International Journal of Management Reviews (2007)
doi: 10.1111/j.1468-2370.2007.00204.x

What have we learned? Themes from the literature on best-practice benchmarking

Graham Francis and Jacky Holloway[1]

The aim of this paper is to review the important themes in the literature on benchmarking, and to assess the contribution to knowledge provided thus far. Four themes from the literature are highlighted: studies of the nature of benchmarking practice; criticisms of benchmarking; evaluating the effectiveness of benchmarking; and the notion of best practice. The authors argue that, while the literature does include a few critical analyses of the practice and impacts of benchmarking, it is dominated by relatively descriptive, and even prescriptive, accounts of the 'realities' of benchmarking. The longer-term (qualitative and quantitative) effects and perspectives of diverse stakeholders on benchmarking are largely neglected. This selective focus has implications for the contribution of benchmarking research to practice, for example, in terms of the role of internal organizational capabilities in analysing the performance of processes and managing communications and change. The authors advocate areas of future research to improve theoretical understanding of benchmarking and thoroughly evaluate its impact within the context of performance management.

Introduction

Twenty years of widespread use have seen benchmarking become an accepted management practice rather than just another management fad. While its practice has occupied the time and energy of many managers, benchmarking has also been the subject of a burgeoning body of literature ranging across the academic to practitioner-oriented spectrum. This literature has evolved from initial 'how to and why' offerings, gradually incorporating greater evaluation and critique.

Given the magnitude, diversity and continual additions to the literature, it is neither practicable nor desirable to offer 'comprehensive' coverage. Instead, the references cited here are a mixture of substantive studies, detailed descriptive pieces and other review papers, selected to illustrate the evolution of the literature, and its limitations.

Reviews by Cox and Thompson (1998), Dorsch and Yasin (1998) and Zairi and Youssef (1995, 1996) traced this growth in benchmarking literature after its first decade and a half. Indeed, new literature reviews have continued to emerge (e.g. Anderson and McAdam 2004; Dattakumar and Jagadeesh 2003; Yasin 2002) along with new journals and books on the subject, perhaps reflecting a transition from management fad to significant management practice. These reviews provide valuable resources for the reader seeking familiarity with the nature of benchmarking and its contribution to performance improvement in public and private sector organizations.

This paper has a more specific aim: to take stock, in terms of contribution to academic knowledge and evidence-informed practice, of four major themes from the literature:

- Theme 1: Studies of the nature of benchmarking practice
- Theme 2: Criticisms of benchmarking
- Theme 3: Evaluating the effectiveness of benchmarking
- Theme 4: The notion of best practice

Some underlying theoretical assumptions may be detected in the benchmarking literature. For example, earlier works treated its impact as universal, suggesting the isomorphic assumptions of institutional theory (DiMaggio and Powell 1991; Frederickson 2002). Later literature is more relativistic, implicitly if not explicitly informed by contingency theory (Donaldson 2001; Elnathan *et al*. 1996; Hill *et al*. 1996; Sitkin *et al*. 1994). Perhaps surprisingly, there is a dearth of critical or postmodern analysis (although Longbottom (2000) tantalizes the reader by referring to both but delivering neither). For example, while many of the employee involvement, organizational change and managerial work aspects of quality management addressed in Wilkinson and Willmott's (1995) edited collection are relevant to benchmarking, there is yet to be a similar scrutiny from the labour process perspective.

Therefore, a literature review organized in terms of explicit theoretical positions on benchmarking would be very brief. Instead, the authors offer a framework of thematic strands. These indicate the strengths and limitations of, and gaps in, the academic literature and, by implication, the research evidence available to practitioners and their advisers. The themes form a spectrum from 'practical and prescriptive' to 'theoretical and critical'. However, the authors argue that the middle ground of evidence-informed and theoretically underpinned practice, associated with 'Mode 2' research (Gibbons *et al*. 1994; Tranfield and Starkey 1998) and favoured by management research policy-makers, is underpopulated. An important aim of this paper is to examine some reasons for this gap in the literature through an exploration of four themes identified in the literature, each theme ending with suggested areas for future research.

This review is not 'systematic' in the terms of Tranfield *et al*. (2003), for reasons that have much in common with the stance taken by Buchanan *et al*. (2005). That is, the wealth and diversity of literature on this subject and its 'ambiguous, multidimensional and contingent' nature (Buchanan *et al*. 2005, 192), makes it impracticable (and probably undesirable) to attempt to draw quantitative conclusions about either the practice of benchmarking or the contribution of academic analyses of it. However, the paper does include a review of efforts to date to develop an understanding of the value of benchmarking to various participants and stakeholders, and points to scope for further work in this area which could have wider relevance in the field of performance management across all sectors of the economy. Following a context-setting section, the body of this paper is therefore structured around the contribution to knowledge reflected by each of the four key themes in the literature.

Best-Practice Benchmarking: Roots and Definitions

From its inception, benchmarking has been a method for identifying aspects of an organization's activity that could be more efficient and/or effective by comparison with other relevant organizations' performance. Data about such 'partner' organizations and their key processes are used (often by multidisciplinary teams) to identify and implement changes, usually to important operational processes.

There are many definitions of best-practice benchmarking that capture most of its essential characteristics. The critical characteristic is the examination of practices (processes). The publication of new management 'text-books' on 'how to do benchmarking' appears to have peaked by the late 1990s (e.g. Camp 1995; Spendolini, 1992; Watson 1993; Zairi 1998; and for the non-profit sector, Letts *et al.* 1999), but such books did establish a broad consensus in terms of what they believed best-practice benchmarking should involve in practice.

Benchmarking originated in the manufacturing sector, but soon found applications in, first, private and then public parts of the service sector. It has its roots in earlier management techniques such as inter-firm comparisons, which have long been available in the form of management information on sets of key parameters pooled by local groups of employers. PIMS (Profit Impact of Marketing Strategies) has been offering 'benchmark' data for member organizations within their own industries since the early 1990s (Merrifield 1994). The first actual labelling of a performance-enhancing activity as 'benchmarking' in Western industry is usually attributed to Xerox in 1983 (Jacobson and Hillkirk 1986). Their portrayal of Xerox as an 'American Samurai' reflects the escalation of interest in Japanese management techniques in the 1980s such as *Kaizen* (continuous performance improvement) (Imai 1987; Yoshikawa *et al.* 1993).

Benchmarking tended to be adopted most readily in what one might term 'blue collar' service organizations, such as engineering maintenance and transport, and high-volume laboratory-based activities such as pathology are quite likely to adopt benchmarking on the back of existing quality assurance or cost analysis systems (see, for example, France *et al.* 2002; Gordon *et al.* 1999; Mathaisel *et al.* 2004). It was slower to take root in the rapidly growing 'white collar' parts of the service sector (financial services, retail, etc.), although other improvement approaches originating in manufacturing appear from the literature to have been even less readily considered (and/or more rapidly dropped) than benchmarking. There are signs of 'benchmarking by stealth' accompanying high levels of adoption by large service businesses of the Balanced Scorecard (Kaplan and Norton 1996; Neely *et al.* 2004).

The following four sections review the literature organized by theme. The first theme, 'Studies of the nature of benchmarking practice', is relatively well populated and, as such, perhaps provides the greatest potential for theory development and, in turn, for more useful contributions to practice.

Theme 1: Studies of the Nature of Benchmarking Practice

The academic literature in this area is largely empirical, encompassing relatively neutral descriptive studies of benchmarking applications, surveys of its prevalence, attempts to categorize benchmarking by inventing new typologies, and the occasional critique particularly in terms of the 'misuse' of benchmarking as a political tool.

A 'Taxonomy of Typologies'?

