



En route to a theory of benchmarking

John P. Moriarty

*Victoria Management School, Victoria University Wellington,
Canterbury, New Zealand, and*

Clive Smallman

Commerce Division, Lincoln University, Canterbury, New Zealand

Abstract

Purpose – The purpose of this paper is to review the epistemology of benchmarking and identify methodological elements of a theory of benchmarking.

Design/methodology/approach – A thematic approach is applied to origins, primal and functional definitions of benchmarking.

Findings – Benchmarking remains theoretically underdetermined, with publications focusing on pragmatism and praxis rather than epistemology. Analysis of the literature leads to a new definition of benchmarking focusing around the teleological processes that lead to state-transformation of organizations.

Research limitations/implications – A theoretical foundation for benchmarking should be consistent with current organizational paradigms. Going forward the paper aims to develop a theory of benchmarking based on illustrative model derived from the thematic review.

Practical implications – The paper initiates the development of a more rigorous theoretical base for future benchmarking practice, which will strengthen organizations' business cases for undertaking such processes.

Originality/value – Recasts much of the extant literature in beginning to focus on the fundamentals of benchmarking.

Keywords Benchmarking, Best practice, Organizational theory, Epistemology

Paper type General review

Introduction

Benchmarking is an often-used modern term associated with a broad range of human endeavour. It is increasingly found to be essential to any serious organizational improvement process (Chen, 2005; Dawkins *et al.*, 2007), where current states of affairs are deemed undesirable (by dint of inefficiency or ineffectiveness) until replaced by more desirable states of affairs, based on evidence or perhaps a belief that such states are at least partially attainable. The *locus* of benchmarking lies between the current and desirable states of affairs and contributes to the transformation processes that realise these improvements. Understandably this concept is well researched and there are various identifiable trends in the epistemology of benchmarking. A recent literature review (Dattakumar and Jagadeesh, 2003) noted that 55 per cent of cited benchmarking publications focus on applications, case studies, education, innovations and extensions with the remaining 45 per cent focusing on models, general issues and fundamentals. Publication trends from 1986 through 2002 indicate that benchmarking applications, case studies and models dominate the literature since it is clearly perceived to be beneficial to practitioners. Similarly, content analysis of *Benchmarking: An International Journal* over



the period 1994-2008 is consistent with this trend. Out of 406 papers, 70 per cent were of a general research nature. These empirical studies included models or frameworks addressing the role and application of benchmarking as an organizational improvement mechanism (Chen, 2005; Dawkins *et al.*, 2007; Alstete, 2008). A further 12 per cent reported case studies of the application of benchmarking techniques, but only 4 per cent of all publications were conceptual (Anand and Kodali, 2008). None addressed the underlying nature of benchmarking. The number of reported benchmarking frameworks is also extremely large. Watson (1993) cited 69 different frameworks, Kozak and Nield (2001) identified approximately 40 different models outlining the process of benchmarking; some originating from organizations and others from researchers and consulting agencies. More recently, Anand and Kodali (2008) benchmarked 35 different frameworks against each other in a search for a single, exemplary framework.

We contend that the conceptual research using benchmarking represents a gap within the literature. This statement comes from the absence of lucid primal definitions of benchmarking. No such definitions emphasise benchmarking's contribution to organizational success through the principal process of organizational adaptation that is triggered by belief in knowledge of better performance elsewhere and driven by the extent of its superiority. Primal definitions are a theoretical, although they clearly possess substantial provenance in human experience and management practice (Jackson *et al.*, 1994; Zairi and Youssef, 1996; Yasin, 2002; Dattakumar and Jagadeesh, 2003). In this paper, we aim to address the under-researched area of primal definitions that attempt to describe benchmarking in absolute words. We conceptually review the primal definitions of what benchmarking is and what is benchmarked within an organizational perspective and based on our conceptual review, we deliver a provisional definition of benchmarking. The process flow of our research is captured in Figure 1.

Our conceptual review is structured along common themes (concepts) looking at primal definitions of benchmarking in literature. Hence, the contribution of our paper does not lie only in a thorough literature review on what benchmarking is, but also in mapping how this simple question was grasped by various researchers. This is then translated into sub-concepts representing streams contributing to the understanding of

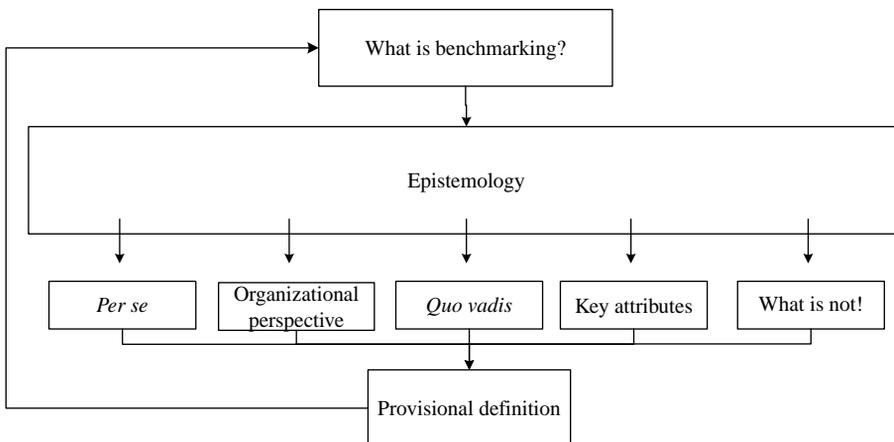


Figure 1.
Process flow of researched
phenomenon

what benchmarking is within an organizational perspective. These sub-concepts form the structure of our conceptual review delivered in this paper which leads into formulation of our provisional definition.

Benchmarking – *per se*

The etymology of “benchmark” was in words used *circa* 1842 to describe the surveying practice of establishing marks in the ground to ensure that subsequent placements of a bench supporting surveyor’s tools or instruments was assured to be on a level plane and assurance that subsequent measurements from the same place were on exactly the same basis.

“Benchmark”, as a noun, describes a point of reference and subsequently extended beyond surveying into a spectrum of organizational practices where the analogy of a level plane is some level of organizational performance or achievement. “Benchmarking”, as an adjective, refers to a process which not only seeks to identify disparate points of reference but also has the objective of aligning them in some favourable manner. This definition also establishes a fundamental aspect of benchmarking that requires two parties: the exemplar demonstrating a desirable state of affairs and the anomalar seeking to approximate or attain that desirable state of affairs. Thus, benchmarking represents a process of organizational adaptation where the focus is not simply on copying others but on learning how to improve organizational performance by sharing ideas (Watson, 1993). This approach establishes benchmarking beyond that of a technique or a tool, but a powerful concept with a change agent impacting on behaviour modification and developing new ways to manage business. Neither is it simply competitor analysis, espionage or theft (Zairi, 1997).

