Setting the scene: the four domains of evaluative practice in higher education

Murray Saunders

Introduction and approach

This book marks a departure from convention in that it seeks to dismantle traditional boundaries in approaches to evaluation. It does so in a sector that, ironically, is replete with many forms of evaluation, but has yet to become a focus for research into evaluative practices. This book is about how value and worth is attributed to what goes on in Higher Education (HE). It is about how society, through its governance structures, decides on what is worthwhile, how its agencies attribute value to its policy and programme interventions, how institutions decide on the quality and merit of its internal practices and how groups of stakeholders (teachers, researchers, students, external collaborators) decide on the value of what they are doing. It is not a book on the evaluation of HE but on evaluative practice in HE.

We have many single accounts of evaluations of HE policy, programme interventions, inspectorial and quality practices, but this book’s focus is the range of practices involved in the attribution of value wherever it may occur. It conceptualises, theorises, and gives empirical examples of evaluative practices across the HE sector in both the UK and, comparatively, in diverse international contexts. The book argues for a specific position on evaluative practice as embedded in particular academic contexts.

The position taken here is that the idea of evaluation needs to be recast. Evaluation theory has a history of introspection (common in areas of social activity striving to position themselves with a distinctive disciplinary identity or professional status) that has resulted in a preoccupation with ‘approach’ and a plethora of evaluative ‘studies’ i.e. of programmes, policies and interventions. It is unusual to focus on evaluation practice as an object of study outside the painstaking exposition of approach in the context of a report on an evaluation or a theoretical, ethical or procedural justification for a particular stance. The preoccupation with demarcations and identity can be a distraction. What we are interested in is constructing a representation of
Setting the scene

evaluation as a type of social behaviour using a generic model of evaluative practice. This requires two steps. The first is to be clear about what we mean by ‘practice’; the second is to be clear about what we mean by ‘evaluative’.

What we mean by ‘practice’

We have adopted the perspective that depicting and understanding what goes on in social domains like HE requires an operational definition of social practice. This perspective denotes a concern with activity, with behaviour, with what people do, what they value and what meanings they ascribe either singly, in groups, in institutions through their systems, or nationally through national managing structures. We are not interested in abstract or a priori discussions about theoretical prescription or turf wars on epistemology. To that extent our approach is grounded in whatever it is that people do in the name of value attribution or evaluation. So, how do we depict or understand ‘what people do’? At its core, what people do is a social phenomenon, multi-hued of course, but we consider it to have discernable characteristics. What people do then can be termed ‘practice’, and all social life can be interpreted as consisting of a series or clusters of practices in different fields of activity, within families, friendships groups, at work and so on.

The idea of practice is a key aspect of sociocultural theory, and it takes as its unit of analysis social practice, instead of (for example) individual agency, individual cognition, or social structures. By social practice we mean the recurrent, usually unconsidered, sets of practices or ‘constellations’ that together constitute daily life (Huberman, 1993, talks about ‘engrooved’ practices). Individuals participate in practices, as ‘carriers’ of ‘routinized ways of understanding, knowing how and desiring’ (Reckwitz, 2002: 249–250). So, a social practice perspective leads us to direct our interest towards how a nexus of practices affects, for example, the way that evaluation is undertaken in a particular locale (Reckwitz, 2002: 258). Moreover, a social practice viewpoint alerts us to the danger of a rational-purposive understanding of change, one which assumes that people on the ground will act in ‘logical’ ways to achieve well-understood goals, or that managers and policy-makers will have clear and stable goals in mind and be able to identify steps towards achieving them. This ‘hyperrationalised and intellectualised’ (Reckwitz, 2002: 259) view of behaviour just does not stand up to scrutiny in university contexts. In this book then, practices can be usefully conceptualised as sets or clusters of behaviours forming ways of ‘thinking and doing’ associated with undertaking evaluative activity, this includes the rooted identities and patterns of behaviour that characterise shape and constrain understanding of evaluative practice. As cultural phenomena, practices can be defined as:

a routinized type of behaviour which consists of several elements, interconnected to one another: forms of bodily activities, forms of mental activities, ‘things’ and their use, a background knowledge in the form of
understanding, know-how, states of emotion and motivational knowledge . . . [A] practice represents a pattern which can be filled out with a multitude of single and often unique actions reproducing the practice . . . The single individual – as a bodily and mental agent – then acts as the ‘carrier’ of a practice – and, in fact, of many different practices which need not be coordinated with one another.