Many writers on benchmarking attempt to produce a classification framework that adequately describes the array of activities carried out in the name of benchmarking (e.g. Dattakumar and Jagadeesh 2003; Kyrö 2003; Yasin 2002). This may be illustrative of the diverse range of activities taking place in the name of benchmarking and the purposes that benchmarking has been seen to serve. Table 1 shows some of the main typologies which have been developed to describe or classify the characteristics of benchmarking.

Perhaps the most commonly cited typology is Camp's (1995) distinction between *internal*, *competitive*, *functional* and *generic* benchmarking (see Table 1). These four categories have subsequently been expanded upon by other writers: for example, 'best in class benchmarking'

Table 1. Benchmarking typologies

Study	Typology	Definitions
Camp (1995)	*Internal*, *competitive*, *functional* and *generic* benchmarking	*Internal*: comparison among similar operations within one's own organization; *Competitive*: comparison to the best of the direct competitors; *Functional*: comparison of methods to companies with similar processes in the same function outside one's industry; *Generic* process: comparison of work processes to others who have innovative, exemplar work processes. (Adapted from Camp 1995, 16.) Gilson *et al*. (2001) and Peters and Waterman (1982) offer examples of generic benchmarking.
Trosa and Williams (1996)	'*Results* benchmarking' and '*process* benchmarking'	'*Results* benchmarking' is concerned (merely) with the comparative data generated by benchmarking. '*Process* benchmarking' considers how results were achieved, so that the performance gaps identified in the 'results' can be closed by investigating and learning from others' practices.
Bowerman *et al*. (2002)	*Voluntary* or *compulsory*	Considers the differences between public and private sector benchmarking and classifies public sector benchmarking in terms of whether it was *voluntary* or *compulsory.*
Elnathan *et al*. (1996)	'*Unilateral*' and '*co-operative*'	Elnathan *et al*. (1996) distinguish in particular between 'unilateral' (often covert and independent) and 'co-operative' approaches (where information is voluntarily shared between participants). The latter includes '*group*', '*indirect/third party*' and '*database*' forms.
Schofield (1998)	'*Implicit* benchmarking' and '*explicit* benchmarking'	Schofield (1998) distinguishes between '*implicit* benchmarking' (a by-product of information-gathering exercises such as a survey undertaken by a national agency) and '*explicit* benchmarking' (a deliberate and structured process to facilitate comparison and identify directions for change that will lead to improvement).
Jackson and Lund (2000)	'*Implicit* benchmarking' and '*explicit* benchmarking'	Jackson and Lund (2000) also use this typology in their discussion of the emergence of benchmarking in higher education. Focusing on the information-gathering roles of participants.
Murdoch (1997)	*Lateral* benchmarking	Another term used to describe *generic* benchmarking.
CIPFA (1996), Watson (1993)	*International* or *global* benchmarking	The comparison with exemplary performance by organizations overseas; can be competitive, functional or generic.

is used by Spendolini (1992) to emphasize the organization-independent nature of generic benchmarking. 'Industry benchmarking' indicates a wider comparative field than competitive benchmarking (since not all organizations within an industry are a firm's competitors). But Coopers & Lybrand (1994) and Mayle *et al*. (2001), among others, have reported that organizations often start by benchmarking internally, probably at least in part because of the complexities of establishing partnerships, particularly with competitors. An alternative to grappling with some of the more problematic aspects of competitive benchmarking is to adopt generic benchmarking with unlike partners; indeed, Camp (1995) and others point out that truly innovative ideas are more likely to be found by looking at key processes outside one's own industry.

Professional accountancy bodies have been advancing the potential contribution of their members to benchmarking activity for some years. CIPFA (1996, 18) have used Camp's distinctions, while CIMA (1996, 6) include the additional category of *customer benchmarks*, where the customer or market is responsible for determining what is required. However, this is in essence an output measure rather than a process benchmark. Other authors make a similar distinction between *process benchmarking* and *results benchmarking* (Trosa and Williams 1996). Another distinction made, in terms of the nature of benchmarking, is that of *strategic*, *operational* and *functional* or *management* benchmarking. Pryor (1989), Shetty (1993) and Society of Management Accountants of Canada (1995) prefer these categories. Camp's typology is not necessarily incompatible with other typologies, as each could, for example, be carried out at the strategic or operational level.

Wolfram Cox *et al*. (1997) explore in depth the significance of the metaphors implied when the emphasis is placed by organizations on bench*marking* (as competition) and *bench*-marking (as collaboration), seeking to predict the nature of the relationships between benchmarking parties. The variations in commitment to invest time in benchmarking and act on its findings, in part at least, will reflect organizational and contextual characteristics, rather than variability originating in benchmarking itself. For example, Wolfram Cox's and several other typologies provide either implicit or explicit contingency explanations for the differences experienced by organizations, particularly where (inter- or intra-organization) structural variables are involved.

Benchmarking: Public and Private

Empirical academic research into the nature and prevalence of best-practice benchmarking activities lagged somewhat behind its advocacy in the practitioner literature and adoption by organizations. For some organizations, benchmarking has become routine and is seen as an integral part of 'the way we do things here'. However, Holloway *et al*. (1999) reported that many organizations were still actively considering introducing benchmarking or had only recently commenced its introduction. This was supported by surveys in the UK and Europe (Cook and Macauley 1996; Coopers & Lybrand 1994; Coopers & Lybrand Europe 1994; Hinton *et al*., 2000; Partnership Sourcing 1997). More recently, in an international survey Jarrar and Zairi (2000, 2001) report that benchmarking was still being adopted by a growing number of companies with no sign of a downturn: 'The ever-growing literature on benchmarking indicates a wide spread of benchmarking applications across geographical and sectoral borders. It is forecasted that such momentum will grow in the future as benchmarking becomes a "way organizations do business"' (2001, 912).

The concept (if not the actual label) of benchmarking has been familiar to public services for some years. In the UK, for example, independent reports on 'best practice' have long been produced by the Audit Commission and National Audit Office. The actual practice of benchmarking in UK local government has increased with the requirement to use it to demonstrate 'Best Value' (the replacement for compulsory competitive tendering implemented in the 1990s; see Bassam 1997; Bowerman *et al*. 2001; Kite and Davidson 1997; Solovy 1997). Stephens and Bowerman (1997) identify several problems of compulsory benchmarking in relation to local authorities and the implementation of Best Value.

Knowing an organizations's position in a league table does little to enable that organization to understand how better performers achieved their status and hence how to move up the table, perhaps overcoming external obstacles or unequal inputs along the way. (See, for example, the work on school league tables by Goldstein and Spiegelhalter (1996); and the extensive critique of the design, analysis and reporting of performance indicators from the Royal Statistical Society (2003)). Substantial analyses of the uses of benchmarking in the UK National Health Service (e.g. Llewellyn and Northcott 2005; Northcott and Llewellyn 2005), the health sector more widely (Wait 2004), local government (e.g. Bowerman *et al*. 2001) and the wider UK public sector (e.g. Bowerman *et al*. 2002; Magd and Curry 2003) make strong links to the influences of the policy context that may distort the potential for acting on the lessons learned from benchmarking.

The influence of behavioural and contingency theories can be seen in such critiques of 'compulsory' public sector benchmarking policies. However, the underlying logic of benchmarking sometimes seems to have been overlooked by designers of schemes to foster best practice in the commercial sector, too. For example, an underlying premise of best-practice benchmarking is that only by understanding *how* exemplary performers achieve desirable outcomes can one make significant improvements. And by investigating how *dissimilar* organizations perform similar processes, 'learning from contexts outside an organization's usual frame of reference' (Wolfram Cox *et al*. 1997, 286), those improvements can outstrip incremental change found within a single organization or industry. Yet, both public sector league tables and private sector incentive schemes (Fitzgerald and Moon 1996) have continued to encourage intra-sector comparisons and the pursuit of relatively static 'best current practice' targets.