Benchmarking – through organizational perspectives

Benchmarking definitions extend through various organizational perspectives. (Leibfried and McNair, 1992) define benchmarking as “an external focus on internal activities, functions or operations in order to achieve continuous improvement”. This is consistent with (Deming, 1986, p. 85) theory of quality management where sustained continuous improvement is embedded in a thorough feedback mechanism. This feedback includes both internal and external referents (benchmarks) of quality production. Similarly (Spendolini, 1992) generalises Xerox’s successful in-house quality improvement process, in observing the need for improvement as continuous and benchmarking attributes to be evaluated for their appropriateness in order to maintain organizational advantage. Hence, benchmarking becomes a process that establishes the ground for creative breakthrough, a process of identifying the highest standards of excellence for products and services, and then making the improvements necessary to reach those standards by addressing the management and operational skills responsible for production (Bhutta and Huq, 1999; Kozak and Nield, 2001). In contrast, Kumar and Chandra (2001) espouse a manufacturing industry perspective, claiming that benchmarking can be considered a form of “reverse engineering” where the performance goals from other successful organizations are assumed to be achievable and applicable to others. This approach conflicts with the concept of benchmarking as a trigger or catalyst for organizational adaptation and suggests that it is feasible to reverse-engineer innovation. This is certainly the case in environments such as information science and manufacturing where the relationships between identical resources and instructions or recipes produce relatively consistent outcomes.

Benchmarking can also be detected as the main tool delivering improvements through various management systems. For example, Argyris (1977) applies benchmarking to double loop organizational learning and McAdam and McCreedy (1999) to knowledge management initiatives. Furthermore, benchmarking is explicitly referred to in the more tactical areas of total quality management (Franceschini *et al.*, 2006), supply chain management (Deming, 1982; Zairi and Baidoun, 2003), balanced scorecards (Kaplan and Norton, 1992), Six Sigma (Xerox, 1979), innovation (Radnor and Robinson, 2000), performance measurement (Carpinetti and de Melo, 2002; Anderson and McAdam, 2004; Dawkins *et al.*, 2007; Alstete, 2008), and business excellence models (EFQM, 2003; NIST, 2007). These applications refine the definition of benchmarking. Here, it means a search for organizations' best practices that lead to superior performance (Camp, 1989). It is also defined as a continuous and systematic process of evaluating organizations recognised as leaders by their peers determining business and work processes that represent best practices and establishing rational performance goals (Zairi, 1994b). Alternatively benchmarking is seen as a practice whose central essence is learning how to improve activities, processes and management (Ahmed and Rafiq, 1998). The framework of organizational learning was chosen by Liang (2004) as a means of establishing a theory of benchmarking. He cites organizational learning as the "effective processing, interpretation of and response to, information both inside and outside the organization" (Easterby-Smith *et al.*, 1999, p. 3). The link to benchmarking is further developed through Huber's (1991) assertion that "an organization learns if any of its units acquires knowledge that it recognises as potentially useful to the organization". Moreover, people learn either from their own experiences or from others. Liang (2004) develops the perspective that the ability to take advantage of others' experiences to build up one's own body of knowledge is one of the most important sources of human and social development. This is consistent with Zairi and Ahmed's (1999) observation that the foundations of benchmarking – "observing a state of affairs and upon deeming it to be desirable and worthy of attainment giving rise to its pursuit" are practices "as old as humankind". Thus, benchmarking is a learning tool. The learning environment is both internal and external to the organization and is encompassed by Zairi's (1994a, b) taxonomy of benchmarking. However, if benchmarking is a method of learning how to learn (Liang, 2004, p. 24), the issue is how this occurs and whether is it describable in a manner that distinguishes between effective and ineffective efforts.

Huber (1991, pp. 96-9) refers to the ways in which second-hand experience is acquired by an organization. Corporate intelligence (understanding competitors), institutional theory (pervasive imitation), grafting (acquisition, merging), scanning (environmental scanning to minimise the impact of change), focussed searching (learning to search for alternatives based on a shortfall of internal welfare – budgets or expectations not being met) and performance monitoring (measuring and learning from errors) are cited as techniques that abet organizational learning. However, whilst each of these techniques describes a mechanism for increasing organizational knowledge, they are all vicarious or sympathetic techniques that are unaccompanied by theory that can distinguish between effective and ineffective efforts. Each of these techniques can be found in the benchmarking processes already referred to and would offer explanation as to why a particular organization was chosen to be an exemplar. However, there is a little doubt that learning from exemplars can increase organizational performance, but the citations claim that sometimes it does not.

Benchmarking definitions also extend through the addition of meta – information to identify functional elements. For example, Watson (1993) repeats the view that benchmarking is a continuous process that searches for and applies significantly better practices for the purpose of achieving superior competitive performance, whilst Garvin (1993) extends this by adding meta-data to qualify “how” this occurs. Garvin (1993) notes that the continuous process is “disciplined” and the search is “thorough”, incorporated into a “careful study” of one’s own practices “and performance”. This is, then, extended through activities that included “systematic” visits to exemplars, and concluded with “analysis” that produces “recommendations and an implementation pathway”.

However, these definitions are predominantly outcome orientated: they address the purpose of benchmarking, not in terms of its essence, but in terms of its potential contribution to organizational success. These definitions admit no purpose to benchmarking other than organizational performance improvement; that is to generate prosperity in the face of competition and to sustain organizational health over time. Indeed, Zairi and Baidoun (2003, p. 12) reiterates an earlier theme: that benchmarking has the objective of establishing rational performance goals.

Arguably, a Darwinian tone permeates these definitions: organizational improvement is essential for survival, but is entirely optional. It is less about the random selection of good practices or re-inventing “the wheel”, but more about a purposeful search amongst exemplars for survival-enhancing attributes that can be adapted and implemented. In summary, these definitions suggest benchmarking to be a purposeful agent that is a contributor towards organizational perfection.

Benchmarking as *quo vadis*

Another perspective adding to the answer of our research question (what is benchmarking?) is the understanding of its historical development. Does benchmarking mean the same today as it did in the past? What triggers the benchmarking? What is the scope of benchmarking and what should be benchmarked? These are questions that direct us toward the literature that reviews trends in scope of benchmarking. We deliver this perspective through Zairi’s (1994a) taxonomy identifying the essential types of benchmarking and enhance this view with the timeline perspective captured by (Watson, 1993; Ahmed and Rafiq, 1998; Kumar and Chandra, 2001; Kyrö, 2003). This perspective is outlined by Figure 2.

Benchmarking taxonomy

Zairi (1994b) identifies essential types of benchmarking in his taxonomy:

- *Internal benchmarking.* Intra-organizational exemplars of replicated activities provide a trigger for improving anomalous performance. Any element of an organization achieving superior performance in any common practice may be used as the template for all others doing likewise. Internal benchmarking may also apply to public sector organizations where the absence of market forces may be replaced by a systematic comparison of best practices. Examples include “branch” performance in distributed organizations, customer service performance between different service locations, and public sector organizations sharing common stakeholders (e.g. hospital boards, government departments). This form of benchmarking is an application of organizational learning where proven innovation may be replicated without the usual competitive constraints to improve overall welfare.

