt by the potential beneficiary of a service. When this is then turned into a request for help it becomes *expressed need*. This does not always follow from a felt need because for one reason or another a 'needy' person will not necessarily ask for help.

Comparative need is the situation where one person (A) sees another (B) in similar circumstances (for example in relation to a specific disability) and sees that B is receiving help to cope with the disability. A might compare his or her plight with that of B and conclude that as help is being given to B it should also be given to himself/herself (A).

The four categories of need are not, of course, mutually exclusive and it will often be the case that two or more kinds of need are present at any one time for a particular individual or group. The difficulty is that if appropriateness (or relevance to need) is seen as a useful criterion for evaluating a service, who is to assess the degree and kind of need in the case of the individual or group? In the public sector normative need is commonly to the fore. It is an essential part of the training of many public sector professionals that they should assess people's needs before planning how to provide any service that might be required. This is very different from private sector market-driven goods and services, where the consumer is more likely to be seen as all powerful or sovereign. Here, if consumers are able and willing to pay the price charged, they are likely to be given the service demanded. The degree to which this model of 'choice' is appropriate in the public sector remains a contentious issue and we shall return to the notion of choice later in the chapter.

It is not uncommon for there to be divergence between what is *needed* (normatively), what a client feels he or she needs (what is *wanted*) and the service that is *supplied*. Some examples, taken from Phillips et al (1994, p 160), are given in Table 4.2.

In evaluating services it can be useful to examine the relationship between these three variables (need, want and supply) to assess the extent to which, for example, there are services which are needed but not supplied, or services which are supplied but not needed.

Acceptability

A public policy or service needs to have a degree of acceptability in the eyes of those who are, directly or indirectly, affected by it; otherwise the

Table 4.2: Relationships between need, want and supply

Needed	Wanted	Supplied	Example
No	Yes	Yes	Giving neurotic hypochondriacs what they want to make them go away.
No	Yes	No	Unmet demands of neurotic hypochondriacs.
No	No	Yes	'Trigger-happy' surgeons or inappropriate interventions by social workers.
Yes	No	No	Undiagnosed high blood pressure.
Yes	Yes	No	Waiting lists.
Yes	No	Yes	Pressures to give up smoking or to reduce alcohol consumption.
Yes	Yes	Yes	Palliative care or caring for elderly dependent people in their own homes.

Source: Phillips et al (1994)

policy's sustainability is questionable, as in the case of the community charge (the so-called 'poll tax') introduced by the Thatcher government in the early 1990s. Positive 'approval' by all members of the public might be too much to hope for, and in many cases 'acceptability' might be the best that can be achieved in the real world.

There are a number of mechanisms that are available to assess acceptability. In a crude way, elections could be seen as one such mechanism so that if the electorate strongly rejects what a government is doing they can send a message through the ballot box. But voters make their preferences known at the ballot box on the basis of a wide range of considerations. What an election result indicates regarding the issue of the acceptability of a particular policy would be difficult to judge, though it might be easier in the case of a local election fought largely on a single issue, which is sometimes the case.

Even if more focused mechanisms were to be used to ascertain the degree of acceptability of a policy or service there are a number of issues that require consideration. For example:

- Should everyone's opinion carry equal weight? Or should some views be given greater emphasis if their advocates have some relevant expertise?
- How should the criterion operate when opinions are inconsistent? At one time a majority of a service's intended beneficiaries might approve of the service being evaluated; at a later date the approval or acceptance rating might be reduced, only to rise again later still. The present authors have often encountered situations where a group of students express an opinion about the most and least useful

parts of a syllabus. The following year, following what seem to be appropriate adjustments, the next group of students may think the opposite. This can make it difficult to know how the service should be changed, if at all.