Networking and knowledge sharing come relatively naturally to social enterprises such as charities. For example, Paton (2003) reported on how the voluntary sector has actively embraced many practices that would be termed 'benchmarking' in other sectors, as the desirability of seeking ideas and making comparisons with external organizations has gained legitimacy. However, 'fully fledged' benchmarking following the prescriptions of the likes of Camp is found relatively rarely, although with some good results (Paton 2003; Paton and Payne 1997).

The 'benchmarking club' has been reinvented using web-based technologies (see, for example, Business Performance Improvement Resource 2005; Mann and Grigg 2004; Welch and Mann 2001). Benchmarking activity and research in ICT-based organizations was the focus for a special issue of *Benchmarking: an International Journal* (Alshawi *et al*. 2003), illustrating its relevance alongside the growth of knowledge management. A further example of the modernization of benchmarking is found in its application to tourism destinations (Kozak 2002).

Thus there is a body of research-based literature about the uptake of benchmarking, albeit lagging behind the more instructional/proselytizing literature in practitioner magazines, professional newsletters, etc. However, much of the literature comprises individual case studies rather than surveys. From his survey Longbottom (2000) commented that 'the volume of benchmarking projects is less than expected given the vast amount of literature on the subject over the past few years' (p. 113). Some indications of the reasons why such 'promising practices' are not adopted as readily as one might expect can be found in Leseure *et al*. (2004) but, as their review suggests, many aspects remain under-researched.

Pointers for Future Research on the Nature of Benchmarking Practice

This theme illustrated the diversity of practices that characterizes benchmarking and how the literature attempts to depict it. Unfortunately, the proliferation of typologies is of limited value in helping us to understand the purposes held for benchmarking, partly because it tends to lag behind practice (which in turn has become more diverse, rather than convergent). Indeed, sometimes it is possible to gain the impression that writers are competing to provide the ultimate definition or approaching the field from completely separate viewpoints that are incapable of integration. There is certainly not a single accepted classification system or taxonomy.

As best-practice benchmarking is an established management tool as well as an intellectual concept, and has seen widespread adoption over a number of years, there is the potential in this research theme to examine how the nature of benchmarking develops longitudinally. It may be possible to study how organizations have learned how to learn. For example, taking Camp's (1995) typology, is there a tendency for benchmarking activity to develop with experience and ambition along a 'maturity curve' from *internal* or *functional* to *competitive* and then *generic* benchmarking? If so, does this increase the benefits gained? Or is a more important long-term effect to be found among organizations that move from 'results' to 'process' benchmarking, gaining and embedding capabilities in change management along the way? In-depth longitudinal qualitative case studies may be an appropriate method.

Theme 2: Criticisms of Benchmarking

Initially, the vast majority of benchmarking literature focused merely on 'what is' and 'how to'. Subsequently, an increasing number of authors have begun to question aspects of the fundamental logic of best-practice benchmarking. The literature in this theme incorporates strands relating to design, distinctiveness and deployment of management effort in the broader context of performance management. It links benchmarking to the literature on management fads and fashions, and the adoption of innovative or promising management practices (Leseure *et al*. 2004).

Notes of Caution from the Literature

While much of the practitioner-oriented literature eulogizes about the benefits of benchmarking, cautionary notes have been sounded by such authors as Anderson and McAdam (2004), Cox and Thompson (1998) and Elnathan *et al*. (1996). For example, Cox and Thompson (1998) claim that there are circumstances when benchmarking is inappropriate, as it carries serious strategic risks, such as the inability to control effectively against loss of sensitive data to competitors or the costly failure to implement someone else's 'best practice' effectively. Elnathan *et al*. (1996) suggest that benchmarking suffers from hidden costs such as: (i) the costs of the time and effort needed to co-ordinate the process and the inputs of participants in order to obtain a comparable set of data; and (ii) the quantification of costs which, while significant, may be hard to trace or measure, including costs relating to the cultural change required and costs arising from resistance to change by those involved and affected. Indeed, benchmarking is no more immune to the GIGO (garbage in, garbage out) principle than any other performance management system; statistics can be abused or used incorrectly, as either a cause or effect of gaming and other political behaviours (Maleyeff 2003; Royal Statistical Society 2003). These criticisms may be particularly pertinent to the public sector context.

Although it is often difficult to obtain empirical data about unsuccessful business practices, informants in the authors' own research quite often expressed concerns about problems that had in some cases led to benchmarking being abandoned or not started, in spite of initial plans (Holloway *et al*. 1999), and similar findings were reported by Ansell *et al*. (2003). Doubts were greatest in smaller firms, expressed in terms of perceived cost and time constraints; and some respondents did not know how to find out enough about what benchmarking involves (in spite of the many publications and advisory services available). Benchmarking efforts that were considered to have been reasonably successful were characterized by basic organizational processes such as having someone with previous benchmarking experience on the team, good interdisciplinary working, and top management commitment and realistic resources, rather than slavish adherence to the 'instructions' such as Spendolini's 'five steps' (Spendolini 1992).

The type of benchmarking undertaken may be constrained by perceptions of risks or costs. Choosing similar organizations means they are more likely to be your competitors. Under such circumstances, the exchange of performance indicators may be a straightforward reciprocal arrangement. The exchange of *ideas*, however, is much more likely to be inhibited by considerations of commercial advantage (and this applies increasingly in the public sector too). So, returning to Camp's (1995) *internal*, *competitive*, *functional* and *generic* typology, a combination of similar comparator organizations and metrics or performance indicator data may tend to squeeze out the possibility of generic benchmarking and thus preclude many organizations from some of the greatest potential performance gains that could be made.

Wolfram Cox *et al*. (1997) discuss the tensions between competition and collaboration in the context of various forms of benchmarking. They describe benchmarking as a mixed metaphor, in which 'theory is insufficiently developed to explain the differences between effective and ineffective efforts' (Wolfram Cox *et al*. 1997, 286). There is an apparent lack of enthusiasm for bringing description and prescription together in search of theoretical explanations, combined with a reluctance to undertake empirical testing as noted by Elnathan *et al*. (1996, 53): 'Nowhere in the literature is there a comprehensive, comparative study of which information sharing method is most effective under a specific set of conditions.' Systematic evaluations of benchmarking programmes using theoretical models and concepts are few and far between in the literature. This is a gap which it would be more appropriate to expect to see filled by academic research than by consultants or professional or trade bodies.

The accusation that best-practice benchmarking is atheoretical, or at least under-theorized, can be extended to many management fads or fashions (see, for example, Holloway 2002; Walgenbach and Hegele 2001). However, this does not render them unworthy of academic study, as considered below.

Benchmarking: Fad or Fiction?

The sustained popularity of benchmarking (in terms of its uptake) may well be due to the fact that it subsumes such a broad range of activities, many of which have been common management practice in a number of guises for many years. However, the slowing of the tide in the 'airport bookstand' genre of benchmarking books could signal a loss of mass appeal, at least as perceived by publishers. To hard-pressed managers, is the introduction of benchmarking just another management 'fad' or 'fashion', with little to indicate why they should go out of their way to adopt it? As such, it may be rejected before being fully considered, particularly if the next fashion is already being heavily promoted. Longbottom (2000), for example, catalogues the assault on benchmarking made by proponents of business process re-engineering in the early 1990s.

However, 'management fads' or fashions are not necessarily as insubstantial as fads in home decoration or fashions in clothing. They can have a serious part to play in changing organizational performance. They reflect the 'emergent collective preferences' of managers for new techniques which will help them to 'respond to organizational performance gaps opened up by real technical and economic environmental changes', serving as a technical learning process for managers – provided that they appear both rational and progressive (Abrahamson 1996, 255; and see Huczynski 1992). Whether such fashions have intrinsic value and validity when they emerge from their originators is less important than the characteristics they develop as they are interpreted by trainers or consultants and then by managers themselves.