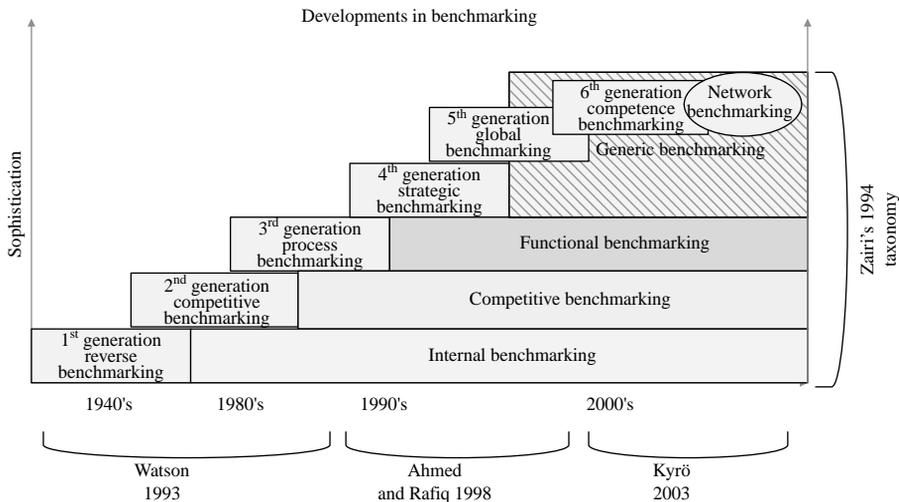


Figure 2.
Developments in
benchmarking

Sources: Adapted from Watson (1993); Ahmed and Rafiq (1998); Kyrö (2003)

- *Competitive benchmarking.* An organization's business practices are re-evaluated in the light of knowledge that their primary competitors have been observed to demonstrate superiority in some important elements of performance. Conventional triggers of such re-evaluation include observable customer-facing factors such as defect rates or process speed. This form of benchmarking is at "arms-length".
- *Functional benchmarking.* An organization's business practices are re-evaluated in the light of knowledge that non-competitor organizations (exemplars) demonstrate superiority in some common elements of business practice. This triggers re-evaluation of these business practices, often in partnership or in conjunction with exemplars. Common elements such as the use of information technology, administrative or logistical processes permits cooperation between organizations since the risk of market-place competition is non-existent.
- *Generic benchmarking.* An organization's business practices are purposefully compared with organizations having demonstrably superior performance from broadly similar practices or dispositions. Comparisons of exemplar practices, either through a conscious search or through observed performance, are conducted irrespective of the type of industry or location. This is the broadest form of benchmarking as it is triggered by broadly applicable practices that markedly improve performance. Examples such as "Just-In-Time" production management and zero-waste environmental practices improve efficiency in a generic manner and have minimal cross-sector or competitive overtones. Dispositions such as "agility", "responsibility", "achievement" or "innovative" are also encompassed by generic benchmarking.

Zairi's (1994a, b) taxonomy may also be viewed as a process or journey of increasing sophistication or adaptation. This journey commences with organizations using benchmarking to identify and replicate superior achievement through what Watson

(1993) and Kumar and Chandra (2001) termed “reversed engineering”. This concept extended to address broader inter-organizational performance gaps arising from demonstrable competitive advantage and identified two components that needed to be benchmarked: the first being evident external performance (competitiveness) and the second being evident internal performance (processes). The former – evidence of “what” an exemplar is achieving and the latter – a more complex issue as to “how” this achievement occurs. The final extension of this progression involved recognition the exemplar need not be in either the same industry or indeed the same country, so long as there was some benefit to be gained by examining their superior dispositions or practices and applying any learning arising from doing so. Watson’s (1993) early perspectives of this progression were extended by Ahmed and Rafiq (1998) as strategic benchmarking and later augmented by Kyrö (2003) as competence and network benchmarking. Yet these perspectives remain within Zairi’s taxonomy.

There is also the important question as to whether one type of benchmarking is more appropriate than another. Implementation is the process of achieving these requirements. Many authors have commented on the appropriateness of one type of benchmarking over another. Bhutta and Huq (1999, p. 257) cite Leibfried and McNair’s (1992) relevance Table I.

High relevance accrues from the close matching of an anomalars requirements to those of an exemplar. For instance, strategic direction and current performance are more associated with market conditions within a competitive milieu than elsewhere: successes being attributed to competitive advantages arising from superior practices. Similarly, exemplar processes, such as those governing the production of a commodity, are highly relevant to anomalars engaged in identical activities (production yields being an example of a discerning metric).

Key attributes of benchmarking

Another common thread dominating benchmarking literature focuses on key attributes necessary to address when effective benchmarking is at stake. Ahmed and Rafiq (1998, p. 228) identify these key attributes as measurement, continuous improvement and systematic implementation. However, these are broad concepts common to most organizational endeavours. Moreover, all definitions imply that benchmarking is a process –, i.e. a sequence of activities that involves “process and assessment”.

Process. Process relates to the underlying steps governing the activities or processes of interest to the anomalars.

Kyrö (2003) extends the concepts and forms of benchmarking by noting that organizational developments suggest two additional attributes of benchmarking: application to internal organizational learning processes and recognition of the

	Internal benchmarking	Competitor benchmarking	Functional benchmarking	Generic benchmarking
Performance benchmarking	Medium	High	Medium	Low
Process benchmarking	Medium	Low	High	High
Strategic benchmarking	Low	High	Low	Low

Source: Leibfried and McNair (1992)

Table I.
Benchmarking
relevance table

dispersive nature of exemplars. For example, if activities (e.g. production, design, research and development) are dispersed not simply throughout a single organization, but also throughout partner organizations, there is an additional challenge of performing measurements and achieving outcomes consistent with process. Internal organizational learning itself then becomes an important enabler in simply applying a benchmarking process. These two additional distinctions add nuance to Zairi's (1994a, b) "generic benchmarking" and recognise the evolving nature of organizational behaviour, and hence these practices identify the scope of benchmarking. What is clear that from its earliest formal conceptualization onwards, is the tendency for organizations to seek some form of assurance that critical success factors are able to be compared with similar factors wherever else they might be found (in competitor organizations, kindred organizations or simply in any analogous situation). There is the presumption that practices underpinning the nature of comparable factors are both understood and congruent between anomalar and exemplar organizations. Furthermore, quantifying gaps in factor performance (assessments) and quantifying achievements also relies on this congruence otherwise it is not benchmarking. How these practices are implemented with any certainty is clearly very important.

However, there is an inescapable observation that arises from Zairi's (1994a, b) taxonomy of benchmarking: it is really immaterial who the exemplar is, provided there is appropriate congruence between its states of affairs and those of the anomalar. What the taxonomy illustrates is an expanding *locus* of likelihood as to where a suitable exemplar might be found and nothing more. The taxonomy does not address how an anomalar can benefit, a priori, from a benchmarking partnership.

Assessment. Assessment is a formal measurement process that identifies the performance gap between the exemplar and anomalar. The magnitude of the gap between respective measurements serves to quantify the latent potential available for release in the outcomes phase. We identified that most of the "assessment-oriented" literature focuses at identification of what should be benchmarked within a particular organizational setting and onto "how" these can be achieved.