- There might be occasions when the people whose opinions might be relevant are not available. This is a particular problem in the case of policies involving major capital expenditure which are likely to have long-term effects. Judgements often have to be made on behalf of future generations. The opinions of some groups may be difficult to ascertain: evaluators need to think carefully about how to consult any 'hard to reach' groups.
- Are people the best judges of what is in their best interests? If so, why is it that so many people still smoke? Public opinion is an important factor in evaluating public policies but perhaps it needs to be supplemented by the opinion of people who have particular expertise in the specific field. Members of the public are not always well informed, though there is always a danger here of adopting an inappropriately paternalistic or patronising attitude.
- Should we define 'public' widely (for example, the electorate, taxpayers, local community, etc) or narrowly (for example, the users of the service in question)? A decision might be made to define the term narrowly in order to focus attention on the opinion of those who have regular and firsthand experience of the service to be evaluated. For example, in the case of a leisure centre opinions might be sought from those who have used the centre in the last 12 months. The difficulty here is that there might be people who went to the centre more than a year ago and were put off by some aspect of the service (such as loud music) and who have, as a result, decided to stop using the centre. A narrowly defined 'public' (such as current users of the centre) would mean that the opinions of those who have stopped going will not be uncovered.

There are a number of mechanisms available for ascertaining the views of members of the public or service users; they include elections which, as we have seen, are a rather blunt instrument in the context of service evaluation – referenda, letters to state institutions, petitions, opinion polls, user surveys, focus groups, citizen juries and exercises in public participation.

An interesting and potentially useful account of the variables which customers are likely to regard as important determinants of the quality of a service has been provided in the form of a checklist by Morgan and Murgatroyd (1994):

- reliability: e.g. is the service performed at the designated time?
- responsiveness: e.g. willingness to provide the service;
- competence: e.g. possession of the required skills and knowledge to perform the service;
- access: approachability and ease of contact with the providing institutions;
- courtesy: e.g. politeness, respect, and friendliness of contact;
- communication: e.g. keeping people informed in language they understand; also listening to them; explaining the service, options and cost; assuring people that their problems will be handled;
- credibility: e.g. belief that the providers have the person's best interests at heart; trustworthiness and honesty;
- security; e.g. freedom from danger, risk or doubt;
- understanding/knowing the individual; e.g. making the effort to understand one's needs by providing individualised attention;
- appearance/presentation; e.g. the physical facilities, appearance of personnel, equipment used, etc.

Choice

Political parties commonly say that they are in favour of increasing the choice which the public have. It is therefore a legitimate criterion to use in evaluating public services. Evaluators might ask 'To what extent and in what ways has choice improved for members of the public?'. Creating markets in which providers compete with each other for 'customers' on grounds of price and quality will, it is often asserted, improve standards.

Conceptually choice is located at the 'exit' end of the 'voice–exit' continuum developed by Hirschman (1971). At the 'voice' end of the scale services are seen to be improved by the public or service users voicing their opinions and wishes and holding providers to account through the ballot box or other mechanisms for voicing their approval or not as the case may be. As competition and market forces increase their role in public services 'exit' becomes the key. What this means is that as 'customers' become dissatisfied with a service, instead of relying on their 'voice' they can 'exit' from the contract or relationship with a particular provider and 'choose' an alternative. The idea is that as providers become more aware that 'customers' can easily go elsewhere, they are likely to make greater efforts to reduce the costs of their services and to improve the quality of the service.

The notion of choice is closely related to other criteria. For example, how much choice do people have in the ways they can *access* services? And to what extent would increasing choice lead to greater *equity*?

However, how realistic is it for people to have real choice in relation to public services? In some cases the intended beneficiary of the service is not, for most practical purposes, the person given the choice. Patients are likely to be influenced by their GPs when deciding which consultant or hospital to 'choose'. And how much choice is there likely to be for services like fire fighting, police forces and sewage disposal?

In addition, there is no evidence that everyone always wants choice. And in some circumstances, for example in relation to choices of surgery or other health interventions, being given choices could even increase people's levels of anxiety. There is also the argument that not all citizens are likely to be equally able to take advantage of 'choice' (Dixon and Le Grand, 2006).

It thus remains an open question as to how important it is to have choice in public services, especially when it is unclear how we should deal with choice when it conflicts with other public policies. For example, to what extent should parents be encouraged (or allowed) to choose not to have their children vaccinated when the experts and politicians claim that it is in the public interest for all parents to cooperate with the vaccination programme (Gulliford, 2009)?