Abrahamson (1996) pointed out that the consumption of management fashions can be influenced by the demands of their ‘consumers’, who may be insufficiently demanding of the suppliers of management fashion rhetorics (i.e. management academics) in terms of requiring causal explanation or theoretical underpinnings. There are very few systematic evaluations of benchmarking interventions – or indeed other performance improvement systems – to be found in the literature. A pattern of management fashions being promulgated, adopted and discarded without due attention being paid (either by practitioners or academics) to how far their potential impact was realized is a missed opportunity for both groups. The complex range of influences on the adoption of ‘promising practices’ designed to improve organizational performance illustrated in Leseure *et al*.’s (2004) systematic review provides considerable food for thought.

Benchmarking as a Performance Management Tool

Benchmarking is often only one of several performance management tools used by managers (especially in larger organizations) either in series or in parallel (Berry *et al.* 2005; Holloway *et al*. 1999). It is perhaps unfair to accuse managers of merely jumping on the latest bandwagon, however, as benchmarking is readily integrated within an overall system for performance management, playing to its strengths and recognizing its limitations. This is supported by the findings of the Best Practice Club’s survey of member organizations (Chase 1997), where benchmarking was being used to improve the value of products and services to the customer, being most effective where total quality management (TQM) and self-assessment using the European Foundation for Quality Management’s Business Excellence Model (EFQM 1993) were already established.

Pointers for Future Critical Research on Benchmarking

Benchmarking may have lost some of its attractiveness as a research topic as newer approaches to performance management and improvement have become established but, so long as benchmarking is still in common practice, there is scope for synergy within the wider field of performance management research. For example, conceptual and empirical studies of the use (and misuse) of the Balanced Scorecard (Kaplan and Norton 1996) are beginning to emerge (Bourguignon *et al.* 2004; Neely *et al*. 2004; Nørreklit 2000; Nørreklit *et al.* forthcoming). The simultaneous presence across a wide range of sectors and countries, of several widely used and well-documented approaches to cross-functional performance management should still give multidisciplinary research groups much to investigate.

The relative shortage of academic studies of benchmarking from a ‘radical critical’ perspective has already been noted; indeed, theoretical examinations from more mainstream perspectives also remain scarce. Most of the ‘criticism’ of benchmarking has come from consultants or practitioners and has been in terms of its technical limitations or assumed impact on the financial bottom line. Ultimately, benchmarking is a tool for performance management primarily serving the interests of employers and shareholders, which directly affects peoples’ working practices and could conceivably lead to reductions in jobs. It should therefore logically attract attention across the spectrum from neo-institutionalists, agency theorists and adherents of transaction cost economics and resource-based strategies, through relativist contingency theorists and Foucauldians, to critical realists, Habermasians and Marxists. Yet, standard texts on critical management thinking (Alvesson and Willmott 1996; Barry *et al*. 2000; Clegg and Hardy 1999) have little to say about performance management in general, and even less on benchmarking. So there is a major opportunity for theoretical analysis waiting.

As interest in corporate governance and stakeholder perspectives grows, benchmarking may provide a rich territory for process-focused studies. For example, how is benchmarking affected by legislation and corporate policies on information management and disclosure? The critique that benchmarking is just another management fashion has generally been indistinguishable from critiques of other performance management systems. The literature tends not to include in-depth studies of the processes by which interpersonal relationships, power structures, etc. are affected by benchmarking. Nor do writers tend to take a broader look at how organizations manage their performance more generally, particularly whether those that are 'good' at benchmarking also 'good' at other approaches to performance management. If this is the case, does this reflect consistent technical competence, or a sound choice of approach taking into account the organization's culture, history, leadership style, key personalities and so on? If the latter contingency explanation holds, there are implications for the extent to which 'best practice' or even 'good practice' can be generalized about.

Benchmarking seems to be a field in which the academic community has not by and large provided students or practitioners with 'scholarly rhetorics' (Abrahamson 1996) which are sufficiently informed by empirical research to provide a sound basis for action. Nor have they applied theories to this area of management practice so as to gain a better understanding of its dynamics, or to improve on those theories. The authors advocate research in all these areas.

Theme 3: Evaluating the Effectiveness of Benchmarking

As a sub-set of the literature on evaluation of the impact of management interventions (which is in any event somewhat patchy), published studies of the impact of benchmarking activity on organizational outcomes are particularly scarce. Furthermore, evaluation is rarely their main focus, and (in common with other performance management approaches) where impact is explored, it is more likely to be in terms of the dimension of efficiency rather than effectiveness or long-term impact. The literature can be classified in terms of intended audience: academic or practitioner.

Academic Evaluations of the Outcomes of Benchmarking

Where it does exist, evaluative benchmarking literature has tended to concentrate on identifying *factors* associated with 'successful' benchmarking (as indeed does much of contingency theory). Unfortunately, factor models do not explain *how* outcomes occur; rather they associate a level of outcome (or output) with a level of predictor (input), inferring the causal linkages between the two. By contrast, *process* models aim to identify conditions that are seen as necessary but insufficient to cause an outcome. Fundamental to process models is an understanding of how these necessary conditions are combined, as a sequence of events, which provide an insight as to how a discrete outcome occurs, whenever it does occur. This would allow research to focus on the dynamics of social change explaining how and why benchmarking may have beneficial results.

Benchmarking has at its core a concern to control processes so that they produce optimal benefits and, in this respect, cybernetic control models may have a part to play (see, for example, Dent and Ezzamel 1987), but it also needs to operate within a context in which 'optimal' performance is under constant change. Thus, both double-loop control, and double-loop learning will have a part to play in the effective use of this approach (see, for example, Argyris and Schön 1996; Sitkin *et al*. 1994).

Insights into benchmarking as a process can be found within the literature of management disciplines such as accounting and human resource management, and domains such as strategy and organizational behaviour. (Leseure *et al*.'s (2004) literature review of the adoption of promising practices draws on these and other disciplines.) Elnathan *et al*. (1996) find that benchmarking can aid managers both to benchmark results and to drive improvement through such techniques as Activity Based Costing (ABC). Zairi and Leonard (1994) examine how an organization can learn the process of ABC through benchmarking another organization, while Fifer (1989) shows how one could benchmark the value chain. Benchmarking can also be used as an input to a target costing approach, and in changing manufacturing practices (Young and Selto 1991).

Elnathan *et al*. (1996) emphasize the desirability of applying a contingency approach in the assessment of the success of benchmarking activities. They propose a research framework for incorporating *antecedent*, *contextual* and *outcome* variables. One antecedent variable in their framework is senior management support, which is also an important factor in the work of Hill *et al*. (1996), particularly where managers are embarking on benchmarking for the first time and are strongly influenced by their own perceptions of senior management's commitment. Elnathan *et al*.'s (1996) contextual variables include several characteristics of partnerships such as number of partners and degree of trust; and Mannering (1996) identifies the desirability of partner organizations having shared values. Elnathan and Kim (1995) model the relationship between the formation of partnerships and potential benefits and costs of benchmarking (with consequent implications for the use of third-party benchmarking organizations as well as the formation of direct partnerships). Elnathan and Kim (1995, 362) call for a greater understanding of 'benchmarking practices and organizations in order to produce a clearer picture of what factors determine firms' benchmarking benefits and costs and in turn affect their benchmarking decisions'. Their conclusions add weight to the growing case for systemic research incorporating multiple stakeholder interests.

So How Can Managers Evaluate Benchmarking?

