A Theory of Benchmarking

A thesis submitted in fulfilment

of the requirements for the

Degree of

Doctor of Philosophy

at

Lincoln University

by

J. P. Moriarty M.E (ELECT), MPP

Lincoln University

2008

Abstract of a thesis submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

A Theory of Benchmarking

By

J. P. Moriarty ME (ELECT), MPP

Background: Benchmarking is a well established and respected mechanism that contributes to organisational improvement. Its epistemology demonstrates it to be theoretically underdetermined with literature focusing on pragmatism and praxis. Benchmarking's critics hold it to be *a-theoretical;* failing to provide its practitioners with a reliable basis for distinguishing between effective and ineffective efforts.

Purpose: To review Benchmarking's epistemology and identify the necessary or sufficient methodological elements contributing to its effectiveness and to establish them within an acceptable theoretical framework.

Approach: A causal approach is applied to the objectives of organisational benchmarking's current definitions and implementation frameworks. The resulting theoretical framework is then validated against current exemplary benchmarking praxis to explain its effectiveness and satisfy historical criticisms. Central to the approach is the application of supervenience and entailment relationships between benchmarking parties within the umbrella of Peircean Causation to determine the feasibility of a benchmarking proposition.

Findings: Benchmarking's *a priori* effectiveness *(*sufficiency) can be established from an organisational axiom and five logical conditions. This research establishes a new encompassing definition of benchmarking reduces its typology to a single consistent form and establishes an Effective Benchmarking Process that explains current practices and addresses historical criticisms. These logical conditions also explain the effectiveness of empirical frameworks such as the Malcolm Baldrige National Quality Award and ISO 9000.

Research Implications: A theoretical foundation for benchmarking provides a platform for extending the theory of organisational improvement.

Practical Implications: A theoretical foundation for benchmarking has significant potential to enhance organisational sustainability by reducing wasted effort.

Originality: This research focuses on the causal linkages between benchmarking and organisational sustainability. The research establishes a new definition of benchmarking, specifies necessary and sufficient conditions for its application and frames practitioner efforts within an Effective Benchmarking Process (EBP).

Keywords: Benchmarking; Theory; Organisations; Taxonomy; Causation; Sustainability.

Acknowledgements

I am pleased to acknowledge the support of many people and institutions whose generous contributions of time, resources, ideas and encouragement lead to this work.

Professor Clive Smallman, PhD, supervised this research and his reminder that ‘worthwhile scholarship rests on theoretical provenance’ arose at our very first conversation. This provided stimulus for the selection of this topic. I am truly grateful for his guidance, encouragement and co-authorship of publications.

Michaela Balzarova, PhD, associate supervisor, to whom I am grateful for her comments and encouragement on this thesis and publications.

Professor David Simmons, PhD, with whom I have shared much on tourism matters, is gratefully acknowledged for his thrice-strident insistence that I undertake doctoral study: sound advice!

My colleagues in the VMS, Victoria University of Wellington, have also contributed greatly through their stimulating discussions and interest. I thank them all most sincerely.

The librarians and administrative staff at Lincoln and Victoria Universities have also been most helpful and generous. In particular I thank Annette Brixton and Michelle Collings of Lincoln University and Linda Walker of VMS.

Mr Geoffrey Mansfield, ‘uncle-in-law’, has been a truly generous host during my numerous trips to Christchurch. I thank him most sincerely.

The love and support of my wife and family has been indispensable to the completion of this work. I thank them most profoundly.

I dedicate this work

to my wife, Helen and to the memory of my parents, Jack and Margaret Moriarty.

Table of Contents

Index of Figures.

Index of Tables.

Index of Equations.

Glossary of Terms, Symbols and Common Abbreviations.

Terms and Working Definitions	
Theory	An evolutionary construction that can be validated logically, epistemologically and paradigmatically to explain a generic set of presumed realities.
Model	A depiction or analogue of a particular set of presumed realities that represents or explains all or part of their behaviours via logical processes.
Exemplar	An organisation, system or process evincing a more desirable state of affairs than an anomalar.
Anomalar	An organisation, system or process evincing a less desirable state of affairs than an exemplar.
States of Affairs, S_A	States of affairs represent the status of organisational competencies (of A) at some point in time that can be gauged according to some consistent metric and establish the teleological trajectory or sustainability of an organisation. The most common consistent metric is economic – a measure of the ability to acquire and maintain the resources necessary for survival.
formal, material, efficient, final	The italicised form of these words refers to the four forms of Aristotelian causation.
Environmental variables	Elements that can be identified and gauged according to some consistent metric representing the status of resources deployed within an organisation and give rise to a state of affairs.
Welfare	A measure of the contribution a state of affairs makes to the survival of an organisation.
Effective benchmarking process	The feasibility of transforming at least one exemplary causal relatum (*efficient* cause, *final* cause) to an anomalar to obtain a possible improvement in the welfare of its state of affairs.
Perfectly effective benchmarking process	The feasibility of transforming all exemplary causal relata to an anomalar to obtain equality of welfare of its state of affairs with that of the exemplar.
Effective benchmarking improvement	An *efficient c*ausal process where the invariant antecedent of an improvement in the welfare of a state of affairs is an effective benchmarking process

Symbols	
$\rightarrow$	Always followed by, unconditionally implies, as in A$\rightarrow$B; A is always followed by B
{x}	Things similar to x or the Set of x things.
$\forall$	For all, as in 'For all' real numbers
$\exists$	There Exists
$\Leftarrow$	Cause of, as in A is the Cause of the Effect B; A $\Leftarrow$ B, or A is the antecedent of B
$\Rightarrow$	Effect of, as in B is the Effect of the Cause A; B $\Rightarrow$ A, or B is the successor of A
$\Leftrightarrow$	Reciprocity, as in A is the unconditional antecedent of B and B is the unconditional successor of A; A $\Leftrightarrow$ B
$\neg$	Negation, Not. As in $\neg${x} not things similar to x
$\wedge$	Conjunction (simultaneous presence) (causal 'and')
$\vee$	Disjunction (individual presence) (causal 'or')
$\in$, $\notin$	Is a member of, belongs to; is not a member of, does not belong to; as in -1 $\in \Re$, but -1 $\notin \aleph$
$\cap$	Intersection of proper sets
$\cup$	Union of proper sets
$\subset$	Proper subset of; A is a proper subset of B as in (A$\in$B) $\wedge$ (A $\neq$ B)
$\subseteq$	Inclusive subset of; as in (A$\in$B) $\vee$ (A=B)
$\supset$	Proper superset of ; A $\supset$ B = (B $\in$ A) $\wedge$ (A $\neq$ B)
$\Diamond$	Necessity; e.g. **of necessity** everything having property G has property F, i.e. $\Diamond \forall y\,(G_y \rightarrow F_y)$
iff	If and only if
$\blacktriangleright_{w,s,g}$	Supervenience, A supervenes on B weakly (w), strongly (s) or globally (g) if "A indiscernibility $\rightarrow$ B indiscernibility". See Appendix 2 for detailed explanation.
$\models$, $\vdash$, $\Vdash$	Entailment. Logically $\models$, Probabilistically $\vdash$, Dispositionally $\Vdash$ Entails, e.g. A $\models$ B iff all logical models of A are also logical models of B, where A and B are sets of clauses. See Appendix 2
$\Re$	The set of real numbers
Z	The set of Integers (-∞, …., -1, 0,1, 2, … ∞)
$\aleph$	The set of cardinal numbers (1, 2, 3, … ∞)
$^-$	Boolean NOT as in $\bar{\mathbf{a}}$
$\oplus$	Boolean Exclusive OR as in (**a** $\oplus$ **b)**
+	Boolean OR, as in (**a + b)**