Nevertheless, given that politicians of several hues argue for increased choice, whether or not it improves services, choice has become an important criterion for evaluating public services.

The story becomes further complicated

Different stakeholders are likely to have different opinions about which evaluative criteria should be given most weight. Whose will prevails about this will depend on the amount and types of power that different stakeholders have, and this in turn might depend on the particular context at any particular time (Thomas and Palfrey, 1996). Possible relationships between stakeholders and the evaluative criteria that they are most interested in are summarised in Table 4.3. These relationships are certainly contestable and we offer the table simply to illustrate the point that not all stakeholders will agree on which criteria should be emphasised during an evaluation.

Those who are commissioning an evaluation might also have different priorities and might be in a position to insist that evaluators take their prioritised list of criteria as given. Evaluators might be able to negotiate

Table 4.3: Possible relationships between evaluative criteria and stakeholders

	STAKEHOLDERS				
CRITERIA	Those who pay for the service being evaluated	Intended beneficiaries	Professionals	Managers	Politicians
Effectiveness		*	*	*	*
Efficiency	*	*	*	*	*
Economy	*			*	*
Equity		*		*	*
Acceptability	*		*		*
Accessibility	*		*	*	
Appropriateness	*	*		*	
Responsiveness	*	*	*	*	*
Accountability		*	*	*	*
Ethical considerations		*	*		
Choice	*	*		*	

Source: Thomas and Palfrey (1996)

at the margins but it may be unlikely that they will be able to persuade the commissioners to make wholesale changes to the criteria to be used. Decisions about the use of criteria become further complicated when, as is commonly the case, several criteria are used at the same time. There is no 'correct' balance between the criteria. How much weight is placed on which criteria is largely a value-laden and political question. There is likely to be a good deal of 'trade-off'. An improvement in efficiency, for example, might be achieved by a reduction in the extent to which the organisation can be held accountable for its actions and inactions. Whether this would count as an improvement in performance or a deterioration depends on the point of view of particular stakeholders and their relative weightings on the criteria in question.

Ethical considerations

Politicians, professionals and managers do not generally like to be thought of as acting unethically. There is therefore a widespread feeling that services should be seen as ethical in the ways they are planned, delivered and evaluated. Ethics, a branch of moral philosophy, can be regarded as an evaluation criterion like those addressed earlier in the

chapter or can be regarded as an overarching consideration which can underpin the other criteria. For example, as we have seen, an evaluation of a service's effectiveness might reveal that people delivering the service are effective in achieving explicit aims and objectives, but this is not very comforting if the aims are not very challenging or if they are morally repugnant. In the case of efficiency, one can argue that policy makers and managers have a moral (that is, ethical) responsibility to increase the ratio between benefits and costs as much as possible. So no matter which criteria are being used to evaluate policies and services, the expectation that people should behave ethically is rarely far away.

Ethical theories tend to be more normative than explanatory; they say something about how things 'ought' to be rather than about how and why things happen in the way they do. The theories fall into three main categories.

Teleological theories

These are commonly referred to as consequentialist theories, for example, utilitarianism. In the view of utilitarians (the most famous of whom was Jeremy Bentham) a policy is formulated and implemented (or a service is delivered) ethically to the extent that the consequences produce the greatest good for the greatest number. Evaluation of a policy or service here focuses on the results or outcomes. 'Does the policy produce utility?' is the key question, and by 'utility' Bentham meant 'that property in any object, whereby it tends to produce benefit, advantage, pleasure, good, or happiness … or … to prevent the happening of mischief, pain, evil, or unhappiness' (Honderich (ed), 1995, p 85).

Utilitarianism underpins the way that the National Institute for Health and Clinical Excellence (NICE) evaluates health interventions (particularly drugs) on the basis of the expected outcomes or consequences that will follow their use. NICE investigates the likely outcomes of each intervention in terms of the extra quantity and quality of life and their costs in attempting to ensure that taxpayers' money is spent in a cost-effective way. Underpinning this is the notion that the state has an ethical duty to ensure that scarce public money is spent on goods and services which will produce the 'greatest good'.