For example, Xerox reports its success through adoption of "TQM" process; namely "Plan, Do, Check Act" (PDCA) process model that had its origins in the formation of the "scientific method" first described by Bacon (1620/2000) and subsequently extended into manufacturing and organizational practices by Shewhart (1980) and more notably by Deming (1986). Also referred to by Camp (1989), Drew (1997) and Carpinetti and de Melo (2002) the benchmarking implementation model (Figure 2) includes elements of Leibfreid and McNair's (1992) relevance approach and the PDCA model as depicted by Figure 3.

These five steps outline the approach taken for using benchmarking to achieve the objective outcome of organizational improvement. The details within each step also expand and may contain highly-complex sub-processes. For example, a suitably empowered leadership prioritises the facets of organizational performance to be benchmarked. In a simple situation, a subjective prioritisation may apply but for complex situations where there are substantial numbers of related processes contributing to the selected organizational objective, prioritisation of what to benchmark may be objectively assessed through analysis. Quantitative techniques such as analytical hierarchy process (AHP) (Partovi, 1994), principal component analysis (PCA) and common factor analysis (CFA) (Büyüközkan and Maire, 1998) may stand as examples. AHP scrutinizes

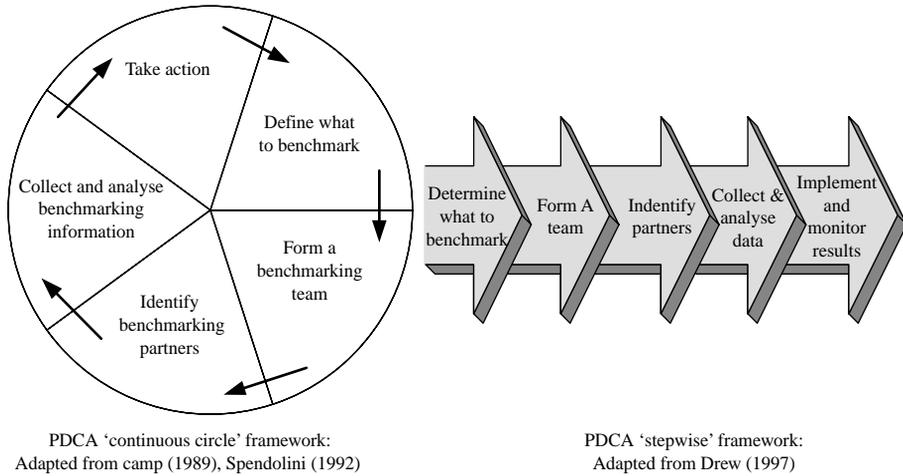


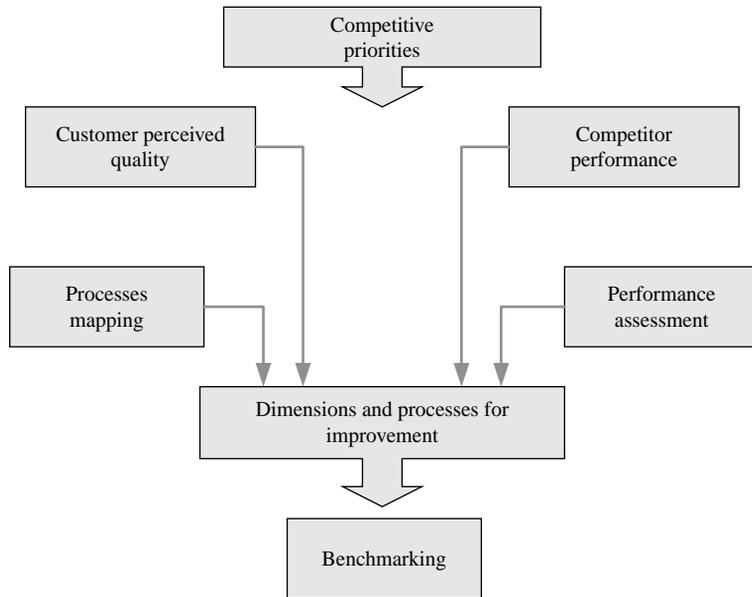
Figure 3.
Benchmarking process
models based on the
plan-do-check-act
TQM model

relationships between organizational objectives and their associated processes to be quantified which leads into subsequent sensitivity analysis determining benchmarking factors (or “relata”). PCA and CFA represent statistical approaches that rely on identification of process variances and inter-process dependencies to derive factors (benchmark relata) that contribute the greatest variance to organizational objectives (such as customer satisfaction). However, these techniques are complex and the results may be difficult to interpret.

Carpinetti and de Melo (2002) suggest the systematic mapping and analysis of a wide range of defining relata (Figure 4). They assert that relata “contributing to efficiency and effectiveness of business process mostly relate to prioritised competitive criteria” in determination of what should be benchmarked. However, the analysis of these relata still present issues as relationships between them remain to be quantified. Of course, this is but the first step in the benchmarking implementation process. Once the prioritised relata are identified, additional processes are needed to identify exemplars and, once found, isolate and associate the corresponding exemplary relata with them.

Similarly to Xerox story, Zairi and Baidoun (2003) identify a 12 step implementation model that was used by Yellow Pages (a division of British Telecommunications plc) to implement benchmarking. This process is captured by Table II. This model combines analytical and practical steps to guide the process.

In this model, steps three and four might be addressed using techniques described earlier in Xerox experience. What is significant about this model is its sheer extent. The PDCA models of Camp (1989) and Drew (1997) are abbreviations of this model, but they all identify selection of benchmarking relata to be critical and also suggest that benchmarking implementation is no small effort, particularly for a sizeable organization. Specific emphasis on “management commitment” and partnership issues such as exemplar selection and site visits suggest that the implementation costs (and risks) associated with benchmarking are non-trivial. Zairi’s (1994a, b) also warns that if process-driven (functional) benchmarking is undertaken, care must be exercised to involve



Source: Adapted from Carpinetti and De Melo (2002)

Figure 4.
Defining what to
benchmark

1.	Ensure management commitment
2.	Process selection
3.	Selecting your targets
4.	Process mapping
5.	Start partnership selection
6.	Successful selection
7.	Preparation for site visits
8.	The site visit
9.	Identify practical solutions and plan action
10.	Implement
11.	Keep in touch
12.	Continuous improvement

Source: Zairi and Baidoun (2003, p. 13)

Table II.
A benchmarking
implementation model

not only process outputs but also information on how they are achieved. This caution is more broadly amplified by Partovi (1994), Cassell *et al.* (2001) and Carpinetti and de Melo's (2002) emphasis on the adoption of a broad systematic approach towards selection of benchmark relata. Indeed, Partovi (1994) claims that failure to identify priority benchmark relata most probably invalidates subsequent benchmarking analysis. For example, output (cost) driven benchmarking might conflict with other organizational objectives such as customer satisfaction if the relationship between cost and quality is not well understood.

The common thread throughout these benchmark implementation models is the need to identify priority relata that impinge on organizational performance; the relationships between these priority relata and other organizational processes; exemplars evincing sufficient similarity to trigger improvement initiatives; and the capacity to implement improvements.