Symbols, Continued	
•	Boolean AND, as in (**a • b)**
[A]	Matrix A
$[A]^T$	Transpose of Matrix [A]
∇	Vector Gradient operator, $\nabla(f(x,y,z)) = (\partial f/\partial x, \partial f/\partial y, \partial f/\partial z)$
$\sum$	Summation notation
$\prod$	Product notation
$\varnothing$	Null or empty Set
∞	Infinity
$\partial f/\partial x$	Partial derivative of function f(x,y,z) with respect to x
$\int$	Integration
$\lvert A \rvert$	Determinant of A
$\pm$	Plus or minus
$\therefore$	Therefore, hence
$\equiv$	Identical to
$\cong$	Similar to
$\neq$	Not equal to

Abbreviations	
AHP	Analytical Hierarchy Process
CFA	Common Factor Analysis
EBP	Effective Benchmarking Process: A practitioner-orientated benchmarking process framed by theoretically necessary and sufficient conditions
EFQM	European Foundation for Quality Management
EVA®	Economic Value Added, Stern-Stewart's term for the difference between net operating profit after tax and the capital charge on assets employed. A welfare-orientated financial measurement.
ISO	International Standards Organisation
MBNQA	Malcolm Baldrige National Quality Award, often referred to as 'Baldrige'
NZBEF	New Zealand Business Excellence Foundation (Baldrige-based)
PCI	Principal Component Analysis
PDCA	A Total Quality Management term for a 'Plan-Do-Check-Act' cycle of business improvement
SME, VLE	Respective abbreviations for Small and Medium Enterprise(s) & Very Large Enterprise(s)

Chapter 1. Introduction.

1 Why a Theory of Benchmarking?

Benchmarking has developed into a popular organisational improvement tool that has come to be regarded as an essential component of internationally respected business excellence programmes and good management practice. The simple concept of improvement based on observed or perceived exemplary performance elsewhere is not at all novel as it is observably a universal human trait. What is remarkable is the degree to which benchmarking has become associated with organisational improvement in the post modern era.

In this context it is also important to confine the locus of benchmarking to '*organisations'* since there are other loci, such as computer science, surveying and geology, where the terms 'benchmark' and 'benchmarking' may not always convey exactly the same meaning or sense of purpose.

Benchmarking's successes have been well publicised: Xerox Corporation was a spectacular example of benchmark-driven organisational rejuvenation when its first-mover advantage evaporated rapidly. There are also steady streams of less spectacular but important successes at all levels of business enterprise where the imperative of survival through continuous improvement is increasingly well understood by management.

However there are also many stories that tell of ineffectiveness as benchmarking does not always deliver successful outcomes despite consuming significant organisational resources. It appears incongruous that the conceptual simplicity of benchmarking can belie its efficacy. It might be tempting to attribute such failures to inadequate execution were it not for a growing body of empirical evidence that exhorts practitioners to be attentive to numerous pitfalls. Success can be elusive – even if reliance is made on the most explicit or pedantic of any one of at least sixty currently available implementation frameworks or by reference to one of seven current forms of benchmarking. Is this efficient? Can the risk of failure be established and altogether avoided, or at least mitigated?

Perhaps these incongruities have been subsumed by an understandable reverence for its considerable benefits? Criticisms have generally surfaced in mild or tentative forms, but more recently they have become strident. Why has there been a proliferation of various

types, forms and frameworks of benchmarking rather than convergence to one exemplary form?

Answers to these incongruities should lie within the encompassment of benchmarking's theoretical framework: but no such framework has been found. Literature demonstrates a broad and progressive fusion between the concept of benchmarking and numerous analytical systems and methodologies that were originally intended to support its implementation. Although practitioners may draw upon analytical support to discriminate the anomalous from the exemplary, there is no *a priori* analytical method distinguishing between effective and ineffective benchmarking efforts.

1.1 Aims and objectives of this Research.

Although absence of a theoretical framework does not appear to have impeded the use of benchmarking as an organisational improvement tool its critics deserve answers to their questions surrounding its theoretical provenance. Better knowledge of the essence of benchmarking would either assuage or confirm doubts as to its efficiency and effectiveness in a wide range of organisational settings and provide to those inexperienced with its use a more reliable basis for adoption than is currently the case.

Benchmarking's *a theoretical* nature provides the principal objective of this thesis: to address the essence of benchmarking and resolve its criticisms through the research question:

What is the theoretical framework for benchmarking?

Can benchmarking be encompassed by a single theory that provides distinction between effective and ineffective efforts and simplifies current complexities?

The aim of this research will be to examine benchmarking practices and associated literature under the working hypothesis that benchmarking's practices and epistemology embody causal elements, which if identified, might permit distinction between effective and ineffective effort and form the basis of a theoretical framework.

The potential benefits of establishing this popular organisational improvement practice within a theoretical framework are significant. Benchmarking efficiency will improve

substantially if it can be determined, *a priori*, that any proposed application is theoretically effective.

1.2 Theoretical considerations in Organisational Studies.

This research has the objective of establishing a theoretical framework for benchmarking, yet to do so it must also address two evident issues. The first issue is that acceptable benchmarking practices have been well-established despite the absence of a theoretical framework. The second issue relates to the controversial question: 'what counts as theory' in organisational studies?

In the first case, current organisational praxis that constitute 'benchmarking' is evinced by literally scores of frameworks or procedures representing the 'data' that an evolving theory must address and use in some validation process. Moreover, these 'data' are not the usual 'presumed realities' – since they have been criticised for their inadequacies. Within the context of this research, current organisational benchmarking praxis represents a 'dataset' of experimental procedures which have satisfied a sufficient number of practitioners to justify their continued use as an organisational improvement tool. As part of its validation, a theory of benchmarking must explain why such praxis contains both effective and ineffective elements and also identify a theoretical praxis that resolves extant criticism. Frankl (1992, 12) would call this 'existential' validation of a theory.

The second issue is the nature of an acceptable theoretical construction within organisational studies. This is controversial and will be addressed more fully in Chapter 4 where the aim will be determination of criteria against which an organisational theory can be validated – bearing in mind that any theory is only a representation of reality, not reality itself.

1.3 Approach.

Although benchmarking has an extensive epistemology to draw upon it is almost entirely pragmatic with focus on praxis rather than theory. Were it not for its celebrated successes, benchmarking would long be regarded as just another 'management fad'. This represents sufficient evidence to suggest that benchmarking practices do contain the elements necessary for successful application.

Chapter 2 will examine benchmarking's epistemology to commence the task of identifying these causal elements, known as 'causal relata', and Chapter 3 will examine the nature of causation to locate them within its established theories.
As this research seeks to establish a theory that justifies, simplifies, explains and extends organisational benchmarking, the nature of an acceptable organisational theory must also be addressed together with the criteria against which it is to be validated. Are all currently accepted organisational paradigms amicable to benchmarking or is its effective locus much narrower? Chapter 4 examines the nature of theoretical and '*a-theoretical*' constructions within an organisational framework to establish the necessary conditions that filters or validates one from the other. This filter or validator will then be applied to the emerging theory of benchmarking.

Benchmarking is also a teleological phenomenon: performed for the sake of organisational improvement. If benchmarking is to be effective, its causal relata must also contribute to organisational improvement. Identifying these causal relata and discriminating between those contributing to effectiveness, and those that do not, establish the locus of benchmarking within organisational teleology.

Chapter 5 defines the nature of organisational improvement and its relationship to benchmarking.

In Chapter 6, these causal relata are established and drawn together as 'The Theory of Benchmarking' ('The Theory') in the form of an organisational axiom and a set of logical conditions that identify potentially effective benchmarking processes and potentially effective benchmarking improvements. These conditions are further warranted and defended as a sufficient set of criteria that obtains effective benchmarking. It will be demonstrated that extent to which benchmarking is effective relies on the nature of two particular relationships between exemplary (exemplar) and non-exemplary (anomalar) states of affairs: i.e. supervenience and entailment.

It is essential that The Theory be amenable to its practitioners. Chapter 7 interprets The Theory in practitioner terms and establishes an Effective Benchmarking Process. This Process identifies a sufficient set of practitioner actions that can be reviewed in advance to assess the feasibility (theoretical effectiveness) of a benchmarking proposal. A further outcome of this Process is the reduction of benchmarking to a single consistent form.