A number of ethical objections have been raised to the use of Quality Adjusted Life Years (QALYs), which have become an established method of evaluating the expected cost-effectiveness of health interventions (Edgar et al, 1998). The criticisms 'may all be construed as the accusation that QALYs are inherently unjust' (Edgar et al, 1998,

p 74) and the way in which QALYs are used by NICE is evolving over time in an attempt to respond to the objections, but the idea based, as it is, on notions of utilitarianism remains an influential approach in formulating resource allocations decisions within health care systems.

According to utilitarianism, if there were four patients in urgent need of organ transplants (for example, two might need a kidney, one a heart and one a liver) then one could argue *prima facie* that a healthy person should be forced to give up their four organs for the greatest good: one patient dies in order to save the lives of four people. However, it is unlikely that most people would agree that this was the right thing to do as it seems to ignore the 'rights' of the healthy individual who has all his or her organs currently intact. It could be argued therefore that a theory which gives more emphasis to rights and duties would be more appropriate.

Deontological theories

This is where the second ethical school of thought comes in. The argument in this case is that certain acts are right or wrong in themselves irrespective of the consequences. Thus it could be argued that the rights of the healthy individual in the case above should be protected even if it means that the four needy patients will die. The emphasis is on the rights of individuals to enjoy autonomy and the moral duty that we have to respect those rights.

However, there are problems here too. For example, one of the acts which people are obliged to carry out (according to deontologists like Kant) is always to tell the truth. So if someone had the opportunity to save lives by telling a lie or by breaking a promise, then they should not do it. This seems counter-intuitive and many people would probably see it as ethically misguided.

A common example of the use of these theories is the code of conduct with which members of various professions are required to comply. These normally set out duties and the standards of behaviour expected of members of the profession in question, and serious contravention can lead to an individual being struck off the register of professional members and a resultant legal inability to practise the profession in question.

Virtue ethics

Whereas utilitarianism and deontological theories focus on actions (the first from the point of view of the results of the actions and the second

based on duties and right motives) virtue ethics emphasise the need to develop in individuals the right characteristics. Based on the early work of Aristotle, Foot (2003) has argued that there are a number of virtues, including courage, temperance, wisdom and justice, which are likely to benefit both the individual with the virtues and other people. Foot sees the virtues as 'dispositions', that is to say the characteristic or trait (such as courage or compassion) will become active in circumstances which require it. The virtuous person is thought or expected to act courageously, compassionately, etc, when the situation calls for it. The virtues are closely related to some of the criteria which were examined earlier in the chapter; see, for example, the notion of justice – the second evaluative criterion in the list set out earlier in this chapter. Generally speaking, the notion of virtue ethics places a responsibility on evaluators and others to fulfill their civic duties and to work for the public good. Evaluation practitioners are expected to act honestly, justly and courageously.

The above theories, in varying ways and degrees, underpin the guidance given to those engaged in the planning, delivery and evaluation of public policies. Gillon (1994) for example has argued that the following 'ethical factors' are relevant to the practice of medicine: honesty; equity; impartiality; respect; and adherence to high levels of competence.

Honesty

It is generally unwise for evaluators to issue promises to service users that an evaluation will lead to improvements in services. Although service improvement is an important aim, it commonly turns out not to be possible for various reasons. For example, there could be a change of personnel within senior management with incoming managers having very different ideas from those of their predecessors about the future shape of services, or there might be increased scarcity of resources compared with the date on which the evaluation was commissioned. Evaluations, and those who commission evaluations, therefore need to be honest with service users about the prospects for service improvements.

However, there are likely to be problems in being honest. Elsewhere we have drawn attention to:

> The possibility of legal action against the authors of evaluation reports when, for example, they have made statements which could be construed as defamatory, or

> if they have expressed doubts about the truthfulness of a statement but proceeded without verifying its validity. Conversely, evaluators may feel the need to suppress information or opinions if, by revealing them, they run the risk of exposing the informants to some form of reprisal (Palfrey and Thomas, 1996, pp 282–3).

Honesty and openness are also required about the criteria to be used in an evaluation and about how these came to be selected.