An a priori methodology that informs on the feasibility of a benchmarking proposition would reduce the both the implicit and explicit risks associated with implementation.

Benchmarking as what it is not, alias criticisms of benchmarking

A significant stream adding to the formulation of what is benchmarking within the literature is represented through understanding of what benchmarking is not. To satisfy this requirement, we looked at research that focuses on the unrequited effects of benchmarking. Understandably, organizational improvement is the principal objective of benchmarking, yet even this clearly desirable objective is not without significant criticism. Some of the criticisms implicitly arise from the difficulties of obtaining reliable exemplar information, as well as from the difficulties inherent in achieving the organizational changes suggested by the benchmarking processes. Other criticisms arise from the increasing complexities applied to benchmarking typologies where refinements do not add to the certainty of implementing benchmarking but rather expand on the circumstances in which it might be applied. Expanding typologies is akin to violating Occam's Razor[1], and multiplying entities without simplifying their application arguably gives rise to the criticism that benchmarking is a "fad". Indeed, Anand and Kodali (2008) questioned the need for so many frameworks and concluded that there only needed to be two typologies of benchmarking: "internal and external". One might go even further and wonder why there is a need for more than one benchmarking framework and a single typology!

However, it is useful to examine benchmarking's critics to determine any common themes.

Information-oriented criticisms. Benchmarking requires an exemplar and an anomalar. The degree to which sufficiently reliable information can be obtained on exemplar performance is a vital component of any benchmarking implementation process. Exemplars may have sound competitive reasons to secure their advantages from others. Campbell (1999) notes that anomalars spending considerable effort in attempting to gather information describing exemplar advantages (often by covert means) may fail to focus on their own unique situation and become prone to distraction and misdirection. Moreover, benchmarking is always a retrospective process: learning what used to be the case may not yield much advantage in fast-moving markets.

Kozak and Nield (2001) take this further by claiming that the information required to implement benchmarking reduces heterogeneity within industries and increases the risks of uncompetitive homogeneity if product differentiation declines. Elnathan *et al.* (1996) focus on the costs associated with information gathering and suggest that costs in the form of employee time taken to obtain comparable sets of data, whilst traceable, might be overlooked when determining the cost-benefits of benchmarking.

A common theme threading through all information-orientated criticisms is that benchmarking is intrinsically retrospective and may even be inefficient. Any exemplary state of affairs examined (or adopted) by the anomalar will not only be historical but may also be, unknowingly, disassociated from the exemplar organization's future purpose (its teleology). This is converse to Kozak and Nield's (2001) claim of uncompetitive

homogeneity, since an exemplar with (successfully) evolving states of affairs in pursuit of some undetected future purpose may provide anomalous with uncompetitive benchmarks!

Implementation-oriented criticisms. Implementing a successful benchmarking programme requires more than adherence to the step-wise programmes outlined previously. Watson (1994) cites two major difficulties in implementing a rigorous benchmarking study: deciding what project to focus the benchmarking resources on; and then what organizations to solicit as partners. Simply applying Zairi and Baidoun's (2003) 12 steps or the five steps of Spendolini (1992), Camp (1989) and proponents of the TQM PDCA approach will likely be insufficient according to Francis and Holloway (2007). They observe that previous implementation experience, good interdisciplinary working, top management commitment and realistic resources are characteristics associated with benchmarking implementation success; in short slavish adherence to some formula is unreliable. More recently Huq *et al.* (2008) reported that failure to prioritise the anomalous's own (internal) situation before seeking exemplars elsewhere considerably increased the risk of implementation failure. Perhaps, the inability to implement benchmarking in a more certain manner has deeper roots? Wolfram Cox *et al.* (1997) lament the absence of a sufficiently developed theory that would explain the differences between effective and ineffective efforts. Wöber (2002) also stated that benchmarking lacked a rigorous foundation in management science and added that a generally accepted methodology for selecting suitable benchmarking partners was only addressed in 2000. He refers to the application of non-parametric frontier analysis techniques such as data envelopment analysis (DEA) (Charnes *et al.*, 1978; Banker *et al.*, 1984) and statistical techniques to examine and rank the technical efficiency of sets of "decision-making units". A frontier orientated approach provides a means of identifying appropriate exemplars and of quantifying the optimal parameters that "might" elevate less technically efficient performers (anomalous) to that of the exemplar. Non-parametric data is particularly appealing as almost any data reflective of enterprise performance efficiency suffices. For example, financial data may be combined with production and customer data to form an efficiency frontier. However, these extremely popular techniques should not be applied without close regard to the circumstances. Although often under-emphasised, the application of DEA requires careful attention to "noise", the reliability of longitudinal data sets and scalability. It may also be challenging for non-mathematically inclined managers to interpret the output of a DEA model or determine whether it is properly specified (Belton and Vickers, 1993; Smith, 1997). In general, DEA interprets uncertainties in parameter values and "noisy data" as contributors to efficiency and thus affects the relative standing of the corresponding organization (Färe *et al.*, 2000). Longitudinal analysis of enterprises is also problematic and requires special techniques such as Malmquist's productivity index (Bjurek, 1996) which compares datasets representing different time periods. DEA requires that both the anomalous and exemplars in the dataset exhibit similar returns-to-scale – a particularly important factor when organizations from different industries, marketplaces or sizes are compared or in the situation where some data simply cannot be scaled (e.g. it may be impractical to scale scarce or constrained capital resources such as land or buildings).

In summary, the implementation of a benchmarking programme can be far from straightforward and although greatly assisted by the application of popular analytical techniques such as DEA, is more likely to be successful if conducted by experienced implementers who can navigate around the practical issues that are certain to arise.

Theoretical criticisms. The acceptance of benchmarking has resulted from its widespread use (Francis and Holloway, 2007) and by dint of this it has arguably escaped becoming “another management fad”. Yet there is little literature that focuses on the theoretical composition of benchmarking. Several publications additional to Liang (2004) have included a “theory of benchmarking” to support other work and it is useful to examine the basis of these claims.

Cox (2003) examined benchmarking in the context of computer science. In this discipline, a benchmark is a generally accepted reference against which various computer technologies are compared. The foundation of Cox’s benchmarking theory is Kuhn’s “structure of scientific revolutions” (Kuhn, 1996). Benchmarking is claimed to “operationalize scientific paradigms”. “A benchmark (in the discipline of computer science) takes an abstract concept (a paradigm) and makes it more concrete so that it can serve a guide for action” (Cox, 2003, p. 35). In other words, the acceptance of a benchmark fulfils Kuhn’s conditions for acceptance of the new paradigm as “normal science”. There are some parallels between this approach to a theory of benchmarking in computer science and organizations in general. Practices that have provenance in respect of economic performance are triggers for implementation by those aware of them, but do not have them. How these tests are established and whether they can be applied to a current situation to achieve exemplary outcomes is not addressed. What is addressed is that if a desirable state of affairs is achieved, the mechanisms that give rise to it are at least transiently exemplary and trigger others to equal or better them. Cox does not identify any a priori method of achieving a set of tests (equating to organizational methods) that trigger enhanced performance.

van Helden and Tillema (2005, p. 338) identified Public Sector benchmarking as an important surrogate for the absence of market forces. Their search for a benchmarking theory applicable to the public sector relies on the combination of three economic efficiency hypotheses and thirteen institutional reasoning hypotheses reflecting normative economic behaviour within public sector organizations. Economic efficiency is driven by innovation and customer satisfaction within a competitive environment. Inefficient performance, whether real or perceived, may attract remedial action from higher authorities (e.g. government) and threaten survival (devolution into more efficient entities). The fundamental issue arising from their hypotheses is that they provide a framework for benchmarking utility, not a theory of benchmarking. However, their hypotheses establish a strong case for ensuring that any benchmarking theory should encompass economic efficiency and organizational teleology.