What is the value of benchmarking, both as a practical management tool (in essence 'Does it work?') and as a designed system for controlling organizational outputs ('How does it work?')? Undertaking benchmarking activities, and acting on the findings, involves not-inconsiderable resource commitments, so these remain key questions for managers. However, the balance of benefits relative to the associated costs appears to have been weighed up comparatively rarely, and judgments are based mainly on anecdotal evidence. Benchmarking is recommended as 'a good thing' by management trainers and educators, practitioner journals, consultants, statutory and professional bodies; indeed, it is an explicit requirement for many organizations. Yet where it has been undertaken, systematic evaluation provides evidence that, in some circumstances, both financial and other costs may outweigh the benefits (Lincoln and Price 1996; Sheridan 1993). Assessing its overall impact requires careful attention to both sides of the equation, in which the benefits may be less tangible than the costs.

A large-scale survey by Coopers & Lybrand (1994) found that 75% of large organizations viewed benchmarking projects as successful because they: set meaningful and realistic targets; improve productivity; provide new insights; give an early warning of competitive disadvantage; and motivate staff by showing what is possible. Identifying the benefits perceived by benchmarking practitioners was also a key aim of the empirical research carried out by the present authors (Holloway *et al.* 1999; Mayle *et al.* 2001), while acknowledging the difficulty of separating those 'outcome variables' that are attributable to benchmarking from those that might have happened anyway. In evaluating the cost-effectiveness of benchmarking, the question also arises of whether such benefits could be achieved more cost-effectively by other approaches (although the same issue of a lack of good data on cost-effectiveness applies to many performance improvement activities).

Is the external comparison element, claimed to be one of benchmarking's distinctive features, necessary other than as a stimuli to consider the internal processes of one's own organization? Is there an element of 'old wine in new bottles', with little really to commend benchmarking over well-designed comparative performance measurement? Evaluation studies of other performance management systems that use external standards or comparators, such as ISO9000, Investor in People and the EFQM Excellence Model, have concluded that their impact is at best uneven. However, controlled evaluations are rarely possible and organizations in some sectors may have no option but to adopt prescribed approaches.

Pointers for Future Research on the Effectiveness of Benchmarking

The methodological difficulties of measuring the impact of benchmarking interventions from a research point of view are as great as they are for managers. However, opportunities do sometimes arise for action research or 'before and after' studies that can reveal both intended and unintended (but not necessarily undesirable) effects of benchmarking programmes. (See also, for example, the quasi-experimental approach reported by Neely *et al.* 2004.) It may also be possible to identify impact indicators (both qualitative and quantitative), and to use techniques such as ABC to capture in detail some of the financial and money-equivalent costs of undertaking activities in the name of benchmarking. Particularly useful would be research to identify which elements of the benchmarking process actually yield the benefits (Elnathan *et al.* 1996; Holloway *et al.* 1999). The possibility of a 'Hawthorne effect' is strong, especially if the importance to the organization of relatively humble processes is recognized, and shop-floor workers are empowered to implement improvements based on their own experiences. But, particularly where records exist of the impacts of previous performance improvement activities, it may be possible to suggest the presence of a distinctive benchmarking effect.

A wide variety of factors may be critical to enhancing the success of benchmarking (see, for example, Cox and Thompson 1998; Elnathan *et al.* 1996; Holloway *et al.* 1999), each of which is perhaps necessary for successful benchmarking but not sufficient in itself. Antecedent conditions do not necessarily lead to the desired outcomes without an understanding of the *context* in which benchmarking takes place, from the tacit knowledge brought to it by participants to the wider organizational and environmental setting (see Elnathan *et al.* 1996). In accepting the importance of the process, it is possible to move beyond the notion of cause and effect implied by much of the current literature.

Another diverse and expanding body of literature that could usefully be linked to benchmarking research addresses networks as organizational forms from a range of theoretical perspectives and social science disciplines (see, for example, Borgatti and Foster 2003; Hoang and Antoncic 2003; Kilduff and Tsai 2003; Uzzi 1996; Zaheer and Bell 2005). Particularly relevant strands concern partnerships and risks surrounding partner selection, information flows and asymmetries, the role of networks in fostering innovation, and the practical and intellectual challenges of evaluating the performance of networks. Although the novelty of the network form, and its status as supplanter of market and hierarchical forms, may have been overstated, evidence about the processes involved in network building and maintenance and the role of trust, boundary-spanners and communications could offer new insights into benchmarking effectiveness.

Theme 4: The Notion of Best Practice

There is a tendency for 'best-practice benchmarking' and 'benchmarking' to be used interchangeably. But by identifying this as a distinct theme in the literature, we recognized that the notion of 'best practice' is being questioned separately from the activity of benchmarking itself.

Is 'Best Practice' the Enemy of 'Good Practice'?

There is an implicit assumption in much of the benchmarking literature that organizations want to be ‘the best’, in terms of ‘best in class’, highest return to shareholders, leading brand in the market and so on. Such rhetoric is not necessarily consistent with practice and, in any event, we know from the world of strategy that ‘best’ is context-dependent, particularly in the global economy. Many managers are more pragmatic in terms of what can be achieved.

The quest for ultimate performance is even more contested in the public sector, where delivering appropriate local services consistent with national policy is the main driver rather than being ‘better’ than other providers (see Bowerman *et al*. 2002; Llewellyn and Northcott 2005; Northcott and Llewellyn 2005). Llewellyn and Northcott (2005) demonstrate that an unintended consequence of UK government attempts to raise standards through hospital cost benchmarking has been to reinforce a tendency for hospital performance to cluster around the mean rather than become more like the ‘best’.

Indeed, the concept of continuous improvement that has driven the quality movement since the 1980s runs counter to the possibility of arriving at an absolute standard of ‘the best’ performance. The sharing of good practice, based on evidence of cause-and-effect, for the benefit of stakeholders (be they public service users, consumers of goods, shareholders or employees) is becoming the order of the day.

If the notion at the corporate level that to be ‘the best’ is no longer the only strategy, this has long been the reality at the managerial level. Competing demands and changing policies lead managers to make frequent trade-offs, resulting in sub-optimization of individual processes, even though overall performance may be comparatively strong. Indeed, as March and Simon pointed out in 1958, managers may not be able to optimize even if they were willing to try, given the complexity and volume of information they are generally confronted with; thus, satisficing is the norm at all organizational levels.

So where does this leave ‘best practice’ benchmarking? Have the first two words been subtly abandoned, in the literature as well as in practice? In terms of how the benefits are perceived by the organization and individual participants in benchmarking teams, it is important to consider their initial expectations and motivations to benchmark (see, for example, Hill *et al*. 1996). Holloway *et al*. (1999) found stakeholder differences in expectations and perceptions of how successful benchmarking has been. This concern is reflected in Yarrow *et al*. (2004), who report on the impact of the first 10 years of the PROBE Partnership – an international group with members in academia, manufacturing industry, employers’ organizations and consultancy dedicated to promoting and evaluating the use of a suite of benchmarking tools. They conclude:

> Above all, benchmarking is not all about the technicalities of advanced ‘best practices’ and statistical comparisons. It is also (and mostly) about people, learning and improvement … The next 10 years may be less about the state of the art of best practices as such, and more about applying the social sciences to stimulate and support diffusion and adoption of those practices. (Yarrow *et al*. 2004, 838)

Several theoretically informed critiques do exist of the notion of ‘best practice’ as opposed to ‘good’ or ‘better’ practice (see, for example, Huxham and Beech 2000). From an institutional theory perspective (DiMaggio and Powell 1991; Frederickson 2002; Meyer and Rowan 1991), the pursuit of ‘best practice’ can be seen as evidence of the tendency for organizations to become more similar (isomorphism); and to legitimize behaviours through ‘myth and ceremony’, that were once regarded as innovative but eventually become dysfunctional or constraining.