Theories are judged by the acceptability of their results. Chapter 8 validates The Theory using the theory filter or validator developed in Chapter 4 to assess the degree to which it addresses current organisational paradigms, explains current empirical practices, extends knowledge and is logically complete. This chapter concludes that The Theory is a valid potentially theoretical construction whose ultimate acceptability is determined, over time, by exposure.

Chapter 9 reviews the contribution of The Theory of Benchmarking developed in this thesis to understanding and extending current knowledge of benchmarking. Benchmarking is concluded to be a mechanism that transfers power between two causal engines. Chapter 9 also proposes additional lines of research to extend the application of The Theory to organisational improvement.

1.4 Publications Arising.

All of Chapter 2 has been accepted for publication by Benchmarking: An International Journal, under the title 'En Route to a Theory of Benchmarking' and is currently in print (Volume 16, Issue 4). (Authors: Moriarty, J. P. and Smallman, C., 2009).

Parts of Appendix 1, in respect of its contribution to Hospitality Benchmarking via the application of DEA: 'Challenges for Hospitality: Beyond Price' has been accepted as a refereed paper for The NZ Tourism and Hospitality Conference, Hanmer, NZ, December 3rd -5th, 2008. (Author: Moriarty, J. P.).

This paper has also been accepted for publication in a revised form by Tourism Economics (expected publication date, June 2010).

Chapter 2. The Epistemology of Benchmarking.

2 Introduction to Organisational Benchmarking.

Benchmarking is a well-used modern term associated with a broad range of human endeavour. It is increasingly found to be an essential contributor to any serious organisational improvement process – where a current state of affairs are deemed temporary until replaced by a more desirable state of affairs - based on some evidence or expectation that such a state is attainable either in whole or in part. The locus of benchmarking lies between the current and more desirable state of affairs. It contributes to the transformation processes that realise these improvements.

Literature suggests that benchmarking definitions exist at two levels: primal definitions attempt to describe benchmarking in absolute terms and functional definitions attempt to describe it in operational terms.

The objective of this chapter is a review of the epistemology of benchmarking. A thematic approach will be applied to origins, primal and functional definitions of benchmarking, its organisational ontology and teleology, its limitations and encompassment by theoretical frameworks.

2.1 Origins and Primal Definitions of Benchmarking.

A variety of dictionaries cite the etymology of benchmark from words used circa 1842 to describe the surveying or geological 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 made on exactly the same basis.

Benchmark, as a noun, describes a point of reference and subsequently extended beyond surveying and geology into a spectrum of organisational practices where the analogy of a level plane is some level of organisational performance or achievement. It is also used in computer science to refer to the performance of reference software operating in a particular environment. Although this context is different to organisational benchmarking,

one author has included the concept of improvement within the ambit of computer science benchmarking and this particular aspect will be explored briefly in Chapter 2.4.3.

Benchmarking, an adjective, describes 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. It is within an organisational context that this study of benchmarking will be conducted.

Table 2-1 lists the results of a review of current literature on the common attributes of current benchmarking definitions.

Epistemological Attributes Defining Benchmarking	
Definitional Attribute	**Evolutionary Epistemology**
Feedback Mechanism	(Deming, 1986; Argyris, 1999; Kumar & Chandra, 2001)
Measuring and Improving (Continuously)	(Camp, 1989; Leibfried & McNair, 1992; Spendolini, 1992; Vaziri, 1992; Zairi, 1992)
Adaptation/Modification	(Zairi, 1992; Watson, 1993; Zairi, 1997),
Superior/Best Practice	(Camp, 1989; Bhutta & Huq, 1999; Kozak & Nield, 2001; Maire *et al.*, 2005; BNQP, 2008)

Table 2-1. Epistemological Attributes Defining Benchmarking.

Deming's (1986) theory of quality management emphasised the need for enhancing and sustaining production quality using feedback mechanisms as a means of behaviour modification. Feedback included both internal and external referents (benchmarks) of production quality. Spendolini (1992) generalised Xerox's successful in-house quality improvement process, further observing that as the need for improvement was continuous, benchmarking also needed to be both continuous and systematic in its evaluation of the attributes of the best practices of others so as to maintain organisational advantage.

Zairi and Ahmed (1999), noted that although benchmarking was a formal process used by Xerox Corporation in 1979 to improve organisational performance, the concept of observing a state of affairs and upon deeming it to be desirable and worthy of attainment is evocative of behaviours as 'old as humankind'. This definition establishes a fundamental aspect of benchmarking: it requires two parties. The *exemplar* is the party (organisation) evincing a desirable state of affairs and the *anomalar* is the other party seeking to approximate or attain that desirable state of affairs.

Organisational perspectives are also used to define benchmarking. Leibfried *et al* (1992) defined benchmarking as 'an external focus on internal activities, functions or operations in order to achieve continuous improvement'.

Others have approached the definition of benchmarking counterfactually by remarking on what it is not.

Watson (1993) says that the benchmarking concept should be viewed as a process of organisational adaptation, not adoption – not simply a question of copying others, but learning how to improve by sharing ideas. Zairi (1997) emphasises that benchmarking is not just a technique or a tool. Rather it is a powerful concept, a change agent whose impacts portend behaviour modification and developing new ways to manage business.

Neither is it simply competitor analysis, espionage or theft, but a process that establishes the ground for creative breakthrough, by identifying the highest standards of excellence for products and processes, and then making the improvements necessary to reach those standards by addressing the management and operational skills responsible for production (Bhutta *et al.*, 1999; Kozak *et al.*, 2001).

In contrast, Kumar and Chandra (2001) espoused a manufacturing industry perspective and claimed that that benchmarking could be viewed from the other direction and considered it to be a form of 'backward engineering' where the performance goals from other successful organisations are assumed to be achievable and applicable to others. This approach conflicts with the concept of benchmarking as a trigger or catalyst for organisational 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 identical resources and instructions/recipes produce relatively consistent outcomes. But it remains to be seen whether this particular approach can be generalised – especially in an organisational context where resources are seldom identically deployed and instructions are subject to human interpretation.

Xerox CEO, Kearns offered a pragmatic definition of the role of benchmarking in his organisation's environment: 'Benchmarking is the continuous process of measuring products, services and practices against the toughest competitors or those companies recognised as industry leaders' (Camp, 1989).

Benchmarking definitions can be observed to extend through the addition of meta-information to identify functional elements. For example, Watson (1993) views benchmarking as a continuous process that searches for and applies significantly better practices for the purpose of achieving superior competitive performance. Garvin (1993) extends this by adding meta-data to qualify *how* this occurs. Garvin noted that the continuous process was 'disciplined' and the search was 'thorough', incorporated a 'careful study' of one's own practices 'and performance', was extended through activities that included 'systematic' visits to exemplars and concluded with 'analysis' that produced 'recommendations and an implementation pathway'.

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

A Darwinian tone permeates these definitions: organisational 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 perspectives suggest benchmarking to be a teleological agent - a contributor towards organisational perfection.

Primal definitions of benchmarking offer little additional lucidity but explanations of its functionality are considerably more numerous. The next section examines benchmarking from a functional perspective and establishes its nature within an organisational context.

2.2 A Functional Approach to Benchmarking.

Primal definitions of benchmarking emphasise its contribution to organisational success through the principal process of organisational adaptation which is triggered by belief of 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, 1996; Yasin, 2002; Dattakumar & Jagadeesh, 2003).

In their review of benchmarking literature Dattakumar & Jagadeesh (2003) noted that 55% of cited benchmarking publications focus on applications, case studies, education, innovations and extensions with the remaining 45% focusing on models, general issues and fundamentals.

Clearly perceived to be beneficial to practitioners, publication trends from 1986 through 2002 indicate that benchmarking applications, case studies and models (including general review and fundamental topics) dominate the literature.

Moriarty and Smallman's (2009, in print) examination of Benchmarking: an International Journal (also known as Benchmarking for quality management & technology) over the period 1994-2008 is consistent with this trend.