Discussion

This synopsis of benchmarking traces the emergence of a practice that has over the past 40 years become identified as a strong contributor to organizational improvement. Successful implementation of benchmarking is not without its difficulties, but the concept of an anomalar using an exemplar to identify internal practices that might be made more efficient and therefore improve its overall welfare is all but universally accepted practice.

The taxonomy of benchmarking has been the subject of considerable research. Zairi’s (1994a, b) advances a four-part taxonomy that identified an expanding *locus* of anomalar opportunity; the closest being the anomalar’s internal processes and the most distant being congruent exemplars located anywhere. This taxonomy has been extended and as

noted, adds nuance to the situations under which states of affairs might be benchmarked, but does not resolve the fundamental dilemma: a priori isolation of effective from ineffective effort.

Critics of benchmarking focus on three areas: information, implementation and theory. The reliability of exemplar information, the intangibilities associated with implementing benchmarking and the lack of a theoretical framework that distinguishes effective from ineffective efforts detract from the potential benchmarking appears to offer.

The literature is overwhelmingly pragmatic (that is process-driven, case-oriented and generic) as opposed to theoretical. Where theories are invoked, they centre on the utility of benchmarking in terms of organizational learning and reasoning as well as economic enhancement. Attachment to, or association with organizational theories has not elevated benchmarking to the stage where practitioners can embark on a programme that can be tested a priori, to judge whether it will feasibly deliver the sought-after results. Indeed, the literature warns of the need to furnish a benchmarking programme with ample financial, technical and leadership resources so as to reduce, rather than eliminate, ineffective effort. Such an approach may not even be efficient: ample resources may enhance the likelihood of extracting isolated benefits arising from benchmarking, but may do so with an overall loss of anomalar welfare.

There is, however, a single recurring theme: improving organizational welfare (both short and longer term) is emphasised. The nature of this is twofold: organizational survival is a continuous, purposeful pursuit that preserves or enhances the welfare of its stakeholders (staff, shareholders, the community and customers) and mechanisms that might contribute to this journey must be consistent with this purpose. But just as the utility of an automobile, as a means to the purpose of reaching a destination, can be enjoyed without an appreciation of thermodynamics, so too can the utility of benchmarking be enjoyed without a formal theory.

Yet all commentators hold that benchmarking is generally desirable. What is the fundamental reason for this? Improvement, whether at organizational or elemental levels, is certainly held to be desirable, but does this beg the question at hand: what is the concept of organizational improvement? Is there an endpoint to it? Is it a teleological or ontological question that can be answered? How does benchmarking contribute to the ontology of an organization and moreover, is it a teleological component or is it perhaps a chance-related element that sometimes results in the attainment of more desirable states of affairs?

It is clear that however it is described; benchmarking is intended to be a means towards the end of achieving a more desirable organizational state of affairs than is currently the case; not necessarily the best but better. Moreover, benchmarking might identify the changes necessary to achieve that end. This suggests that benchmarking is within an organizational teleology and its purpose associated with the ontological question “what is the essential nature of the organization”?

The concept of “change” is also implicit in benchmarking. Change as a constituent of benchmarking-directed improvement processes was summarised by Harrington (1995) as “all improvement is change, but not all change is improvement”. Process factor independence would greatly simplify the task of applying improvements to effect superior production. Yet, such simplicity is rare as factor interdependence and factor indeterminacy complicate purposeful change. Harrington (1995) also raises a

probabilistic tenor: benchmarking “might not lead to improvement” and associates it with an objective rather than deterministic role.

These perspectives indicate that benchmarking is instantiated for the sake of pursuing organizational improvement. As such it requires an anomalar as well as real or notional exemplars – whose locations are immaterial, that are perceived to evince superior states of affairs. Implementing benchmarking requires a mechanism to change the anomalar’s undesirable state of affairs to that of the exemplar. What is the nature of the mechanism driving such change? Superior states of affairs appear to be a necessary component of organizational survival in a competitive environment. Even where organizations operate in environments that are not subject to open market competition (examples of which often cite the public sector), efficient performance is ultimately intolerable to stakeholders. No controversy would arise from the perspective that organizational survival is predicated on the achievement of increasingly better states of affairs; organizational survival is teleological.

These perspectives suggest a provisional definition of benchmarking:

Benchmarking is an exemplar driven teleological process operating within an organization with the objective of intentionally changing an existing state of affairs into a superior state of affairs.

This definition can be further clarified by our model captured by Figure 5 that visualises the concept presented above.

Presented model shows benchmarking within its origins where internal organizational activities (drawn as vertical lines on axis X) have recognised values detected through organizational performance indicators (readable on axis Y). Benchmarking as an organizational initiative lies firmly on two fundamental pillars that represent first, the general understanding of what benchmarking is within an organizational perspective (primal pillar) and second, the operationalization of benchmarking endeavours

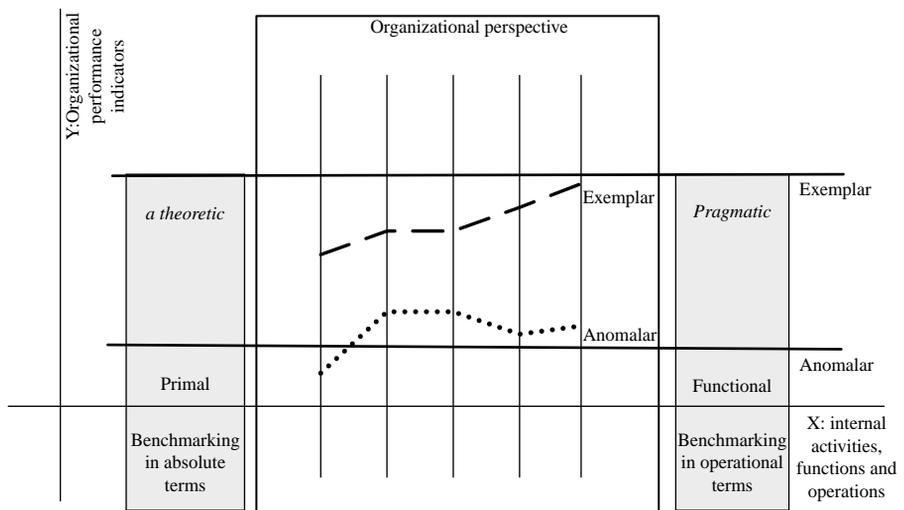


Figure 5.
Benchmarking model capturing the origins of benchmarking

(functional pillar). The difference between the actual and desired state of affairs between the anomalar and exemplar can be read through values on axis *Y*. For the clarity of this point the performance of anomalar and exemplar is shown as linear all across the organizational perspective (activities) – hence to further underline the level plane of benchmarking. Yet, we recognize that each anomalar and each exemplar perform differently at various organizational activities and processes. These performances are illustrated by dotted lines for both the anomalar and the exemplar. Understandably the anomalar can increase its performance while using the tool of benchmarking through appropriate definition of scope of benchmarking (identification of what to benchmark) and effective use of measurements to reach for desired performance outcomes (quality of benchmarking operationalization).