Best-practice studies may well pay no attention to generalizability, ignoring the organizational contexts, resources deployed and internal influences such as the impact of individuals’ personalities on team effectiveness. Indeed, the criteria by which exemplary practice has been judged often remain unreported, so organizations looking for benchmarking partners may find it difficult to judge the true relevance of the achievements claimed by some organizations, including ones on the 'industrial tourism' circuits promoted by advisory services. (The same criticism can less readily be levelled at winners of formal competitions such as the EFQM Excellence or US Malcolm Baldrige awards, which are based on detailed criteria and use trained independent assessors.)

In short, 'best practice' implies the existence of a 'single best way' rather than the 'most appropriate way contingent on the circumstances at the current time', an implication that is hard to support from most theoretical perspectives (see Leseure *et al*. (2004) Box 6: 'The myth of universally applicable best practices', 68). Even organizations with very similar structures and objectives will have their own unique cultures and histories, priorities and needs, all of which will affect the impact of performance improvement approaches such as benchmarking (see Hope and Mühlemann 2001; Mayle *et al*. 2001). Cox and Thompson (1998) argue that benchmarking only allows organizations to catch up with the leader, not to 'leap-frog' them, and as such 'market leadership and sustainable competitive advantage cannot be gained from benchmarking; it can only ever offer second best' (Cox and Thompson, 1998, 11).

Just as doubts can reasonably be cast on the notion that the 'best' operational processes which are the subject of benchmarking exercises can be replicated in other contexts, there is also no proven single 'best way' to do benchmarking. This is not to deny the real potential for valuable lessons to be learned by sharing effective practices between and within organizations, but rather to call for a serious examination of benchmarking's success factors and a willingness to challenge the notion of 'best' practice. As the saying goes, 'the best may be the enemy of the good'.

Pointers for Future Research on the Notion of 'Best Practice'

In terms of the validity of the broad concept of 'best practice', beneficial future research is likely to be more theoretical, philosophical and ontological than empirical. It may be a forum in which relativism confronts realism, postmodernism claims victory over modernism, and the concept of paradigms in social science is revisited.

More useful here and now is to consider the issue of whether it is appropriate to establish 'good', 'better' or 'best' ways to undertake benchmarking itself. If methodologically sound approaches to such evaluation are established, they may well be transferable to other areas of performance management. At its most basic, benchmarking can be conceived as a means of sharing good practice, which has been validated in certain organizational contexts and which logic suggests may bring benefits elsewhere. However, while there are certain factors which are necessary for benchmarking to contribute to any sort of 'results', there are also many process and contextual features which are associated with effective outcomes, not least in terms of partner selection. Thus, it has been possible so far to identify associations (rather than causal relationships), but not to identify specific organizational characteristics or elements in benchmarking itself which are in themselves sufficient to ensure that the desired results are delivered and 'best practice' in benchmarking is attained. Whether the search for this kind of deterministic model is desirable is a matter of opinion rather than fact.

Some Concluding Remarks

In terms of studies into the nature of benchmarking practices (Theme 1), the tendency for writers to develop new typologies to describe benchmarking practices leads us to conclude that, while such efforts indicate diversity of activity in the name of benchmarking, they have not perhaps told us enough about its various outcomes particularly in the longer term.

Criticisms of benchmarking (Theme 2) fall into two main camps: those that are relatively practical, related to the details of benchmarking practice; and those verging on the theoretical, linked to performance management more generally. Writers have raised notes of caution about the assumptions held by some academics and practitioners. Those critiques have come mainly from writers who broadly support the existence of management-led performance improvement objectives, rather than the more 'radical critical' traditions that focused on, for example, BPR and TQM in the 1990s. Initial reflections on why this may be so prompt the conclusion that the benchmarking literature lacks theoretical claims that might attract the attention of 'radical critical' writers. In terms of longevity, again compared to BPR and TQM where applications outside manufacturing continually attracted scepticism or hostility until they seem to have gradually faded away, it is almost as if benchmarking has got in 'under the radar' into new sectors and industries, in the process of which its methods are slowly evolving.

It is hard for benchmarking to escape from the 'management fad' label, but our assertion is that this is an unhelpful characterization. Full-blown 'best-practice benchmarking' does seem to have faded from the limelight, but the principle of comparing data in the pursuit of performance improvement has become an increasingly normal management activity.

Literature linked to Theme 3, evaluating the effectiveness of benchmarking, encapsulates many generic issues about evaluation of organizational and managerial interventions. There have been limited efforts to evaluate the approach, both as an academic project and from a practical management perspective. The authors would encourage much more explicit and deliberate evaluation studies of benchmarking and other performance management systems. Even if benchmarking gradually fades from view, there will be new systems coming along to take its place, in spite of the relative lack of knowledge about their intended and unintended (positive and negative) consequences for organizational performance and learning. Such studies of benchmarking might have wider relevance, and a range of management and social science disciplines have important contributions to make to this quest.

The final theme on the very notion of 'best practice' is particularly diffuse but likely to grow in importance alongside contemporary academic and practitioner interests such as managing competing stakeholder needs, evidence-based practice and establishing new modes of corporate governance. Future literature within this theme is likely to cross many discipline boundaries as well as theoretical ones.

This review paper has explored the literature of best-practice benchmarking through four main themes and considered their implications for future empirical and theoretical research. We believe this thematic handling of the vast literature provides some 'way-marks' for readers wanting directions while steering a course through the ocean of literature that contains some hazardous prescriptions and poorly charted explanations. While the references cited include exemplars of substantive studies, detailed descriptive pieces and other review papers, theory-building is still often shallow.

So what are the ramifications of such diversity and incompleteness for benchmarking research? To attempt a synthesis of messages linking the four themes, we would identify a need to understand better the nature and influences of group processes, organizational culture and management style. This raises a number of issues, most notably around organizational knowledge and learning, knowledge elicitation, communication, team building and dynamics. Clearly, although a lot is now known about benchmarking, there is a lot more still to be learned.

Acknowledgements

The authors wish to acknowledge the contributions of Jackie Fry, Matthew Hinton, David Mayle and the late Mary Bowerman to this research.

Note

1 Corresponding author: Dr Jacky Holloway, e-mail: j.a.holloway@open.ac.uk; Tel: +44 (0)1908 655833; Fax: +44 (0)1908 655898.

References

Abrahamson, E. (1996). Management fashion. *Academy of Management Review*, **21**, 254–285.

Alshawi, S., Irani, Z. and Baldwin, L. (2003). Editorial. Benchmarking: information and communication technologies. *Benchmarking: an International Journal*, **10**, 312–324.

Alvesson, M. and Willmott, H. (1996). *Making Sense of Management*. London: Sage.

Anderson, K. and McAdam, R. (2004). A critique of benchmarking and performance measurement: lead or lag? *Benchmarking: an International Journal*, **11**, 465–483.

Ansell, J., Moles, P. and Smart, A. (2003). Does benchmarking help? *International Transactions in Operational Research*, **10**, 339–350.

Argyris, C. and Schön, D.A. (1996). *Organisational Learning II. Theory, Method, and Practice*. Reading, MA: Addison-Wesley.

Barry, J., Chandler, J., Clark, H., Johnston, R. and Needle, D. (2000). *Organization and Management: a Critical Text*. London: Thomson Learning.

Bassam, S. (1997). Best value will help create a new service delivery ethic. *Local Government Chronicle*, 6 June, 8–9.

Berry, A.J., Broadbent, J. and Otley, D. (2005). *Management Control: Theories, Issues and Performance*, 2nd edition. Basingstoke: Palgrave Macmillan.

Borgatti, S.P. and Foster, P.C. (2003). The network paradigm in organizational research: a review and typology. *Journal of Management*, **29**, 991–1013.