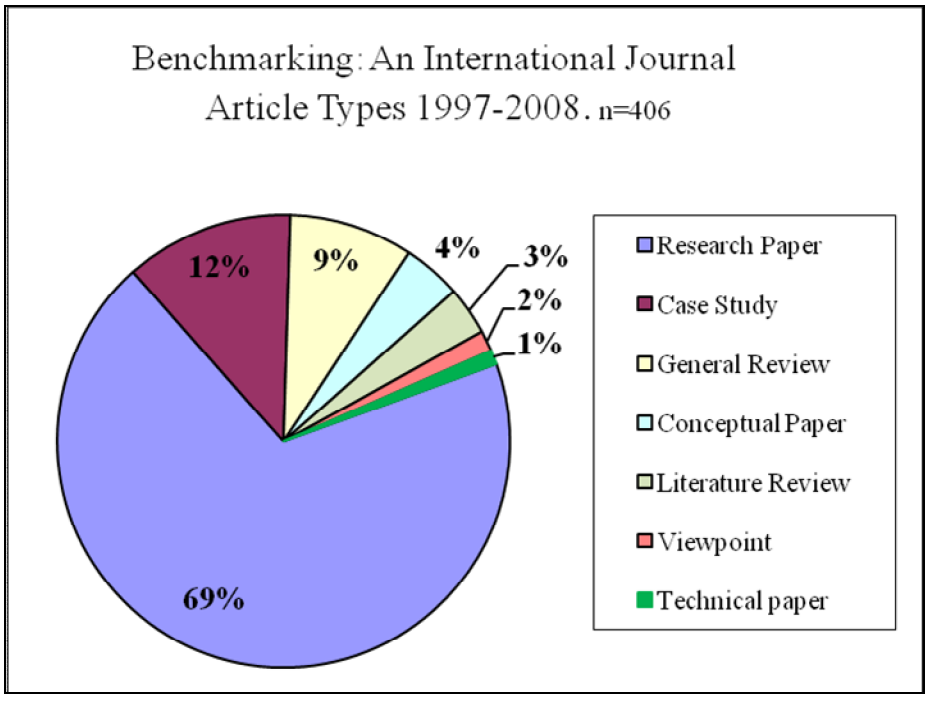

Figure 2-1. Distribution of Benchmarking Journal Article Types 1997-2008.

Figure 2-1 illustrates that out of 406 articles in the principal journal dedicated to the topic of benchmarking, the majority of publications (68%) were of a general research nature – empirical studies including models or framework testing of the role and application of benchmarking as an organisational improvement mechanism. A further 12% reported case

studies of the application of benchmarking techniques but only 4% of all publications were conceptual and none addressed the underlying nature of benchmarking.

Kozak and Nield (2001) also noted that there were approximately forty different models outlining the process of benchmarking: some originating from organisations and others from researchers and consulting agencies.

The organisational framework within which benchmarks and benchmarking appear is also diverse. By inference as to technique, benchmarking appears at strategic levels of understanding such as double loop organisational learning (Argyris, 1977) and knowledge management (McAdam & McCreedy, 1999). Benchmarking is explicitly referred to in the more tactical areas of total quality management (TQM), supply chain management (Deming, 1982; Zairi *et al.*, 2003), balanced scorecards (Kaplan & Norton, 1992), six sigma (Xerox, 1979), innovation (Radnor & Robinson, 2000), performance measurement (Carpinetti & de Melo, 2002; Anderson & McAdam, 2004), and the conjunction of TQM and business excellence models (Welch & Mann, 2001; NIST, 2007; NZBEF, 2007).

These contexts refine the definition of benchmarking to be: a search for industry's best practices that will lead to superior performance (Camp, 1989); a continuous and systematic process of evaluating companies recognised as industry leaders, to determine business and work processes that represent best practices and establish rational performance goals (Zairi & Leonard, 1994); a practice whose central essence is learning how to improve activities, processes and management (Ahmed & Rafiq, 1998).

Harrington *et al* (1995) note that benchmarking applies to processes and emphasised the systematic and evolutionary nature of its contribution to achieve superior production. This is consistent with Juran's (1950) view of an incremental route to superior production and that an important *few* rather than the trivial *many* factors should be the subject of attention. But Harrington (1995) also noted that the concept of a process could encompass more than just production components and include administrative components such as bureaucracy. This observation is echoed by Clarke *et al* (1997) where the process of change, instigated by production benchmarks, must itself be benchmarked so as to achieve superior change management.

Anderson and McAdam (2004) also resolved benchmarking into lead or lag components: being respectively, a predictive or reflective contribution to organisational improvement so as to address earlier recognised shortcomings where sole focus on historical perspectives may result in conflict if new strategies rely on old measures (Bourne *et al.*, 2000).

However, these perspectives also beg numerous questions. For example, in an organisation, what triggers the benchmarking process, how does one decide what to benchmark, ensure that the implementation is successful and finally, know whether what has been done is 'the best'? The next section examines benchmarking practices within an organisational framework.

2.3 Benchmarking Practices.

Latent issues raised in the previous section may be addressed through two questions. The first relates to the types of benchmarking that might be undertaken: do different organisational requirements trigger different types of benchmarking? The second is more practical: given a particular type of benchmarking, exactly how is it to be conducted such that it achieves the purpose of organisational improvement?

2.3.1 Benchmarking Taxonomy.

To address the triggers of benchmarking, we may start with Zairi's (1994) taxonomy that identifies its essential types: Internal, Competitive, Functional and Generic

- Internal Benchmarking. Intra-organisational exemplars of replicated activities provide a trigger for improving anomalous performance. Any element of an organisation 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 organisations where the absence of market forces may be replaced by a systematic comparison of best practices. Examples include 'branch' performance in distributed organisations, customer service performance between different service locations, and public sector organisations who share common stakeholders (hospital boards, government departments, etc). This form of benchmarking is an application of organisational learning where proven innovation

may be replicated without the external competitive constraints to improve overall welfare.

- Competitive Benchmarking. An organisation'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. Traditional candidates for triggering this re-evaluation have been observable customer-facing factors such as defect rates or process speed. This form of benchmarking is at 'arms-length'.
- Functional Benchmarking. An organisation's business practices are re-evaluated in the light of knowledge that non-competitor organisations demonstrate superiority in some common elements of business practice. This triggers re-evaluation of these business practices – often in partnership or in conjunction with the exemplar. Common elements such as the use of information technology, administrative or logistical processes permits co-operation between organisations since the risk of market-place competition is non-existent.
- Generic Benchmarking. An organisation's business practices are purposefully compared with organisations having demonstrably superior performance from similar practices of dispositions. Comparisons of exemplar practices or dispositions, 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 or dispositions associated with better performance. Practices 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. Similarly, dispositional factors such as quality (of service), timeliness (of production), knowledge, analysis, success (financial results) or leadership may also provide a broad basis for benchmarks between organisations in completely different sectors. Whereas the other forms of benchmarking can provide elemental comparisons, generic benchmarking provides factor-level comparisons. Practitioners are then required to augment their findings with other techniques such as root-cause or cause-effect analysis to identify elemental deficiencies. The Baldrige business excellence model is an example of a generic, dispositional framework that is empirically determined to be associated with exemplary performance (BNQP, 2008).

But Zairi's taxonomy may also be viewed as a process: a journey of increasing sophistication or adaptation. This journey commences with organisations using benchmarking to identify and replicate superior achievement through what Watson (1993) and Kumar and Chandra (2001) termed 'backwards engineering'. This concept extended to address broader inter-organisational 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 practices and applying any learning arising from doing so. Watson's early perspectives of this progression were extended by Ahmed and Rafiq (1998) and later augmented by Kyrö (2003), yet they remain within Zairi's taxonomy.