(Reckwitz, 2002: 249–250)

So, the social practice perspective of sociocultural theory sets into the foreground social practices and focuses on the way practice itself, in whatever domain, becomes an object of scrutiny. This perspective integrates a number of theories: those that explore professional learning processes (see Eraut, 2000; Schön, 1991), those that develop the idea of practice itself (Giddens, 1976; Lave and Wenger, 1991; Wenger, 1998 and 2000), along with the concept of the knowledge resources (formal, explicit and technical, on the one hand, informal, tacit, social, cultural and discursive on the other) that are produced and accessed, metaphorically as ‘rules’ (Blackler, 1995; Bereiter and Scardamalia, 1993).

Practices are inherently social and evolving. They are nested in cultures that form a major part of the intellectual, moral and material resources on which the practices themselves depend: cultures and practices constitute one another – they are not separable. Constellations and clusters of practices are bound together by social groupings, which are sometimes called communities of practice, sometimes called activity systems. We are able, however, to depict social life as constellations of practices within particular domains and at the same time, cross cut by horizontal and vertical considerations associated with distributions of power and resources, gender, ethnicity, identity/biography and place. The focus of this book is the constellation of practices associated with HE and further, the constellations of practices which are evaluative in nature.

What we mean by ‘evaluative’

Working within this approach, we can see that evaluation is essentially a social practice. It is undertaken by people, within structures of power and resource allocation. We can say that evaluative practice is social practice bounded by the purpose, intention or function of attributing value or worth to individual, group, institutional or sectoral activity. In order to depict evaluative practice within a sector of activity like HE, we can discern four domains of practice:

- National/systemic
- Programmatic
- Institutional
- Self.
These domains are not definitive but will form the basis of the structure of this book, and are described below. But first, we will summarise what social practice theory means when linked to evaluation. Stern (2006: 293) has outlined the complexities of the widely shared view that evaluation is a ‘practical craft’ by identifying at least seven taken for granted assumptions about where the emphasis should be. These range from evaluation as technical practice, as judgement, as management, as polemical and as social practice. We are defining evaluation as clusters, or ‘constellations’, of practices which form a ‘first order’ definition of which the various manifestations in his formulation are expressions. As Stern points out, these expressions may well be very different in focus and not particularly consistent; it is these differences this book sets out to explore.

Social practice and evaluative practices within HE

The application of social practice theory to evaluation leads us to take a perspective which emphasises the way evaluative practices are embedded in specific HE contexts. Without this appreciation of the contextual factors which affect outcomes, then evaluative practices will be ineffective (Pawson and Tilley, 1997: 114). This means that in our perspective, evaluation,

- is characteristic of all social policy areas.
- involves dimensions of evaluative practice consisting of symbolic structures, particular orders of meaning in particular places, and has unintended effects.
- consists of practices which use implicit, tacit or unconscious knowledge as well as explicit knowledge.
- can have progressive enabling characteristics but are also perceived and experienced as controlling, as part of a ‘surveillance’ culture.

The approach depicts the growth in externally derived requirements for judgements about value in a domain (HE) that has traditionally been self-policing. This can be understood as a form of ‘breakdown’ of the tacit contract between professional groups and society (through its governance structures) that enabled them to police themselves through self-regulatory, internalised ethics, standards and frameworks for action. The way we ‘value what we do’ therefore has changed. We depict the growth of a complex set of evaluative practices that are embedded within the sector and are occupying an increasing amount of time, energy and new expertise. This book will examine these practices which are located within four domains.

How to describe the ‘object’ of this book in this way is not straightforward. The four domains constitute an analytical frame which could be used to represent evaluative practice in any social policy sector (health, criminal justice, social services etc.). We will present these four domains as a way of organising the focus for our contributors. It will show that depicting evaluative practice within HE in this way foregrounds the situated nature of the
purposes and uses of evaluation highlighting strong differences in emphasis. The next section explains what these four domains are.

**National systemic evaluative practice**

This domain concerns the embedded practices associated with the way a national sector is regulated, controlled or managed within what has been called ‘the evaluative State’ (Neave, 1998). In many cases the practices that are associated with the way in which value is attributed across HE institutions within a nation state are inspectorial, standardised, associated with audit, externally imposed and used for comparative purposes or ranking. Within the UK, the Research Assessment Exercises (RAE, now called REF (Research Excellence Framework)), the Quality Assurance Agency (QAA) institutional reviews, the periodic reviews of departments, are all examples of this kind of systemic evaluative practice. For example, HEFCE (The Higher Education Funding Council for England) sanctions the QAA to coordinate the institutional audit processes of teaching quality which have a regulatory function:

The mid-cycle follow up is an integral part of the overall institutional audit process and will support the same aims. It will serve as a short health check, for the institution and for QAA, on the institution’s continuing management of academic standards and quality of provision, normally some three years after an institutional audit. It will be an opportunity to reflect upon developments made in the management of standards and quality within the institution since the previous institutional audit, and, in the context of the findings of that audit, for QAA to advise the institution of any matters that have the potential to be of particular interest to the team that conducts the institution’s next audit.