Equity

This principle involves the concepts of impartiality and natural justice. As seen earlier when considering equity as an evaluative criterion, it suggests that special consideration needs to be given to the position of disadvantaged groups. What should evaluators do if they disagree with the aims of a programme? For example, a public policy might be designed with the aim of reducing expenditure on a service which is of particular benefit to low income groups suffering social and economic disadvantage. As we have enquired elsewhere: 'Would (the evaluators) pass the portfolio to someone who did not suffer from such moral scruples or who held views similar to those of the commissioning agents?' (Palfrey and Thomas, 1996, p 283). There is no simple answer to this question but, as a minimum, evaluators need to ensure that the views of all principal stakeholders and interest groups are included in evaluation projects in which they are engaged, and that the contributions of all such groups are considered to be of potentially equal validity and relevance (Palfrey et al, 2004).

Impartiality

Partiality and inappropriate discrimination can take various forms. An important example as far as evaluators are concerned is to exclude the views of people whose opinions might be inconvenient to those in authority. As we have pointed out elsewhere:

> Evaluating a staff appraisal system wholly from the management criterion of its impact on productivity would wrongly exclude relevant and equally valid criteria constructed by staff who have undergone appraisal. (Palfrey and Thomas, 1996, p 283)

In a sense impartiality is a means towards achieving equity and 'although the evaluator can never be a completely neutral observer and interpreter, there has to be an attempt to keep an open mind about the evidence to be gathered and about its meaning' (Palfrey et al, 2004, p 162).

Respect

People's opinions about the policy or service being evaluated should, according to this principle, carry equal weight unless there are convincing reasons for not doing so. It could be argued that evaluators should pay more attention to the opinions of those who seem to be better informed about the issues involved, but this is not what happens in parliamentary elections. The notion of respect is taken seriously by evaluation approaches such as that espoused by Smith and Cantley (1985) in their requirement that a range of data collection methods should be used to collect various kinds of data from a variety of sources. This approach does not address the question of whether it is justified to give more weight to some views than others, but at least it recognises that potentially divergent views should be listened to seriously.

Adherence to high levels of competence

If evaluation projects were to be undertaken in nursing homes in order to make judgements about the quality of care, the commissioners of the evaluation might require a questionnaire to be used as the principal – or even the sole – means of data collection. But evaluators might feel that in order to get a full picture they would also need to spend time in the home observing practices and listening and talking to people (patients and carers) in informal settings at meal times, recreation times, and in sitting rooms etc. The response from the commissioners might be that the evaluation budget will not stretch to this. Evaluators would therefore need to decide whether to design a questionnaire which meets all the expectations for a competently designed questionnaire, or whether to withdraw on the grounds that the questionnaire alone will be an inadequate means of collecting the necessary data. To what extent should evaluators compromise their expectations and standards in this kind of situation? Should evaluators be satisfied if they do all that commissioners require of them? There is no simple answer to these dilemmas. In each case evaluators should at least try their best to persuade the commissioners to find extra resources to ensure that a high degree of competence can be achieved.

The ethical factors discussed above are espoused by Gillon (1994) in relation to medical practice. But there are also a number of ethical principles which have been advocated more broadly across professional life generally (Newman and Brown, 1996, Palfrey and Thomas, 1996). They include:

- respect for people's *autonomy* (including keeping people fully informed so that they can make informed choices);
- *non-maleficence* (attempting to ensure that one's actions cause no harm and to inform people if there are any unavoidable risks);
- *beneficence* (for example, using resources as beneficially as possible);
- treating people *fairly* (for example, providing everyone with the service to which they are entitled and not allowing any personal views about people's lifestyle to affect the quality of service given to them);
- acting with *integrity* (for example, genuinely striving to achieve the stated objectives of our activities).

Sometimes, the following are also included:

- *impartiality*;
- *confidentiality*.

Dilemmas can arise when these various ethical concerns conflict. If this occurs one can either aim for a compromise or give priority to one principle over the others. For example, the principle of autonomy might lead a medical practitioner to spend a lot of time with a patient to ensure that the patient can make an informed choice about treatment; but this might conflict with the principle of fairness/justice because every minute spent with the first patient means one minute less for other patients. There is unlikely to be an 'ethically correct' course of action. We must make a judgement which is 'appropriate' having taken into account the various principles.