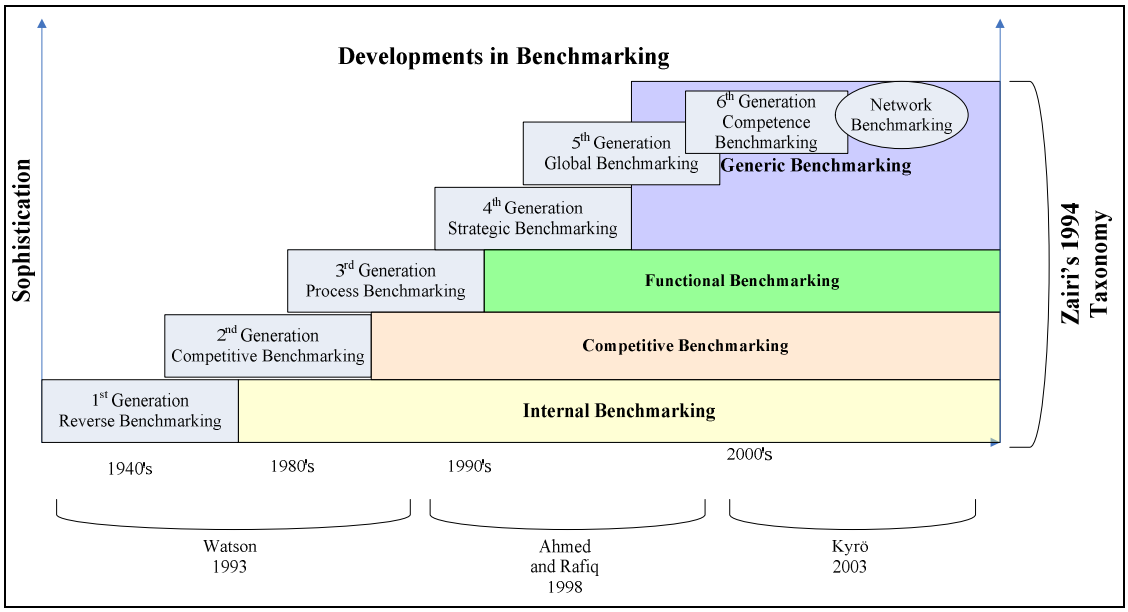

Figure 2-2. Developments in Benchmarking.

Modified from Watson (1993), Ahmed and Rafiq (1998) and Kyrö (2003)

Figure 2-2 combines these perspectives of benchmarking and the various generational developments that have arisen from the 1940's until the present. These generations generally align with Zairi's taxonomy and also expand the mechanisms that account for benchmarking functionality. Ahmed and Rafiq (p 228) identify key characteristics such as measurement, continuous improvement and systematic implementation as intrinsic

benchmarking attributes. However, as these are broad concepts common to most organisational endeavours, they further qualify them by noting that benchmarking involves:

- an understanding of pertinent theory (enabling),
- the ability to perform appropriate measurements (assessment) and
- achievement of results consistent with theory (outcomes).

Ahmed and Rafiq's 'theory or enabling' relates to the underlying principles governing the activities or processes of interest to the anomalar. '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 dimension the latent potential available for release in the 'outcomes' phase.

Kyrö extends the concepts and forms of benchmarking by noting that organisational developments suggest two additional generations of benchmarking based on its application to internal organisational learning processes and recognition of the dispersive nature of exemplars. For example, if activities (production, design, research and development) are dispersed not simply throughout a single organisation, but perhaps also throughout partner organisations, there is the additional challenge of performing the measurements and achieving outcomes consistent with process. Internal organisational learning itself then becomes an important enabler in applying a benchmarking process. These two additional generations add nuance to Zairi's 'generic benchmarking' and recognise the evolving nature of organisational behaviour.

These practices identify the scope of benchmarking. It is clear that, from its earliest formal concepts onwards, the tendency for organisations to seek assurance that factors with a significant contribution to overall success may be compared with similar factors elsewhere. This might be in competitor organisations, kindred organisations or in any analogous situation. The presumption is that theory underpinning the nature of comparable factors is both understood and congruent between the anomalar and exemplar organisations. Dimensioning 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 taxonomy of benchmarking. It is the observation that it is really immaterial who the exemplar is, provided there is congruence between its states of affairs and those of the anomalar. The taxonomy illustrates an expanding locus of likelihood as to where a suitable exemplar might be found – nothing more. The taxonomy does not address how an anomalar can benefit, *a priori*, from a benchmarking partnership.

2.3.2 Benchmarking Implementation.

All definitions imply that benchmarking is a process – a sequence of activities that involves theory, measurement and identified outcomes. 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, shown in Table 2-2.

	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**

Table 2-2. Benchmarking Relevance Table.

After Leibfried *et al.* 1992

'High' relevance accrues from the close matching of an anomalar's 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).

Xerox adopted what has become the familiar 'Plan, Do, Check Act' (PDCA) TQM process model that had its origins in Francis Bacon's (1620/2000) expression of the 'scientific method'. In modern times this method was extended into manufacturing and organisational practices by Shewhart (1980) and more notably by Deming (1986). PDCA-based benchmarking implementation models, exemplified in Figure 2-3, have been devised by Camp (1989), Spendolini (1992) and Drew (1997).

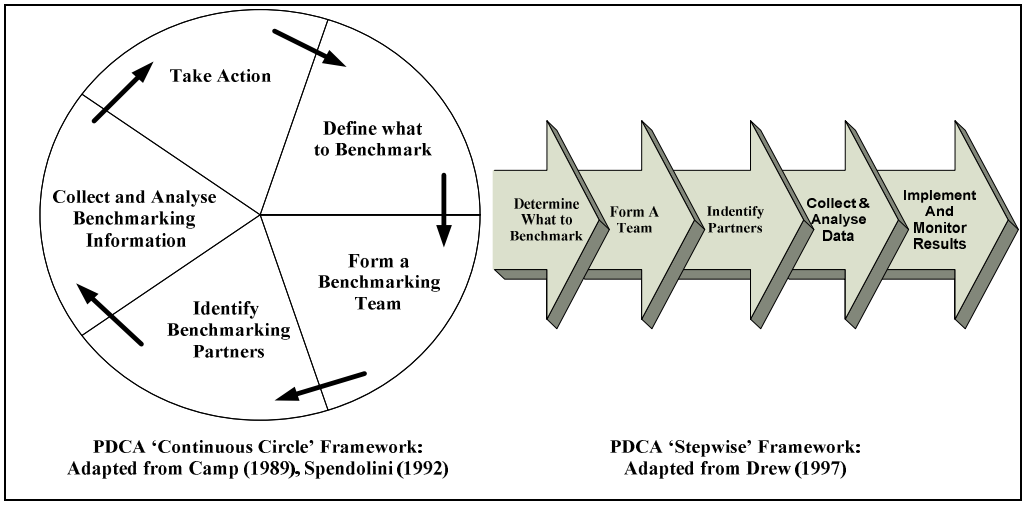

Figure 2-3. Benchmarking Process Models based on PDCA TQM model.

These models all have a five step approach for using benchmarking to achieve organisational improvement. The details within each step also expand and may contain highly-complex sub-processes.

For example, a suitably empowered leadership prioritises what is to be benchmarked. In a simple situation, a subjective prioritisation may suffice, but for complex situations where there are substantial numbers of related processes contributing to the selected organisational objective, prioritisation may be based on some form of analysis. Quantitative techniques include analytical hierarchy process (AHP) (Partovi, 1994), principal component analysis (PCA) and common factor analysis (CFA) (Büyüközkan & Maire, 1998). AHP requires relationships between organisational objectives and their associated processes to be quantified and subsequent sensitivity analysis determines benchmarking factors (or 'relata'). PCA and CFA are statistical approaches that rely on identification of process variances and inter-process dependencies to derive factors (benchmarking relata) contributing greatest variance to organisational objectives (such as customer satisfaction). These techniques are complex and the results may be difficult

to interpret. Carpinetti and De Melo (2002) suggest that systematic mapping and analysis of a wide range of contributory relata should be undertaken as the first step.