(QAA, 2009: 1)

Often, such audit processes will have unintended effects that create a backlash of influence on routine practice within institutions. All of the countries within the UK undertake equivalent processes, and all sectors within the UK have their equivalent systemic evaluative practices. In the UK school system, for example, all schools are subject to review and inspection by the Office for Standards in Education (OFSTED), while social work is reviewed by the Social Work Inspection Agency. Interestingly, these processes within the UK system operate within a structure which in many ways celebrates the semi-autonomous governance of UK HE institutions. Their power lies in the effect such processes have on the procurement of resources and the reputation on which institutions thrive. The example from the UK of the Research Assessment Exercise (RAE) has as its primary purpose the following:

- to produce quality profiles (see paragraph 30 and Annex A) for each submission of research activity made by institutions. The four higher education funding bodies intend to use the quality profiles to determine their grant for research to the institutions which they fund with effect
from 2009–10. Any HEI in the UK that is eligible to receive research funding from one of these bodies is eligible to participate.

Critiques of this orientation within the UK HE sector have centered, in particular, on the way in which assessments of research outputs distort the quality and nature of the research process as Walford (2000), for example, has suggested. He argues in the context of research in the Maths discipline that an RAE usually involves a double use of peer review in that a researcher has to submit publications, and these will in general have been peer reviewed. He notes that the review by the RAE panels is itself a peer review and argues that this exclusive reliance on peer review is a major defect of an RAE, for it is likely to lead to a systematic failure to recognize ground-breaking research. He reminds us that the study of the history of science shows that peer review can give results which are conservative and it often happens that researchers produce work which is judged at the time by their fellow researchers to be worthless, but which is later (sometimes much later) recognized to have been a major advance.

The RAE is one example of a systemic evaluation which is located, in this case, within UK HE. However, the book is not limited to the UK experience. It compares practice across cultures and experiences, to scan the range of practice at this domain. Cross-cutting themes involve degrees of control, regulation, autonomy and consensus within specific country systems.

Internationally most systemic domain evaluative practice functions in this way. In Australia, the THES reports a debate which is currently in full flood which is arguing over the pros and cons of peer versus metric-based systems of assessment of research quality. Peer review will continue to form a ‘fundamental’ element of Australia’s national system for assessing the quality of university research. Kim Carr, the Innovation, Industry, Science and Research Minister, has launched a consultation on how to replace the now-defunct research quality framework. The new system will use metrics – measurements of research outputs, such as the number of times an academic’s published research is cited by other scholars – but it will continue to incorporate an element of peer review. The move comes as the UK consults on plans to replace a peer review-based research assessment exercise with a system based largely on metrics for science subjects. Australia’s new system, to be called Excellence in Research for Australia (ERA), will use ‘metrics, including citation data, to help keep costs down’. In the USA, in contrast, sector-wide systemic evaluative practice does not exist. Evaluative practices are solely undertaken at institutional levels and tend to be driven by issues associated with the HE market place, recruitment and fund raising.

**Programmatic evaluative practice**

This domain of evaluative practice is concerned with more familiar territory for most individuals in HE, in that it covers evaluative practices associated with policy or programme interventions of various kinds. Evaluative practices in the
programmatic domain deal with policy and project interventions designed to change the emphasis of national processes and practice. An example is the evaluation of the Institute for Learning and Teaching in Higher Education (ILTHE) in the UK undertaken by the Teaching Quality Enhancement Committee (TQEC), chaired by Sir Ron Cooke (2003), or evaluation of the Enterprise in Higher Education initiative in the early 1990s.

More recent examples are the evaluation of the Scottish Enhancement Themes, undertaken by two of the editors of this book, or the Centres for Excellence in Teaching and Learning in England and Wales. What is clear is that evaluations in the programmatic domain are often used as policy tools (Bleiklie, 2002: 30), since the evaluation is normally trying to ascertain whether investment in the programme has been worthwhile. Evaluations of interventions, programmes or policies within HE are not uncommon, and the evaluation is normally justified by various types of theoretical position. An example of the approach adopted for a policy evaluation concerning the Scottish Enhancement strategy is embedded in the following extract:

The overall approach to be taken and a provisional schedule of work are outlined below. It builds upon that taken in our previous evaluation of the SHEFC’s (now SFC) Strategy for Quality Enhancement (2003–2006). In brief, it combines two powerful traditions of evaluation – Utilisation Focussed Evaluation and Theory Based Evaluation.