Benchmarking is not simply about change; rather it is the identification and successful implementation of a superior state of affairs within an anomalar's organization. Successful benchmarking requires that the anomalar determine the "cause" of an exemplary state of affairs and transfer its effects to their organization. The current literature is quite unspecific as to how this is achieved; instead relying on experienced practitioners and extensive process prescriptions to improve the success of what is clearly a complex task in other than very simple situations.

We conclude that a theoretical foundation for benchmarking should be consistent with current organizational paradigms and the nature of what constitutes current and superior states of affairs. In short, benchmarking exists, but its essence is obscure. Hence, our future research objective is to establish a theoretical basis for benchmarking based on classical (Aristotle, 1984) and modern (Hume, 1739/1999, 1748/2000; Peirce, 1898/1992, 1935, 1992/1998) theories of causation in conjunction with economic welfare theory to quantify the advancement from a current to superior states of affairs.

Note

1. "*entia non sunt multiplicanda praeter necessitatem*" or "Entities should not be multiplied unnecessarily".

References

- Ahmed, P.K. and Rafiq, M. (1998), "Integrated benchmarking: a holistic examination of select techniques for benchmarking analysis", *Benchmarking for Quality Management & Technology*, Vol. 5 No. 3, pp. 225-42.
- Alstete, J.W. (2008), "Measurement benchmarks or 'real' benchmarking?: an examination of current perspectives", *Benchmarking: An International Journal*, Vol. 15 No. 2, pp. 178-86.
- Anand, G. and Kodali, R. (2008), "Benchmarking the benchmarking models", *Benchmarking: An International Journal*, Vol. 15 No. 3, pp. 257-91.
- Anderson, A. and McAdam, R. (2004), "A critique of benchmarking and performance measurement", *Benchmarking: An International Journal*, Vol. 11 No. 5, pp. 465-83.
- Argyris, C. (1977), "Organizational learning and management information systems", *Accounting, Organizations and Society*, Vol. 2 No. 2, pp. 113-23.
- Aristotle (1984), *The Complete Works of Aristotle – The Revised Oxford Translation*, Barnes, J. (Ed.), Vol. 1, Princeton University Press, Princeton, NJ.
- Bacon, F. (1620/2000), *The New Organon*, Cambridge University Press, West Nyack, NY, p. 105.

- Banker, R.D., Charnes, A.C. and Cooper, W.W. (1984), "Some models for estimating technical and scale inefficiencies in data envelopment analysis", *Management Science*, Vol. 30 No. 9, pp. 1078-92.
- Belton, V. and Vickers, S.P. (1993), "Demystifying DEA – a visual interactive approach based on multiple criteria analysis", *Journal Operational Research Society*, Vol. 44 No. 9, p. 885.
- Bhutta, K.S. and Huq, F. (1999), "Benchmarking best practices: an integrated approach", *Benchmarking: An International Journal*, Vol. 6 No. 3, pp. 254-68.
- Bjurek, H. (1996), "The Malmquist total factor productivity index", *Scandinavian Journal of Economics*, Vol. 98 No. 2, pp. 303-13.
- Büyükoçkan, G. and Maire, J.L. (1998), "Benchmarking process formalization and a case study", *Benchmarking for Quality Management Technology*, Vol. 5 No. 2, pp. 101-25.
- Camp, R.C. (1989), *Benchmarking: The Search for Industry Best Practices that Lead to Superior Performance*, ASQC Quality Press, Milwaukee, WI.
- Campbell, A. (1999), "Tailored, not benchmarked. A fresh look at corporate planning", *Harvard Business Review*, Vol. 77 No. 2, pp. 41-50.
- Carpinetti, L.C.R. and de Melo, A.M. (2002), "What to benchmark?", *Benchmarking: An International Journal*, Vol. 9 No. 3, pp. 244-55.
- Cassell, C., Nadin, S. and Gray, M.O. (2001), "The use and effectiveness of benchmarking in SMEs", *Benchmarking: An International Journal*, Vol. 8 No. 3, pp. 212-22.
- Charnes, A.C., Cooper, W.W. and Rhodes, E. (1978), "Measuring the efficiency of decision making units", *European Journal of Operational Research*, Vol. 2 No. 6, pp. 429-44.
- Chen, H.L. (2005), "A competence-based strategic management model factoring in key success factors and benchmarking", *Benchmarking: An International Journal*, Vol. 12 No. 4, p. 364.
- Cox, S.E. (2003), *A Theory of Benchmarking with Applications to Software Reverse Engineering*, University of Toronto, Toronto.
- Dattakumar, R. and Jagadeesh, R. (2003), "A review of literature on benchmarking", *Benchmarking: An International Journal*, Vol. 10 No. 3, pp. 176-209.
- Dawkins, P., Feeny, S. and Harris, M.N. (2007), "Benchmarking firm performance", *Benchmarking: An International Journal*, Vol. 14 No. 6, pp. 693-7120.
- Deming, W.E. (1982), *Quality, Productivity, and Competitive Position: Center for Advanced Engineering Study*, Massachusetts Institute of Technology, Cambridge, MA.
- Deming, W.E. (1986), *Out of the Crisis*, MIT Center for Advanced Engineering Study, Cambridge, MA.
- Drew, S.A.W. (1997), "From knowledge to action: the impact of benchmarking on organizational performance", *Long Range Planning*, Vol. 30 No. 3, pp. 427-41.
- Easterby-Smith, M., Burgoyne, J. and Araujo, L. (1999), *Organizational Learning and the Learning Organization*, Sage, London.
- EFQM (2003), *Introducing Excellence*, European Foundation for Quality Management, Brussels, available at: www.efqm.org/uploads/introducing_english.pdf
- Elnathan, D., Lin, T.W. and Young, S.M. (1996), "Benchmarking and management accounting: a framework for research", *Journal of Management Accounting Research*, Vol. 8 No. 2, pp. 37-54.
- Färe, R., Grosskopf, S., Kirkley, J.E. and D, S. (2000), "Data envelopment analysis (DEA): a framework for assessing capacity in fisheries when data are limited", paper presented at IIFET X Conference, July, para 4.