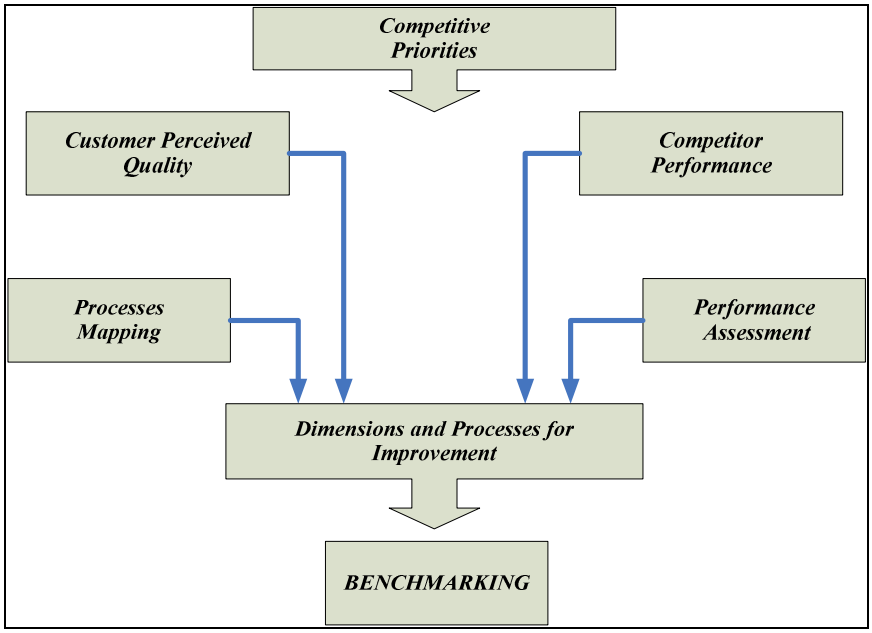

Figure 2-4. Defining what to Benchmark.

After Carpinetti and De Melo (2002)

Carpinetti and De Melo assert that contributory relata 'mostly contributing to efficiency and effectiveness of business processes mostly related to prioritised competitive criteria' determine what should be benchmarked. However, the 'analysis' of these relata still presents issues as relationships between them remain to be quantified Once the prioritised contributory relata are identified, additional processes are needed to identify exemplars and isolate and associate the corresponding contributory relata within them.

Zairi and Baidoun (2003) identify a 12 step benchmarking implementation model that was used by Yellow Pages (a division of British Telecommunications PLC).

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 & Plan Action.
10. Implement.	11. Keep in Touch.	12. Continuous Improvement.

Table 2-3. The 'Yellow-Pages' Benchmarking Implementation Model.

Zairi and Baidoun (2003, p 13)

The Yellow Pages model in Table 2-3 combines analytical and practical steps to guide the process. For example, steps 3 and 4 might be addressed using the analytical techniques described earlier. What is significant in this model is its extent. The PDCA

models of Camp and Drew are abbreviations of this model but they all identify the selection of benchmarking relata to be critical and also suggest that benchmarking implementation is no small effort, particularly for a sizeable organisation. 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 (1994) also warns that if process-driven benchmarking is undertaken (functional benchmarking), care must be exercised to involve 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 with emphasis on the adoption of a broad systematic approach towards selection of benchmark relata. Indeed, Partovi claims that failure to identify priority benchmark relata most probably invalidates subsequent benchmarking analysis. For example, output-driven (cost-driven) benchmarking might conflict with other organisational objectives such as customer satisfaction if the relationship between cost and quality is not well understood.

There is a common thread throughout these benchmark implementation models: the need to identify:

- priority relata that impinge on organisational performance,
- relationships between these priority relata and other organisational processes,
- exemplars with sufficient similarity to trigger improvement initiatives,
- capacity to implement improvements.

A methodology informed by the prior feasibility of a benchmarking proposition would reduce the both the implicit and explicit risks associated with implementation.

2.4 Criticisms of Benchmarking.

Organisational improvement is the principal objective of benchmarking, yet this clearly desirable objective is not without significant criticism. Some of the criticisms implicitly arise from difficulties in obtaining reliable exemplar information and difficulties inherent in achieving the organisational 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 implementation but rather expand the circumstances in which it might be applied. Expanding typologies is akin to violating

Occam's razor – *'entia non sunt multiplicanda praeter necessitatem*'. Multiplying entities without simplifying their application invites the criticism that benchmarking may be a 'fad'. It is useful to examine these criticisms for any common themes.

2.4.1 Information-oriented Criticisms.

Benchmarking requires an exemplar. 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 attempting to gather information describing exemplar advantages, often covertly, may fail to focus on their own unique situation and become prone to distraction and misdirection. Moreover, benchmarking is always a retrospective process: previous history 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 such as employee time spent gathering comparable data, whilst traceable, might be overlooked when determining cost-benefits of benchmarking.

A common theme through all information-orientated criticisms is that benchmarking is intrinsically retrospective and may even be inefficient. States of affairs examined (or adopted) by the anomalar will be historical and may also be disassociated from the exemplar's future purpose (its teleology). This is converse to the claim of uncompetitive homogeneity, since an exemplar successfully evolving its states of affairs in pursuit of some undetected future purpose will provide anomalars with uncompetitive benchmarks.

2.4.2 Implementation Criticisms.

Watson (1994) cited two major difficulties in implementing a rigorous benchmarking study: deciding what project to focus the benchmarking resources on and then what organisations to solicit as partners. Simply applying Zairi and Baidoun's 12 steps or the 5 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: slavish adherence to some formula is unreliable. Implementing a successful benchmarking programme requires more than adherence to the step-wise programmes outlined in Section 2.3.2.

Perhaps the root cause of the inability to implement benchmarking with the prospect of a more certain outcome lies deeper. Wolfram Cox *et al*. (1979) lamented the absence of a sufficiently developed theory that would explain the differences between effective and ineffective efforts. Wöbler (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 the year 2000. The application of non-parametric frontier analysis techniques such as Data Envelopment Analysis (DEA) *(*Charnes *et al., 1978;* Banker *et al., 1984*) and stochastic techniques have been used to examine and rank the technical efficiency of sets of 'decision-making units'. This provided a means of identifying appropriate exemplars and of quantifying the optimal parameters that 'might' elevate less technically efficient anomalars to that of the exemplar. Non-parametric data is particularly desirable as almost any data reflective of the performance of the enterprises under investigation suffices. For example, financial data may be combined with production and customer data to form an 'efficiency frontier' – the locus exemplars deploying different levels of resources but still achieving a relative operating efficiency of 100%.

However, these extremely popular techniques should not be applied without close regard to the circumstances. For example, 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 & Vickers, 1993; Smith, 1997). In general, DEA is sensitive to uncertainties in parameter values, and noisy data is interpreted as a contributor to efficiency and thus affects the relative efficiency of the corresponding organisation (Färe *et al.*, 2000). Longitudinal analysis of enterprises is problematic and requires special techniques such as Malmquist's Productivity Index (Bjurek, 1996) which compares datasets representing different time periods. DEA requires that all anomalars and exemplars in the dataset exhibit similar returns-to-scale behaviour– a particularly important factor when organisations from different industries,

marketplaces or sizes are compared or in the situation where some data simply cannot be scaled (e.g. constrained resources such as land may be un-scalable).

Elnathan *et al.* (1996) also noted that implementation programmes that failed to win the trust of the exemplar, including disclosure of sensitive items such as 'cost structures', were less likely to be successful. An inability to get to the heart of exemplary performance is a key implementation impediment and diminishes the degree of success that might be anticipated from any programme. What follows from this is some *a priori* view as to the level of information needed so that a benchmarking implementation programme can be properly gauged before resources are deployed.

In summary, implementation of a benchmarking programme can be far from straightforward. Although greatly assisted by the application of popular analytical techniques such as DEA, implementation is more likely to be successful if conducted by experienced practitioners who can navigate around the practical issues that are certain to arise and have secured a considerable degree of trust from the exemplar so that sensitive performance–contributing factors can be determined.

2.4.3 Theoretical Criticisms.

The acceptance of benchmarking has resulted from its widespread use (Francis *et al.*, 2007) and by dint of this it has escaped becoming 'another management fad'. Yet there is little literature that focuses on the theoretical composition of benchmarking. Several publications have included a 'theory of benchmarking' to support other work and it is useful to examine them to determine if they can truly address current criticisms.