Utilisation Focussed Evaluation takes seriously the needs of commissioners of evaluations and has a strong sense of the need for evaluations to have usability. It is responsive to needs that evolve and opens up multiple layers of communication. Its style is communicative yet independent. Theory Based Evaluation focuses on the connections between strategies and intentions. It has a firm basis in evidence but is open to unintended and unanticipated processes and outcomes. It helps to articulate the informal theories of change embedded in policy strategies and the adaptations and modifications which occur as a policy is created in practice. Theory based approaches also recognise the importance of a strong theoretical framework within which, in this case, change might be understood. Often such theories are derived from the social sciences and act as a way of proposal to evaluate an aspect of the Scottish Funding Council’s programme of themed support in which changes in practice formed the core focus:

We have built the evaluation on the idea of ‘practice’ as a key indicator of cultural change but we include in this idea, ways of thinking and writing about quality (we see this as a form of rhetorical practice) as well as day to day practices by stakeholders at different levels in the sector and systemic practices undertaken by institutions, departments or schools and course teams along with sector wide systems associated with the overall strategy.

(Saunders et al., 2008)
Setting the scene

It is interesting to note that programmatic evaluative practice, typified in the case above, can be informed by theory in different ways. It is a cliché to write there is ‘nothing so practical as a good theory’. We suggest that theory can enter consideration at a very early moment in evaluation design. In other words, theories orientate evaluative practice in this programmatic domain and determine the kinds of claims that might be made on the basis of the evaluation.

A starting point concerns the theories of evaluation as inquiry embedded in an evaluation either explicitly or implicitly. This is usually connected to a theory of social science. There are some obvious parallels, for example, between the uses of random controlled trials or their variants in evaluation, for example, and a view of the social world that assumes there are social facts out there to be discovered and tabulated, i.e., a form of positivism. This is a method supported and practised by Carol Fitzgibbon (1998) in the UK. This can be contrasted to an approach developed by Saville Kushner (2000) who suggests that constructing the stories or narratives of key participants and personalising evaluation is an important approach in the evaluation process. Whatever our predilection, we have to make some choices about how we connect to the social and technical world in order to identify what counts as evidence of what.

A second, rather more methodologically neutral issue (in the sense of the form evidence or data might take) concerns theories of evaluation as a process, i.e., the way evaluations should be carried out (mainly theories about how evaluations connect with elements or stakeholders in an evaluation process). Examples of this might be evaluation theory that espouses the advantages of a goal free approach, in which according to Scriven (1991), programme intentions are not relevant but the experience of a programme is. Michael Patton’s utilisation focused approach (1997) to evaluation is another, in which key elements of a design explicitly express the interests and intentions of the evaluation commissioners and programme designers. This is the approach adopted in the Scottish Enhancement strategy example offered above.

Third, Carol Weiss (1997) and Connell and Kubisch (1999) discuss the way we should be looking for the underlying theories about change that guide programme designs and it is this that should form the focus of our designs. This can be understood as programme logic. The idea of a bad change theory is an interesting one. There is a debate in the UK at the moment on whether or not a pilot project approach embodies a good or bad theory of change. On what basis should a good example of practice, embedded in the special circumstances of a pilot, create wider changes? We have heard it rather pithily expressed that ‘the reality of a roll out from a pilot involves the application of fewer and fewer resources to more and more resistance’. The challenges associated with scaling up are many.

The fourth way in which theory has a presence in evaluation, especially programmatic evaluation, is through social theory of a more general kind. These are theories which provide explanatory frameworks for evaluation that can structure our effort, suggest what kind of data is useful and, in our view,
enhance the chance of the evaluation making a contribution to positive developments.

Evaluative work based in the Centre for the Study of Education and Training (CSET) in the Department of Educational Research at Lancaster University over the last 15 years for example, and continued in HERE@LANCASTER, in the same department, is influenced by social practice theory and is consistent with the approach advocated in this book. As we note above, it is an approach that emphasises the situated activities and experiences of the stakeholders who are involved in or affected by the programme.

These stakeholder activities and experiences should form the centre of an evaluation and will yield the resources for judgement about value and worth. It also does justice to the diversity of experience and the voices of all those in the programme’s orbit. We can illustrate this by reference to a metaphor embodying implementation theory called the ‘implementation staircase’. The staircase metaphor is intended to capture the idea that the programme’s messages are understood and acted upon divergently by stakeholders as the programme’s message goes down and up the staircase of implementation. We understand the project implementation process as highly adaptive and practice-based. This conceptual framework emphasises the way in which policy messages are adapted and modified through the process of enactment. It is also important to understand the way in which policy messages are transmitted through a system and are modified and adapted as they move from one group of participating stakeholders to another (see Saunders et al., 2004).