Another difficulty relates to the debate between absolutism versus relativism. It can be argued that there are 'moral absolutes'; that is, some rules are universally 'right' in all circumstances. One of the difficulties with this view is that sometimes there might be a conflict between such rules. For example, if you believe that one must always tell the truth and one must always try to preserve human life, would it be right to lie to save a life? Which universal right should you violate? This is the problem to which we referred earlier when we summarised the position of those,

like Kant, who espouse deontogical theories. In such circumstances it is necessary to weigh and prioritise various 'absolute' obligations. The alternative is to take the view of cultural 'relativism'. Some relativists see absolutism as smacking of attempts to impose morally superior ideas on other people. Instead it could be argued that 'when in Rome one should do as the Romans do' and that what is 'right' varies from one culture, context or set of circumstances to another.

What is the role of the state in all this? Under the auspices of the Commissioner for Public Appointments in the UK, the Committee on Standards in Public Life has set out a number of principles – known as the Nolan Principles – which holders of public office are expected to follow. These seven principles are often seen as the elements of probity and are as follows (Nolan Committee, 1995):

- **Selflessness** – Holders of public office should act solely in terms of the public interest. They should not do so in order to gain financial or other benefits for themselves, their family, or their friends.
- **Integrity** – Holders of public office should not place themselves under any financial or other obligation to outside individuals or organisations that might seek to influence them in the performance of their official duties.
- **Objectivity** – In carrying out public business, including making public appointments, awarding contracts, or recommending individuals for rewards and benefits, holders of public office should make choices on merit.
- **Accountability** – Holders of public office are accountable for their decisions and actions to the public and must submit themselves to whatever scrutiny is appropriate to their office.
- **Openness** – Holders of public office should be as open as possible about all the decisions and actions they take. They should give reasons for their decisions and restrict information only when the wider public interest clearly demands.
- **Honesty** – Holders of public office have a duty to declare any private interests relating to their public duties and to take steps to resolve any conflicts arising in a way that protects the public interest.
- **Leadership** – Holders of public office should promote and support these principles by leadership and example.

Public officials and their managers and political leaders cannot reasonably be expected to base all their decisions and actions directly

on the ethical theories which we examined earlier as this would lead to the paralysis of analysis, but the Nolan Principles provide a useful framework. But again, in the real world there are problems. For example, much of what is contained in the Principles is open to interpretation. In the case of openness, they state that public officials should be as 'open as possible about all the decisions and actions they take'. There is likely to be considerable disagreement about what is possible and what is not in this context. The Principles, therefore, are useful but they certainly do not put an end to debate, nor should they. Nor can public officials, in our view, justifiably wash their hands of their responsibility for behaving ethically by simply 'delegating' such matters to an 'ethics committee'.

Concluding comment

A key issue associated with ethical theories is that of 'values' – the things that people view as important. Values are at the root of decisions about which evaluative criteria are to be regarded as the most important; people's value stances need to be made explicit if opinions about what constitute 'success' is to be as transparent as possible. One of the problems here is that values cannot be arrived at rationally (MacIntyre, 1985). If managers have goals (based on the values of those who set the goals) they can seek to be rational in assessing and selecting the 'best' means to those goals. As we have pointed out elsewhere:

> MacIntyre's view of managers is that they treat ends as given and as outside their scope. In their role as manager, they are unable to engage in moral debate. Their task is to restrict themselves as to the realms in which rational agreement is possible – that is, from their point of view, to the realms of fact, of means and of measurable effectiveness. (Palfrey et al, 2004, pp 166–7)

However, managers need to be aware of the dangers of disadvantaging key stakeholders, especially when an evaluation remit seems to compromise professional standards.

We can conclude that what happens in any evaluation process depends on the value judgements of those with most of the power in the system – be they politicians, professionals, managers, the mass media, pressure groups, a ruling elite, the electorate, or whoever. What evaluators need to do is to make the values which are dominant as explicit as possible so that they become visible and therefore contestable

by all interested parties, particularly those whose services are being evaluated. This is one way in which we can ensure that evaluations are ethically informed.