Although computer science benchmarking is different from organisational benchmarking, one researcher has cited a theoretical framework based on the advancement of normal science and for this reason it is useful to examine its relevance. Cox's PhD dissertation (2003) examined benchmarking in the context of computer science where a benchmark is a generally accepted reference against which various computer technologies are compared. This reference might be a set of tests that are triggered by some motivating reason for comparison, operate on a representative sample of the relevant technologies and have performance outcomes (Cox, p 28). The sets of tests might involve a suite of data that is required to be manipulated in some manner.

The algorithm that manipulates the suite in the fastest and most complete manner determines its status as an exemplar.

The foundation of Cox's benchmarking theory is Kuhn's (1996) 'structure of scientific revolutions'. Benchmarking is claimed to 'operationalise 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 p35). 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 organisations in general. Practices that have provenance in respect of economic performance are triggers for implementation by those aware of them, but who 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 by Cox. 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 want to equal or better them. Cox does not identify any *a priori* method of achieving a set of tests (equating to organisational methods) that will trigger enhanced performance. This is most clearly demonstrated in Cox's (p42) reference to retirement of a benchmark: 'punched cards and vacuum tubes have become obsolete, so that standardised evaluations for them are not needed'. The question is: 'what caused those earlier benchmarks to exist in the first place'? Is that cause also obsolete, or has it simply been re-directed to more appropriate technology?

Cox has demonstrated that the benchmark is consistent with Kuhn's theory of scientific revolution (either it is accepted or it is not), but has not shown the process, i.e. how the algorithms can be transformed from anomalars into exemplars as a result of applying the benchmark?

The framework of organisational learning was chosen by Liang (2004) as a means of establishing a theory of benchmarking. He cites organisational learning as the 'effective processing, interpretation of and response to, information both inside and outside the organisation' (Easterby-Smith *et al.*, 1999). The link to benchmarking is further developed through Huber's (1991) assertion that 'an organisation learns if any of its

units acquires knowledge that it recognises as potentially useful to the organisation'. Moreover, people learn either from their own experiences or from others. Liang develops the perspective that the ability to take advantage of others' experience 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 gives 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 organisation and is encompassed by Zairi's taxonomy of benchmarking described earlier. But if benchmarking is a method of 'learning how to learn' (Liang, p 24), the issue is how this learning occurs and whether is it describable in a manner that distinguishes between effective and ineffective efforts.

In a pertinent passage, Huber (pp 96-99) refers to the ways in which second-hand experience is acquired by an organisation. 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 where budgets or expectations are not met) and performance monitoring (measuring and learning from errors) are cited as techniques that abet organisational learning. Whilst each of these techniques describes a mechanism for increasing organisational 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 organisation was chosen to be an exemplar. However there is little doubt that learning from exemplars can increase organisational performance, but the citations claim that sometimes it does not. Why?

Van Helden and Tillema (2005) 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 economic efficiency and institutional reasoning (p 338). They hypothesise that benchmarking can be viewed as a substitute for the attributes of market-force behaviour with economic efficiency driven

by innovation and customer satisfaction within a competitive environment. Benchmarking is framed by a series of hypotheses that reflect normative economic behaviour within public sector organisations. Inefficient performance, whether real or perceived, can attract remedial action from higher authorities (e.g. government) and may threaten survival as organisations might be devolved into more efficient entities (as was the case in New Zealand where solely highway police were merged with the national Police Force in 1992). Their economic hypotheses are quoted at length:

1. 'benchmarking will improve the average performance of organisations,

2. benchmarking is a stronger incentive to improve performance for poorly performing organisations than for better organisations,

3. benchmarking will diminish performance differences between organisations' (van Helden and Tillema, p 341).

These economic hypotheses are consistent with established benchmarking literature already reviewed, but the third of these hypotheses touches on the matter of organisational homogeneity – the subject of a previous strategic criticism that portends uncompetitive heterogeneity (see 2.4.1). In the public sector, homogeneity implies relative technical efficiency (Farrell, 1957). This is paramount as these organisations operate under a legal monopoly to provide public goods and service on behalf of the community. In the private sector, both technical and allocative efficiencies are paramount since being the best within a poorly performing class of providers (relative technical efficiency under equally priced resources) may trigger consumers to abandon such production in favour of alternatives that are even more technically (improved production) or allocatively efficient (improved resource prices). For example, locally manufactured products might be replaced by functionally identical but less expensive products manufactured elsewhere.

In contrast, institutional reasoning theories presuppose that organisational behaviour is determined by different types of institutional pressures exerted by, for example, government, and interest-groups, the general public and professional groups (van Helden and Tillema, p 343). They further delineate between neo-institutional theory's conformist behaviour and resource dependence theory's emphasis on all kinds of non-

conformist behaviour (van Helden and Tillema, p 343) to form a set of hypotheses that predispose public sector behaviour towards benchmarking as a means of achieving improved performance and at the same time conforming to accepted norms. Similarly, the institutional reasoning hypotheses are also quoted at length.

4. 'the higher the degree of social legitimacy perceived to be attainable from engaging in a benchmarking project or from improving performance, the more likely it is that a public sector organization will engage in a project or improve performance ('cause');

5. the higher the degree of economic gain perceived to be attainable from engaging in a benchmarking project or from improving performance, the more likely it is that a public sector organization will engage in a project or improve performance ('cause');

6. the less the stakeholders of a public sector organization disagree about engaging in a benchmarking project or about the nature of a desired performance improvement, the more likely it is that the organization will engage in a project or improve performance ('constituents');

7. the more a public sector organization is dependent upon the stakeholders that exert a pressure to engage in a benchmarking project or to improve performance, the more likely it is that the organization will engage in a project or improve performance ('constituents');

8. the less a pressure on a public sector organization to engage in a benchmarking project or to improve performance conflicts with its organizational goals, the more likely it is that the organization will engage in a project or improve performance ('content');

9. the smaller the extent of discretionary constraints imposed on a public sector organization by a pressure to engage in a benchmarking project or to improve performance, the more likely it is that the organization will engage in a project or improve performance ('content');

10. the more a higher governmental authority or the legislator plays a part in a pressure to engage in a benchmarking project or to improve performance, the more likely it is that a public sector organization will engage in a project or improve performance ('control');

11. the higher the number of organizations within its organizational field that have decided voluntarily to engage in a benchmarking project or to improve

performance, the more likely it is that a public sector organization will engage in a project or improve performance (‘control’);

12. the higher the level of uncertainty in a public sector organization’s environment, the more likely it is that the organization will engage in a benchmarking project or improve performance (‘context’);

13. the higher the degree of interconnectedness among the organizations within its organizational field, the more likely it is that a public sector organization will engage in a benchmarking project or improve performance (“context”)’ (van Helden and Tillema, pp 344,345).

These thirteen hypotheses address why benchmarking might be adopted by public sector organisations, or for that matter, any other organisations within an interconnected environment. It is also difficult to identify elements of these hypotheses that are exclusive to the public sector as it could be claimed that participants in, say, the Malcolm Baldrige National Quality Awards (NIST, 2007) are similarly interconnected thorough a desire to conform to a reputable system of documented principles and practices that not only include benchmarking, but are generally deemed to increase organisational performance in the broadest possible manner.

Economic hypotheses re-confirm existing benchmarking provenance: improvement in the economic performance of organisations. Institutional reasoning hypotheses contribute to the understanding of the relationship between benchmarking and organisational purpose. In fact, the case study described in their paper cites an adaptation of a ‘balanced scorecard’ (van Helden and Tillema, p 345) as the performance measurement instrument within which the previously described issues of identifying what to benchmark and how to do so are embedded. There is merit in these hypotheses as to the reason or purpose for embarking on a benchmarking programme since the common thread pervading each of them is improved welfare: better performance, conformance with exemplary behaviour, peer esteem and survival.

The fundamental issue arising from their hypotheses is that they provide a framework for benchmarking utility, not a theory of benchmarking. Their hypotheses establish a strong case for ensuring that any benchmarking theory should encompass economic efficiency and organisational teleology.