Stakeholders occupying steps on the implementation staircase are both recipients and agents of policy and through this process the project message will undergo adaptation and be understood very differently according to the unique and situated experience of each stakeholding group. Crucially, it is their interpretation of policy priorities, emphases and embodiments that are passed on to other stakeholders. There are two implications for evaluation here. First, policy should be depicted as multiple policies in action through the experiences of different stakeholders and second, policies and programmes are shifting and evolving entities depending on stakeholder experience of them. The metaphor of the implementation staircase is used to capture this process of policy implementation in which messages go both down and up the staircase but are modified and adapted as they go. An evaluation approach conceived within this understanding will necessarily evolve and take cognisance of the developing policy situation.

In many programme evaluations there is a preoccupation with impact. There are technical, practical and methodological difficulties in being able to isolate a line of determination between a particular intervention or programme and a change in practice at the level of an individual teacher or the enhanced experience of a student, let alone an increase in learning. However, through tracer studies or trajectory studies it can be possible to provide indicative evidence of the effects/impacts of activities by following up on participants and observing/interviewing/looking at students’ work etc.
back in the institutional setting. An important factor to take into account in this process is that impact can also be shaped by the complex and differing culture of departments and disciplines within an HE institution or community. Widely varying departmental structures, ways of communicating, and dominant beliefs about the nature of their core work impede or assist the impact of an intervention.

We would look for evaluative practice at this level, then, that is informed by the trajectory of an intervention or a series of activities which are concerned with charting stakeholders’ experiences of change. This involves looking for cases or examples and tracing their evolution from pre-adoption analyses through to adoption strategies and then on to implementation. The Trajectory Study is a common evaluative practice within this programmatic domain. It involves following the progress of an activity/intervention or cluster of activities from conception to end user experience. Trajectory studies can be accumulated over time and can be undertaken on clusters of activities. The technique involves the following process and chronology:

- What was the genesis of the activity? (Where did the idea come from, why was it selected?)
- What was the intervention strategy for the activity, and why was the method chosen? (How was it disseminated and introduced to the target group, e.g. seminar, newsletter, course, individual consultant, workshop, curricular materials etc.?)
- How was it received by the client group? (What was the quality of the intervention activity in terms of participant experience?)
- How was the activity reconstructed by the target group? (How did the target group reproduce ideas, practices, and ways of doing things etc. once back in their normal working environment?)
- What new practices are beginning to emerge which embody the ideas of the original activity and what is their orbit of influence? (Are new practices being routinely used by the target group and are they influencing a wider constituency? Have there been any institutional changes?)
- What is the experience of the end user (students) of these changed practices? (What is the quality of the learning experience, and the quality of assessed work?)

The Trajectory Study would consider the impact of the programme in terms of enabling outcomes (new things in place), process outcomes (new ways of doing things) and changes in student achievements and experience (improvements). We would argue that impact evaluations should identify both desirable and undesirable outcomes differentiated by stakeholders and should include intended, unintended and unanticipated outcomes, and evaluation design should be responsive to each. We stress this because evaluations of impact often underestimate the timescale involved in embedding changes in routine practices and it can be useful to link impact with the underlying theories which connect activities with intended outcomes.

There are of course a variety of practices associated with programme or
policy evaluations along with a well established literature in this area. What this book seeks to do, however, is to ground the discussion in case studies of policy or programme evaluations in four different HE systems in order to foreground contemporary practice, identify contrasts and continuities of practice and draw out ways in which evaluative practices might be strengthened and improved.

**Institutional evaluative practice**

This domain is about the quality, review and reflexive systems in place within institutions, a phenomenon which grew in the UK from the early 1990s, with the onset of an era of formal quality evaluation (Sporn, 1999: 16), and an emphasis on accountability and the responsible use of public money (Tavenas, 2003: 50). These institutional evaluative systems have gained importance over time, as national auditing bodies have increasingly allowed institutional quality assurance processes to inform external audits. The debate about how to balance self-regulation and external evaluation has seen a number of stakeholders taking an interest in how universities evaluate the quality of their provision. Institutions themselves have been keen to show that their internal evaluative practices are effective, to avoid government intervention. While some academics have been ‘co-opted’ to carry out the evaluative work of quality assurance (Kogan and Hanney, 2000: 104), many others have been sceptical about the ethos and the processes of such evaluation:

> In much of the discussion of quality with respect to universities in recent decades, the external focus has predominantly been on the existence of quality assurance mechanisms, the financial management of the university and its research record. However, inside universities and among academics, there has been a strong scepticism about such a focus on quality assurance mechanisms, which are sometimes seen to be at the expense of a focus on the quality of student learning *per se*.  
> (Bowden and Marton, 1998: 215)

Given the strong external steer on quality evaluation, institutional evaluative practices have tended to develop in line with external requirements, whilst also reflecting institutional philosophies and priorities. Unsurprisingly, the context of particular institutions and sections of the sector (e.g. their degree of managerialism) can be detected in how light or heavy the institution’s review systems are. What is in no doubt, however, is that the quality of academic work is now a matter of public interest, and is open to scrutiny not only within the institution but by a number of external stakeholders. The profound impact of ‘the evaluative university’, to paraphrase Neave (1998), has been such that the balance of power in the academic community has changed – with more power to the administration and less to academics, and staffing/financial decisions made on the basis of evaluation data (Kogan,
2002: 88). The manner of evaluating that work is, clearly, of paramount importance to all involved.

At best, the evaluative processes used in making these assessments may be said to have at their heart an impulse to strengthen the practice of the institution by making judgements of worth through various types of reflexive practice. However, a set of justifications embedded in the following extract from a quality support office of a UK University is typical. The discourse is regulatory and concerned with standardisation and control (source remains confidential):

The University is responsible for ensuring that all of its provision (i.e. programmes and contributory modules) is subject to regular and systematic monitoring and review in order to:

- ensure equivalence of academic standards across all programmes offered by the University and compared with the rest of the sector
- ensure an acceptable quality of student learning experience across all programmes offered by the University and compared with the rest of the sector
- ensure equality of opportunity for a diverse range of learners
- ensure that all programmes and contributory modules remain up-to-date and deliver on agreed learning outcomes, relevant in terms of content and pedagogy as referenced by sector standards and benchmarks, and can attract enough students
- inform departmental, faculty and university priorities and strategies with respect to continued enhancement of the quality of the student learning experience.

The University is responsible for ensuring that its approach to regular and systematic monitoring and review will satisfy regular scrutiny by the QAA through institutional audit.

Objectives

1. be able to demonstrate critical self-reflexion on the quality and standards of the programmes and modules for which they are responsible
2. take appropriate action to further enhance the quality of the student learning experience
3. review the extent to which programmes and modules still meet their intended learning outcomes, and revise either or both as appropriate
4. identify factors that are constraining the operation and development of their programmes and modules, and stimulate an appropriate response
5. identify effective practice and demonstrate strategies for sustaining and further disseminating this practice within and beyond the faculty.
It is possible to see this type of evaluative practice as using external quality standards while focusing on the individual teacher or teaching and learning experience. But the framework is not individually-led; it is compulsory, external and centrally devised. From the perspective of the staff member, this is received as an external evaluation of their practice and can be contrasted with the thrust of the practices associated with the domain of self-evaluative practice (next section), where bottom-up initiatives prevail. As Filippakou and Tapper (2008) have suggested, the use of top-down systems to improve or control or both is a contested area.

Evaluation within institutions relates to how well they are performing in teaching or research, but has permeated many other aspects of university life, such as in the human resource area. Probationary systems for staff are nothing new: they have been in place for decades, although to greater or lesser extents in different institutions. But other evaluations of staff have been added, such as appraisal, with assessment of whether targets have been met and whether quality of work has been acceptable. Such systems are now becoming a ubiquitous part of university life: pressures on senior management to make their institutions more accountable have fed down through the system to individual academics (Thorley, 1995: 59). In some senses this has happened organically, as systems have adapted to external expectations and demands, but there has also been a conscious attempt to challenge ‘the culture of individualism’ and increase ‘professionalism’, as expressed by Newby (1999: 269).

There is no shortage of topics which could be covered in this domain. Evaluations are being carried out on how well the institution is doing on diversity, disability and equality; on whether resources are well deployed, giving ‘good value for money’; on how the university compares with its benchmark institutions; on whether staff roles reflect the appropriate salary grade; on the impact of staff training and development; on whether academic courses meet professional body accreditation criteria; on whether PhD students are being supported appropriately; on whether students are satisfied with learning technologies; on whether staff feel bullied at work etc. In this book we will focus on case examples which are central to the work of most academics: institutional and departmental quality review, course validation, and staff appraisal. We will explore the extent to which there are cross-cultural differences in practice with respect to the control, management and use of evaluations.