Bourguignon, A., Malleret, V. and Nørreklit, H. (2004). The American Balanced Scorecard versus French Tableau de Bord: the ideological dimension. *Management Accounting Research*, **15**, 107–134.

Bowerman, M., Ball, A., and Francis, G.A.J. (2001). Benchmarking as a tool for the modernisation of local government. *Financial Accountability and Management*, **17**, 321–330.

Bowerman, M., Francis, G.A.J., Ball, A. and Fry, J. (2002). The evolution of benchmarking in UK local authorities. *Benchmarking: an International Journal*, **9**, 429–449.

Buchanan, D., Fitzgerald, L., Ketley, D., Gollop, R., Jones, J.L., Saint Lamont, S., Neath, A. and Whitby, E. (2005). No going back: a review of the literature on sustaining organizational change. *International Journal of Management Reviews*, **7**, 189–205.

Business Performance Improvement Resource (2005). http://www.bpir.com/ (accessed 15 February 2005).

Camp, R.C. (1995). *Business Process Benchmarking: Finding and Implementing Best Practices*. Wisconsin: ASQC Quality Press.

Chase, R.L. (1997). Best practice survey. *Organisational Excellence*, June, Bulletin, 3.

CIMA (1996). *Benchmarking*. London: Chartered Institute of Management Accountants.

CIPFA (1996). *Benchmarking to Improve Performance*. London: Chartered Institute of Public Finance and Accountancy.

Clegg, S.R. and Hardy, C. (1999). *Studying Organization: Theory and Method*. London: Sage.

Cook, S. and Macauley, L. (1996). Benchmarking customer service. *Customer Service Management*, December, 44–47.

Coopers & Lybrand (1994). *Survey of Benchmarking in the UK*. London: Coopers & Lybrand and CBI National Manufacturing Council.

Coopers & Lybrand Europe (1994). *Survey of Benchmarking in Europe*. London: Coopers & Lybrand.

Cox, A. and Thompson, I. (1998). On the appropriateness of benchmarking. *Journal of General Management*, **23**, 1–20.

Dattakumar, R. and Jagadeesh, R. (2003). A review of literature on benchmarking. *Benchmarking: an International Journal*, **10**, 176–209.

Dent, J. and Ezzamel, M. (1987). Organisational control and management accounting. In Ezzamel, M. and Hart, H. (eds), *Advanced Management Accounting: An Organisational Emphasis*. London: Cassell.

DiMaggio, P. and Powell, W.W. (1991). The Iron Cage Revisited: institutional isomorphism and collective rationality in organizational fields. In Powell, W.W. and DiMaggio, P. (eds), *The New Institutionalism in Organisational Analysis*. London: University of Chicago Press.

Donaldson, L. (2001). *The Contingency Theory of Organizations*. Thousand Oaks, CA: Sage.

Dorsch, J.J. and Yasin, M.M. (1998). A framework for benchmarking in the public sector: literature review and directions for future research. *International Journal of Public Sector Management*, **11**, 91–115.

EFQM (1993). *Total Quality Management – the European Model for Self-appraisal*. Brussels: European Foundation for Quality Management.

Elnathan, D. and Kim, O. (1995). Partner selection and group formation in co-operative benchmarking. *Journal of Accounting and Economics*, **19**, 345–364.

Elnathan, D., Lin, T.W. and Young, S.M. (1996). Benchmarking and management accounting: a

framework for research. *Journal of Management Accounting Research*, **8**, 37–54.
Fifer, R.M. (1989). Cost benchmarking functions in the value chain. *Planning Review*, May–June, 18–27.
Fitzgerald, L. and Moon, P. (1996). *Performance Measurement in Service Industries: Making It Work*. London: CIMA.
France, N., Francis, G.A.J., Lawrence, S.R. and Sacks, S. (2002). Costing, counting and comparability. *Pacific Accounting Review*, **14**, 1–21.
Frederickson, H.G. (2002). Easy innovation and the iron cage. Research workshop on the empirical study of governance, management and performance, Texas A&M University, 21–23 February.
Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P. and Trow, M. (1994). *The New Production of Knowledge*. London: Sage.
Gilson, C., Pratt, M., Roberts, K. and Weymes, C. (2001). *Peak Performance – Business Lessons from the World's Top Sports Organizations*. London: Harper Collins.
Goldstein, H. and Spiegelhalter, D.J. (1996). League tables and their limitations: statistical issues in comparisons of institutional performance. *Journal of the Royal Statistical Society A*, **159**, 385–443.
Gordon, M., Holmes, S., McGrath, K. and Neil, A. (1999). Benchmarking pathology services: implementing a longitudinal study. *Pathology*, **31**, 133–141.
Hill, M., Mann, L. and Wearing, A.J. (1996). The effects of attitude, subjective norm and self-efficacy on intention to benchmark: a comparison between managers with experience and no experience in benchmarking. *Journal of Organizational Behaviour*, **17**, 313–327.
Hinton, C.M., Francis, G.A. and Holloway, J.A. (2000). Best practice benchmarking in the UK. *Benchmarking: an International Journal*, **7**, 52–61.
Hoang, H. and Antoncic, B. (2003). Network-based research in entrepreneurship. a critical review. *Journal of Business Venturing*, **18**, 165–187.
Holloway, J.A. (2002). Investigating the impact of performance measurement. *International Journal of Business Performance Management*, **3**, 167–180.
Holloway, J.A., Hinton, C.M., Francis, G.A. and Mayle, D.T. (1999). *Identifying Best Practice in Benchmarking*. CIMA Research Monograph. London: Chartered Institute of Management Accountants.
Hope, C.A. and Mühlemann, A.P. (2001). The impact of culture on best-practice production/operations management. *International Journal of Management Reviews*, **3**, 199–217.
Huczynski, A.A. (1992). *Management Gurus: What Makes Them and How to Become One*. London: Routledge.
Huxham, C. and Beech, N. (2000). Tensions in theorising management practice and practising management theory. Working paper presented at British Academy of Management Conference, Edinburgh, 13–15 September.
Imai, M. (1986). *Kaizen: the Key to Japan's Competitive Success*. New York: Random House.
Jackson, N. and Lund, H. (2000). *Benchmarking for Higher Education*. Buckingham: Open University Press.
Jacobson, G. and Hillkirk, J. (1986). *Xerox, American Samurai*. New York, Macmillan.
Jarrar, Y.F. and Zairi, M. (2000). Internal transfer of best practice for performance excellence: a global survey. *Benchmarking: an International Journal*, **7**, 239–246.
Jarrar, Y.F. and Zairi, M. (2001). Future trends in benchmarking for competitive advantage: a global survey. *Total Quality Management*, **12**, 906–912.
Kaplan, R.S. and Norton, D.P. (1996). *The Balanced Scorecard*. Boston, MA: Harvard Business School Press.
Kilduff, M. and Tsai, W. (2003). *Social Networks and Organizations*. London: Sage.
Kite, J. and Davidson, V. (1997). It sounds great, but what does best value mean? *Local Government Chronicle*, 20 June, 20.
Kozak, M. (2002). Destination benchmarking. *Journal of Tourism Research*, **29**, 497–519.
Kyrö, P. (2003). Revising the concept and forms of benchmarking. *Benchmarking: an International Journal*, **10**, 210–225.
Leseure, M.J., Bauer, J., Birdi, K., Neely, A. and Denyer, D. (2004). Adoption of promising practices: a systematic review of the evidence. *International Journal of Management Reviews*, **5/6**, 169–190.
Letts, C., Ryan, W.P. and Grossman, A. (1999). *High Performance Nonprofit Organizations: Managing Upstream for Greater Impact*. New York: Wiley.
Lincoln, S. and Price, A. (1996). What benchmarking books don't tell you. *Quality Progress*, **29**, 33–36.
Llewellyn, S. and Northcott, D. (2005). The average hospital. *Accounting, Organizations and Society*, **30**, 555–583.
Longbottom, D. (2000). Benchmarking in the UK: an empirical study of practitioners and academics. *Benchmarking: an International Journal*, **7**, 98–117.
Magd, H. and Curry, A. (2003). Benchmarking: achieving best value in public-sector organisations.