We also need to remember that each of the criteria examined in this chapter can be useful in evaluation work but, as has been seen, each also has its limitations. This is not a reason to eschew the criteria, but these limitations should be recognised in our evaluation work.

References

Abma, T. (1997) 'Playing with/in plurality', *Evaluation*, vol 3, no 1, pp 25-48.

Bradshaw, J. (1972) 'A taxonomy of human need', *New Society*, March, pp 640-3.

Department of Health (2006) *Achieving high performance in health care systems: The impact and influence of organisational arrangements*, NHS Service Delivery and Organisation R&D programme.

Dixon, A. and Le Grand, J. (2006) 'Is greater patient choice consistent with equity? The case of the English NHS', *Journal of Health Services Research and Policy,* vol 11, no 3, pp 162-6.

Edgar, A., Salek, S., Shickle, D. and Cohen, D. (1998) *The ethical QALY: Ethical issues in healthcare resource allocations,* Haslemere: Euromed Communications.

Elcock, H. and Haywood, S. (1980) *The buck stops where? Accountability and control in the NHS*, Hull: University of Hull.

Ferlie, E., Ashburner, L., Fitzgerald, L. and Pettigrew, A. (1996) *The new public management in action,* Oxford: Oxford University Press.

Foot, P. (2003) *Virtues and vices and other essays in moral philosophy*, Oxford: Oxford University Press.

Gillon, R. (1994) *Philosophical medical ethics*, NY: John Wiley.

Goddard, M. (2009) 'Access to health care services – an English policy perspective', *Health Economics, Policy and Law*, vol 4, pp 195-208.

Gulliford, M. (2009) 'Modernizing concepts of access and equity', *Health Economics, Policy and Law*, vol 4, pp 223-30.

Hill, M. (ed) (1997) *The policy process: A reader*, London: Wheatsheaf.

Hirschman, A. (1971) *Exit, voice and loyalty: Responses to decline of firms, organisations and states*, Harvard, MA: Harvard University Press.

Honderich, T. (ed)(1995) *The Oxford companion to philosophy*, Oxford: Oxford University Press.

Karlsson, O. (1996) 'A critical dialogue in evaluation; how can the interaction between evaluation and politics be tackled', *Evaluation*, vol 2, no 4, pp 405-16.

MacIntyre, A. (1985) *After virtue: A study in moral theory*, London: Duckworth.

McLaughlin, K., Osbourne, S. and Ferlie, E. (eds.) (2002) *New public management: Current trends and future prospects*, London: Routledge.

Mooney, G. (2009) 'Is it not time for health economists to rethink equity and access?', *Health Economics, Policy and Law*, vol 4, pp 209-21.

Morgan, C. and Murgatroyd, S. (1994) *Total quality management in the public sector*, Milton Keynes: Open University.

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

Nolan Committee (1995) *Standards in public life*, London: HMSO.

Palfrey, C. and Thomas, P. (1996) 'Ethical issues in policy evaluation', *Policy & Politics*, vol 24, no 3, pp 277-85.

Palfrey, C., Thomas, P. and Phillips, C. (2004) *Effective health care management: An evaluative approach*, Oxford: Blackwell.

Peckham, S., Exworthy, M. and Greener, I. and Powell, M. (2005) 'Decentralising health services: more accountability or just more central control', *Public Money and Management*, vol 25, no 4, pp 221-30.

Phillips, C., Palfrey, C. and Thomas, P. (1994) *Evaluating health and social care*, London: Macmillan.

Roberts, H. (1991) 'Outcome and performance in health care: survey of current activities', in Public Finance Foundation, *Outcome and performance in health care*, Discussion Paper 40, London: PFF.

Smith, G. and Cantley, C. (1985) *Assessing health care: A study in organisational evaluation*, Milton Keynes: Open University Press.

Thomas, P. and Palfrey, C. (1996) 'Evaluation: stakeholder-focused criteria', *Social Policy and Administration*, vol 30, no 2, pp 125-42.