Self-evaluative practice

This domain offers examples of participatory or bottom-up developmental approaches to evaluation. These evaluations take place when practitioners, either individually or in groups, undertake evaluation to inform their own practice – although the evaluation may have sub-agendas, such as the wish to demonstrate that the individual merits reward and recognition. Such
evaluations are sometimes mirrored in higher level systems. For example, individuals self-evaluate as reflective practitioners; this process has been recognised and adopted by HE accrediting bodies such as the UK Higher Educational Academy (HEA) or the Staff and Educational Development Association (SEDA), to whom individuals can submit their reflections in a specified format in order to gain professional accreditation. Practitioners also evaluate their practice in order to gather data on and improve their professional activities. Action research is commonly used in this way, when practitioners investigate a problem, plan and implement an innovation, and then evaluate its impact before restarting the cycle.

These are a range of activities that closely resemble evaluation and might legitimately be called a form of informal evaluation, remembering of course that informality does not preclude ‘systematisation’. The position adopted in this book is that evaluation can take place within a set of social practices or within an organisation as part of a cultural orientation rather than an over-evaluated or performance ridden raft of controlling measures and systems. An interesting aspect of individually driven evaluation is that it is, by its nature, more likely to be rooted in academic values and norms than other, more top-down evaluative practices. The implication is that the results of such evaluations may be more readily accepted and taken on board by academics. Self-evaluation is perhaps seen at its best when individuals or groups work to build an evaluative culture that has as a central tenet a series of reflective practices. New knowledge is developed as people engage in a process of reflection related to real problems and issues in their own context. As Barab and Plucker (2002: 173) note, there may be little observable teaching in work situations, but a great deal of learning goes on. This learning, increase in knowledge, can be tacit and, therefore, difficult to evaluate. This does not mean that evaluation should be avoided, only that evaluative practices need to be sensitive to the complexity of all the factors involved. In fact, evaluating individual development is a challenge that, in our view, has not yet been met. For example, the efforts of the ‘Rugby Team’ in the UK¹ to work out how to assess the impact of training and development for researchers in universities have confirmed that it is simple to evaluate levels of satisfaction with courses (output) but extremely difficult to evaluate the extent to which such courses affect the way people think and work, and their contributions to the institution (outcomes). Nevertheless, the investigation is a valuable one for our growing understanding of the nature of reflective cultures and ‘what works’.

The use of reflective practice as an evaluative tool is not new in educational and professional development circles in which Schön (1991) and Eraut (2000), for example, have identified the value of developing these processes in terms of organisational health, the adaptive capacity of organisations and in the development of professional identities that understand and are

¹ A UK-wide working group on how to assess the impact of efforts to develop researchers in universities.
sympathetic to collaborative and creative responses to change. That tacit learning is continuous within social practice is axiomatic. The attribution of value and worth through judgements on what is professionally useful, rewarding or what works is part of social practice that can form the basis of such a reflective culture. In our view these processes of judgement are profoundly evaluative – although we do not underestimate the problems associated with using fuzzy concepts. As Brookfield says (1995: 29–30) the reflective practitioner concept seems to mean ‘all things to all people . . . a premature ultimate . . . [which] stops any critical debate dead in its tracks’. For this reason, our focus on evaluative practices is a useful one: we bring our theoretical perspective to bear on observable, practical situations, while keeping in mind these complex contextual considerations.

An interesting and evocative way of depicting such embedded evaluative practices is offered by the work of McCluskey et al. (2008) in which they are described as ‘evaluative moments’. The phrase was coined to cover a set of widely found practices that retained some of the characteristics of evaluative practices as recognised by experts in evaluation, without necessarily being seen as evaluation by those carrying them out (or by those people who are professional evaluators) within a project environment. One of the essential characteristics of evaluation is that it generates new knowledge (or reveals or makes apparent existing or ‘hidden’ knowledge) that can be used by participants or others. In evaluative moments, this generation of knowledge may well not be seen explicitly as knowledge or learning by those undertaking it, just as they may not see it as evaluation. But it serves an evaluative purpose in a working situation and uses some of the tools and processes also used by evaluation. Evaluative moments generally entail a number of related phases that might include the negotiation of work to be done with colleagues concerned, the process design and the creation of tools, the collection or generation of data, the analysis and/or reorganisation of that data and the integration of this (new) knowledge in a process.