Benchmarking: an International Journal, **10**, 261–286.
Malayeff, J. (2003). Benchmarking performance indices: pitfalls and solutions. *Benchmarking: an International Journal*, **10**, 9–28.
Mann, R. and Grigg, N. (2004). Helping the Kiwi to fly: creating world-class organizations in New Zealand through a benchmarking initiative. *Total Quality Management*, **15**, 707–718.
Mannering, K. (1996). Benchmarking and the thin blue line. *Professional Manager*, September, 12–13.
March, J.G. and Simon, H.A. (1958). *Organizations*. Cambridge, MA: Wiley.
Mathaisel, D.F.X., Cathcart, T.P. and Comm, C.L. (2004). A framework for benchmarking, classifying and implementing best sustainment practices, *Benchmarking: an International Journal*, **11**, 403–417.
Mayle, D.T., Hinton, C.M., Francis, G.A.J. and Holloway, J.A. (2001). What really goes on in the name of benchmarking? In Neely, A. (ed.), *Business Performance Measurement*. Cambridge: Cambridge University Press.
Merrifield, D.B. (1994). Measurements of productivity: key to survival. *International Journal of Technology Management*, **9**, 771–783
Meyer, J. and Rowan, B. (1991). Institutionalised organizations: formal structure as myth and ceremony. In Powell, W. and DiMaggio, P. (eds), *The New Institutionalism in Organizational Analysis*. Chicago: University of Chicago Press.
Murdoch, A. (1997). Lateral benchmarking or what Formula One taught an airline. *Management Today*, November, 64–67.
Neely, A., Kennerley, M. and Martinez, V. (2004). Does the balanced scorecard work: an empirical investigation. *Proceedings of the 4th International Conference on Performance Measurement*, Edinburgh, pp. 763–770. Cranfield, Beds.: Centre for Business Performance, Cranfield University.
Nørreklit, H. (2000). The Balance on the Balanced Scorecard – a critical analysis of some of its assumptions. *Management Accounting Research*, **11**, 65–88.
Nørreklit, H., Mitchell, F. and Nørreklit, L. (forthcoming). The validity of performance measurement frameworks. In Neely, A. (ed.), *Business Performance Measurement – Frameworks and Methodologies*. Cambridge: Cambridge University Press.
Northcott, D. and Llewellyn, S. (2005). Benchmarking in UK health: a gap between policy and practice? *Benchmarking: an International Journal*, **12**, 419–435.
Partnership Sourcing (1997). *Benchmarking the Supply Chain. First Cycle of Surveys*. London: Partnership Sourcing.
Paton, R. (2003). *Managing and Measuring Social Enterprises*. London: Sage.
Paton, R. and Payne, G. (1997). Benchmarking – passing fad or probable future? In: Pharaoh, C. (ed.), *Dimensions of the Voluntary Sector*. West Malling: Charities Aid Foundation.
Peters, T. and Waterman, R.H. (1982). *In Search of Excellence: Lessons from America's Best Run Companies*. London: Harper & Collins.
Pryor, L.S. (1989). Benchmarking: a self-improvement strategy. *Journal of Business Strategy*, Nov./Dec. 28–32.
Royal Statistical Society (2003). *Performance Indicators: Good, Bad, and Ugly*. London: RSS Working Party on Performance Monitoring in the Public Services.
Schofield, A. (ed.) (1998). *Benchmarking in Higher Education: An International Review*. CHEMS: London and UNESCO: Paris.
Sheridan, J.H. (1993). Where benchmarkers go wrong. *Industry Week*, 15 March, 28–34.
Shetty, Y.K. (1993). Aiming high: competitive benchmarking for superior performance. *Long Range Planning*, **26**, 39–44.
Sitkin, S.B., Sutcliffe, K.M. and Schroeder, R.G. (1994). Distinguishing control from learning in total quality management: a contingency perspective. *Academy of Management Review*, **19**, 537–564.
Society of Management Accountants of Canada (1995). Implementing benchmarking. *Management Accounting Guideline 16*.
Solovy, A. (1997), Benchmarking guide. *Hospitals and Health Networks*, 15 January.
Spendolini, M.J. (1992). *The Benchmarking Book*. New York: American Management Association.
Stephens, A. and Bowerman, M. (1997). Benchmarking for best value in local authorities. *Management Accounting*, November, 76–77.
Tranfield, D. and Starkey, K. (1998). The nature, social organisation and promotion of management research: towards policy. *British Journal of Management*, **9**, 341–353.
Tranfield, D., Denyer, D. and Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management*, **14**, 207–222.
Trosa, S. and Williams, S. (1996). Benchmarking in public sector performance management. *Performance Measurement in Government. OECD Occasional Papers* No. 9.

Uzzi, B. (1996). The sources and consequences of embeddedness for the economic performance of organizations: the network effect. *American Sociological Review*, **61**, 674–698.

Wait, S. (2004). *Benchmarking: a Policy Analysis*. London: The Nuffield Trust.

Walgenbach, P. and Hegele, C. (2001). What can an apple learn from an orange? Or: what do companies use benchmarking for? *Organization*, **8**, 121–144.

Watson, G.H. (1993). *Strategic Benchmarking*. New York: Wiley.

Welch, S. and Mann, R. (2001). The development of a benchmarking and performance improvement resource. *Benchmarking: an International Journal*, **8**, 431–452.

Wilkinson, A. and Willmott, H.C. (1995). *Making Quality Critical*, London: Routledge.

Wolfram Cox, J.R., Mann, L. and Samson, D. (1997). Benchmarking as a mixed metaphor: disentangling assumptions of competition and collaboration. *Journal of Management Studies*, **34**, 285–314.

Yarrow, D., Hanson, P. and Robson, A. (2004). Made in the 21st century: how far have we come on the journey to excellence? *Total Quality Management*, **15**, 829–839.

Yasin, M.M. (2002). The theory and practice of benchmarking; then and now. *Benchmarking: an International Journal*, **9**, 217–243.

Yoshikawa, T., Innes, J., Mitchel, F. and Tanaka, M. (1993). *Contemporary Cost Management*. London: Chapman & Hall.

Young, S.M. and Selto, F.H. (1991). New manufacturing practices and cost management: a review of the literature and directions for research. *Journal of Accounting Literature*, **10**, 265–297.

Zaheer, A. and Bell, G.G. (2005). Benefiting from network position: firm capabilities, structural holes, and performance. *Strategic Management Journal*, **26**, 809–825.

Zairi, M. (1998). *Benchmarking for Best Practice: Continuous Learning Through Sustainable Innovation*. Oxford: Butterworth-Heinemann.

Zairi, M. and Leonard, P. (1994). *Practical Benchmarking: The Complete Guide*, London: Chapman & Hall.

Zairi, M. and Youssef, M. (1995). A review of key publications on benchmarking: Part I. *Benchmarking for Quality Management and Technology*, **2**, 65–72.

Zairi, M. and Youssef, M. (1996). A review of key publications on benchmarking: Part II. *Benchmarking for Quality Management and technology*, **3**, 45–49.

Graham Francis is from the Waikato Management School, University of Waikato, Hamilton, New Zealand. Jacky Holloway is from the Open University Business School, The Open University, Walton Hall, Milton Keynes MK7 6AA, UK.