- Franceschini, F., Galetto, M. and Cecconi, P. (2006), "A worldwide analysis of ISO 9000 standard diffusion: considerations and future development", *Benchmarking: An International Journal*, Vol. 13 No. 4, pp. 523-41.
- Francis, G. and Holloway, J. (2007), "What have we learned? Themes from the literature on best-practice benchmarking", *International Journal of Management Reviews*, Vol. 9 No. 3, pp. 171-89.
- Garvin, D.A. (1993), "Building a learning organization", *Harvard Business Review*, Vol. 71 No. 4, pp. 78-91.
- Harrington, H.J. (1995), "The new model for improvement: total improvement management", *Management Decision*, Vol. 33 No. 3, pp. 17-24.
- Huber, G.P. (1991), "Organizational learning: the contributing processes and the literatures", *Organization Science*, Vol. 2 No. 1, pp. 88-115.
- Hume, D. (1739/1999), "A treatise of human nature: being an attempt to introduce the experimental method of reasoning into moral subjects", Batoche Books, Kitchener.
- Hume, D. (1748/2000), "An Enquiry Concerning Human Understanding", Batoche Books, Kitchener, p. 113.
- Huq, F., Abbo, M.-H. and Huq, Z. (2008), "Perceptions about benchmarking best practices among French managers: an exploratory survey", *Benchmarking: An International Journal*, Vol. 15 No. 4, pp. 382-401.
- Jackson, A.E., Safford, R.R. and Swart, W.W. (1994), "Roadmap to current benchmarking literature", *Journal of Management in Engineering*, Vol. 10 No. 6, pp. 60-7.
- Kaplan, R.S. and Norton, D.P. (1992), "The balanced scorecard-measures that drive performance", *Harvard Business Review*, Vol. 70 No. 1, pp. 71-9.
- Kozak, M. and Nield, K. (2001), "An overview of benchmarking literature: its strengths and weaknesses", *Journal of Quality Assurance in Hospitality Tourism*, Vol. 2 Nos 3/4, pp. 7-32.
- Kuhn, T.S. (1996), *The Structure of Scientific Revolutions*, The University of Chicago Press, Chicago, IL, pp. 152-378.
- Kumar, S. and Chandra, C. (2001), "Enhancing the effectiveness of benchmarking in manufacturing organizations", *Industrial Management & Data Systems*, Vol. 100 No. 1, pp. 80-9.
- Kyrö, P. (2003), "Revising the concept and forms of benchmarking", *Benchmarking: An International Journal*, Vol. 10 No. 3, pp. 210-25.
- Leibfried, K.H.J. and McNair, C.J. (1992), *Benchmarking*, HarperBusiness, New York, NY.
- Liang, Y. (2004), *Benchmarking Theory Applied to State Parks: An Exploratory Study*, Oklahoma State University, Stillwater, OK.
- McAdam, R. and McCreedy, S. (1999), "A critical review of knowledge management models", *The Learning Organization: An International Journal*, Vol. 6 No. 3, pp. 91-101.
- National Institute for Standards and Technology (2007), *Malcolm Baldrige National Quality Awards, Scoring Guidelines, Consensus*, National Institute for Standards and Technology, Gaithersburg, MD, p. 9, available at: www.quality.nist.gov/Word_files/2007%20Consensus%20Materials/Consensus_Toolkit_Final.doc (accessed 15 December, 2006).
- Partovi, F.Y. (1994), "Determining what to benchmark: an analytic hierarchy process approach", *International Journal of Operations Production Management*, Vol. 14 No. 6, pp. 25-39.
- Peirce, C.S. (1898/1992) in Ketner, K.L. (Ed.), *Reasoning and Logic of Things*, Harvard University Press, Cambridge, MA, p. 197.

- Peirce, C.S. (1935), "Collected papers of Charles Sanders Peirce", in Hartshorne, C. and Weiss, P. (Eds), *Exact Logic*, Vol. VI, Harvard University Press, Cambridge, MA.
- Peirce, C.S. (1992/1998), "The essential Peirce: selected philosophical writings", in Houser, N. and Kloesel, C. (Eds), *The Peirce Edition Project*, Vol. I and II, Indiana University Press, Bloomington, IN.
- Radnor, Z. and Robinson, J. (2000), "Benchmarking innovation", *Creativity and Innovation Management*, Vol. 9 No. 2, pp. 3-13.
- Shewhart, W.A. (1980), *Economic Control of Quality of Manufactured Product*, American Society for Quality Control, Milwaukee, WI.
- Smith, P. (1997), "Model misspecification in data envelopment analysis", *Annals of Operations Research*, Vol. 73, pp. 233-52.
- Spendolini, M.J. (1992), "The benchmarking process", *Compensation Benefits Review*, Vol. 24 No. 5, p. 21.
- van Helden, G.J. and Tillema, S. (2005), "Search of a benchmarking theory for the public sector", *Financial Accountability Management*, Vol. 21 No. 3, pp. 337-61.
- Watson, G.H. (1993), *Strategic Benchmarking: How to Rate Your Company's Performance against the World's Best*, Wiley, New York, NY.
- Watson, G.H. (1994), "A Perspective on benchmarking: Gregory H. Watson in conversation with the editor", *Benchmarking: An International Journal*, Vol. 1 No. 1, pp. 5-10.
- Wöber, K.W. (2002), *Benchmarking in Tourism and Hospitality Industries: The Selection of Benchmarking Partners*, CABI Publishing, Wallingford, CT.
- 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*, Vol. 34 No. 2, pp. 285-314.
- Xerox (1979), "Benchmarking", Xerox Corporation, internal publication, London.
- Yasin, M.M. (2002), "The theory and practice of benchmarking: then and now", *Benchmarking: An International Journal*, Vol. 9 No. 3, pp. 217-43.
- Zairi, M. (1994a), "Benchmarking: the best tool for measuring competitiveness", *Benchmarking for Quality Management & Technology*, Vol. 1 No. 1, pp. 11-24.
- Zairi, M. (1994b), *Measuring Performance for Business Results*, Springer, New York, NY, pp. 62-3.
- Zairi, M. (1997), "Benchmarking: towards being an accepted management tool or is it on its way out?", *Total Quality Management*, Vol. 8 Nos 2/3, pp. 337-8.
- Zairi, M. and Ahmed, P.K. (1999), "Benchmarking maturity as we approach the millennium?", *Total Quality Management*, Vol. 10 Nos 4-5, pp. 810-6.
- Zairi, M. and Baidoun, S. (2003), "Understanding the essentials of total quality management: a best practice approach – part 1", Working Paper Series: No 03/05, University of Bradford, Bradford.
- Zairi, M. and Youssef, M. (1996), "A review of key publications on benchmarking: part II", *Benchmarking for Quality Management & Technology*, Vol. 3 No. 1, pp. 45-9.

Further reading

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

About the authors

John P. Moriarty is a Tourism Research Fellow in the Management School, Victoria University. His research interests include the transfer of academic knowledge to industry, benchmarking and organizational improvement.

Clive Smallman is Professor of Business Management, Commerce Division, Lincoln University. His research interests include decision making, practice and process studies of management, and simulation of social systems. Clive Smallman is the corresponding author and can be contacted at: smallmac@lincoln.ac.nz

En route
to a theory
of benchmarking

503