This territory is one in which there are unclear boundaries between institutional learning and evaluation, where evaluation can be embedded as a culture of reflection. This can be translated into some highly practical approaches, from relatively informal embedded evaluation through review meetings and reflection, to responsive approaches to quality enhancement and assurance, as well as ideas such as the learning organisation (see Easterby-Smith, Aráujo and Burgoyne, 1999; Senge, 1990). From the point of view of the individual academic, this could involve action research into a particular teaching issue. It could involve the preparation of a portfolio to evaluate and demonstrate their achievements, when they are seeking promotion. It could also relate to how they evaluate their own research project. For groups of academics, evaluative moments tend to occur when the work group (e.g. a course team) decides to review their curriculum, or one aspect of it, such as how they prepare Masters students for research projects, or how they run laboratories or studio teaching. The pressures on course teams and other work groups to be more ‘efficient and effective’ puts constant pressure on
them to review and evaluate what they do, how they do it, and how they could do it better. This book will contribute to their understanding of helpful evaluative practices in these situations.

The structure and approach of this book

The book is divided into four sections, corresponding to the four domains described above:

- National/systemic evaluative practice
- Programmatic evaluative practice
- Institutional evaluative practice
- Self-evaluative practice.

Each section introduces the evaluative practice under consideration and contains cases of evaluative practice both in the UK and other countries. Each section is concluded by a synthesis of what the case studies can tell us about the nature, scope and significance of evaluative practice within that domain. In the final chapter of the book we provide an overview of what seems to us to be the particularly interesting differences and commonalities in the evaluative practices within the four domains. These analytical observations aim to help people across the sector, in different roles (such as managers, educational developers, academics and researchers) to see evaluative practice in a new light and manage their evaluative practices in a more informed way. Our hope is that they will find that, whatever the discipline and locus of their work, the combination of theory and empirical examples we offer will enlighten their evaluative efforts.

In order to help orientate and cohere the book, we will borrow the categories in the RUFDATA framework. RUFDATA (Saunders, 2000) was originally an acronym for the procedural decisions in evaluation planning. We use it here to enable a depiction of a domain of evaluative practice within HE. They stand for the following elements:

- Reasons and purposes
- Uses
- Focus
- Data and evidence
- Audience
- Timing
- Agency.

RUFDATA captures evaluation as a series of knowledge-based practices. These knowledge-based practices form the resources of ‘communities of practice’, i.e. groups of practising evaluators. In essence, it involves a process of reflexive questioning during which key procedural dimensions of an evaluation are addressed leading to an accelerated induction to key aspects of evaluation design. It was initially designed to enable initial planning
of an evaluation to help it ‘get off the ground’. RUFDATA is the acronym given to questions that consolidate this reflexive process. To that extent, the approach is a ‘meta-evaluative’ tool.

In order to provide a practical framework to structure this book, we focus on aspects of evaluative practice that consists of problem solving, procedural characteristics and routine behaviour which might use a community of practice of evaluators as a resource. But, in what sense might an evaluation framework be ‘practical’, i.e. based on practices? In day-to-day usage, the idea of the ‘practical’ has a cluster of meanings: solving everyday problems, having an obvious utility or relevance, emphasis on the procedural rather than the theoretical and uses of a straightforward and accessible kind are all immediately recognisable. Giddens (1984) refers to practical consciousness; that is what someone might know about their work practices including the conditions of their own action, but cannot express ‘discursively’. If Giddens is right, most work-based problem solving will involve practical consciousness. We should be trying to find ways of ‘freeing’ this resource for use by our contributors. However, knowledge, whether it is practical in the way we describe above, discursive or ‘technical’, has been described in different ways.

Reification is of particular interest in the context of this book in that it refers to the way in which processes and practices are consolidated into ‘artefacts’ or things. Wenger talks about experience becoming ‘congealed’ into ‘points of focus around which the negotiation of meaning becomes organised’ [Wenger 1998: 58]. A community of practice produces abstractions, tools, symbols, stories, terms and concepts that reify something of that practice in a congealed form . . . with the term reification I mean to cover a wide range of processes that include making, designing, representing, naming, encoding and describing . . .

[Wenger 1998: 59]

In this sense, RUFDATA is an example of reification derived from the consolidated practices of a group of evaluators providing an interrogatory ‘tool’ for use. In our adaptation of RUFDATA, each domain of evaluative practice can be analysed or presented by a consideration of categories within this frame. However, the RUFDATA framework is used loosely in this book. It is not applied as a structure for the individual contributions although some have used it. Its use has been more in the synthesising sections that bring together some key analytical points from the contributions within a particular domain. However, the lead editors for each section have been selective in their use of the RUFDATA categories, using those that appear to be particularly apt or thought provoking.

The book will use 16 case studies, organised into the analytical domains we have presented. They are authored so that they do not comply rigidly with a particular format. However, they all have a preoccupation with and focus on the depiction of recurrent practices which draw out the limits, possibilities and challenges of the evaluative practices embedded in HE.