2.5 Summary and Conclusions.

This synopsis of benchmarking traces the emergence of a practice that has over the past forty years become identified as a strong contributor to organisational 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 advanced 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 which is *a priori*, or at least prior, distinction between effective and 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.

Literature is overwhelmingly pragmatic – process-driven, case-oriented and generic as opposed to theoretical. Where theories are invoked, they centre on the utility of benchmarking in terms of organisational learning and reasoning as well as economic enhancement. Attachment to, or association with, organisational 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, literature warns of the need to furnish a benchmarking programme with ample financial, technical and leadership resources so as to minimise 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 welfare.

There is, however, a single recurring theme: improving organisational welfare (both short and longer term) is emphasised. The nature of this is twofold: organisational survival is a

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 organisational or elemental levels, is certainly held to be desirable, but does this beg the question at hand: what is the concept of organisational 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 organisation 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 organisational state of affairs than is currently the case. Moreover, benchmarking might identify the changes necessary to achieve that end. This suggests that benchmarking is within an organisational teleology and its purpose associated with the ontological question “what is the essential nature of the organisation”?

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 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 organisational betterment. It requires an anomalar as well as real or notional exemplars 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. It requires understanding of the nature of the mechanism driving such change. Superior states of affairs would seem to be a necessary component of organisational survival in a competitive environment. Even where organisations operate in environments that are not subject to open market competition (examples often cite the public sector), efficient performance is ultimately intolerable to stakeholders and funders. No controversy would arise from the perspective that organisational survival is predicated on the achievement of increasingly better states of affairs – that is, survival is teleological.

These perspectives suggest a provisional definition of benchmarking:
"**Benchmarking is an exemplar-driven teleological process operating within an organisation with the objective of intentionally changing an existing state of affairs into a superior state of affairs".**

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

It is concluded that a theoretical foundation for benchmarking should be consistent with current organisational paradigms, causation theory and the nature of what constitutes current and superior states of affairs.

In short, benchmarking exists, but its essence is obscure.

The next chapter examines the contribution of causation theories to benchmarking – particularly the contributions classical (Aristotelian) and modern (principally Humean and Peircean) theories to explain the mechanisms of advancement from a current to a superior state of affairs.

Chapter 3. Developing a Causal Framework.

3 Introduction to the Development of a Causal Framework for Benchmarking.

The key element of the definition proposed in Chapter 2 is purposeful change from one state of affairs to another. This bespeaks causation or the evidence of an effect. What causes change to occur, what are its causal relata, are there issues of supervenience[1] and entailment between cause and effect, is cause prior to or simultaneous with effect and are there universal laws that govern the process of change? Can knowledge of causation address one of the most important criticisms levelled against current benchmarking practices: the inability for practitioners to distinguish between effective and ineffective efforts?

This chapter examines the development of causation theories from ancient to modern times and relates important milestones to the practice of organisational benchmarking. It will conclude that the application of causation theory assists practitioners to distinguish between effective and ineffective benchmarking efforts.

3.1 What is Causation?

The mechanism that gives rise to an effect that we sense or experience is held to be its cause and causation is its epistemology. However causation has a long and fractious philosophical history. The ancient Greeks envisioned an elemental world comprised of earth, air, fire and water where relationships between them accounted for all that was sensed by man. But medieval and renaissance philosophers driven by ontological issues advanced beyond this purely elemental world yet retained many of the forms, arguments and causal classifications of the ancients. In the eighteenth and nineteenth centuries further developments occurred with rationalist and empiricist schools holding different views that in many ways have yet to be settled. Charles Peirce (1898/1992) cautioned that those making causality an ancient universal element or fundamental category of thought have to contend with the fact that a proposition at one period of history may be entirely

[1] Supervenience: a holistic, nomological relationship between two sets of properties {A} and {B} such that there can be no changes in {A} without there being changes in {B}. Entailment: a logical relationship between two sets of facts {A} and {B} such that, everything that makes {A} true also makes at least one instance of {B} true. See Appendix 2 - Supervenience and Entailment.

different in another. However, each contribution to the theory of causation is relevant to an understanding of the mechanisms of change. This work contends that it is also relevant to the human and organisational concept of benchmarking. Contributions that lend themselves to an interpretation of benchmarking practice within an organisation seeking to effect purposeful change will be briefly explored within each historical context and an overall synthesis will combine with a review of organisational change processes in Chapter 4 to form the basis for a theory of benchmarking.

3.2 Ancient Causation.

In retrospect, ancient philosophy may often appear naïve but much was attempted (Bertrand Russell, 1979) and its luminaries established western scholarship through the first recorded insights into cosmology, mathematics, logic and science. Advances in science have long outstripped any notions that Greek philosophers might have entertained, but their views on human, societal or organisational matters are, arguably, still pertinent today. Aristotle's metaphysics is particularly pertinent to benchmarking and his conceptual (rather than empirical) perspectives raise issues that are remarkably absent from its epistemology.

3.2.1 Plato.

The Timaeus (Plato, p 455) dialogues attributed to Plato the first articulation of causality – 'for it is in every way impossible that anything should be generated without a cause'. Moreover, cause had an 'artificer' whose determination was often difficult (Plato, p 455-6). This articulation was preceded by inferences from the Greek notion, expounded by Empedocles, that the universal elements of earth, air, fire and water required two further mechanisms, friendship and discord, respectively drawing elements together or to keeping them apart; i.e. changing them (Diogenes_Laertius, XII).

3.2.2 Aristotle.

Plato's formality was all encompassing and offered no insight into underlying mechanisms. Aristotle identified this incompleteness and approached the concept of causation via the question 'why'? He isolated four different ways of exhausting the question 'Why is this so?' Aristotle (VI, Physics II 194b16 p 332) called these the *material* (species), *formal* (archetype), *efficient* (primary source) and *final* (sake) types of explanation, or causes that relate to knowledge of something.

Material cause identifies that from which it came: its components or constituents.
Formal cause relates to the pattern or form of something: governing principles or laws.
Efficient cause is that from which something starts: its primary source or agency.
Final cause is the sense of end or sake for which something is done: teleology.

Generally these 'causes' occurred in neatly separated situations. For example, a person walking (*material* and *formal* cause) from one place to another might do so not simply to change location (*efficient* cause) but for the sake of exercise - or for the sake or desire of 'good health' (*final* cause). Yet Aristotle recognised that *formal*, *efficient* and *final* causes often coincided, such as when 'man generates man' (Physics, II, 198a 25 p338). Perhaps this coincidence, later amplified by a European theological ontology of substance was the genesis for an often-held view that there is really only *efficient* cause since this most commonly identifies the transition of a form or material from one state to another in some way to produce an effect.

Causation dominated by *efficient* cause will be shown to be incomplete and restrictive in advancing the epistemology of organisational benchmarking. This claim will be addressed at the end of this chapter when the metaphysics of Charles Peirce is examined.

This focus on *efficient* causation also gave rise to a mechanistic and empirical process of enquiry into the nature of the transition: what initiates it, what controls it and how does it behave? Enquiry was also hierarchical process. The question 'why' is posed firstly with reference to the matter, the form, to the primary moving cause and the very sake for which it was done – 'what comes to be after what, and so on at each step of the series' (V1, Physics II, 198a 32, p 338). However, this series was certainly not infinite and ended in a first term simply because to the empiricist, infinity was unthinkable (V2, Metaphysics, II, 994a) (Bunge, 1959).

As to the sequence of cause and effect, Aristotle's description of *efficient* cause indicates that cause either precedes or is simultaneous with the transmission of an effect into reality – its production, and '*separates the producer from the produced*' (V1, Physics II, 195b 16-21, p 334). This also raises the question of reciprocity. If the producer is separated from the produced, can they be reconstituted? That ice is produced from water, is not water, yet changes back to water is a case where reversibility is well known

and fitted neatly with current elemental theories. However a fire produced from wood is clearly not a reversible change process and highlighted the need for a theory of substance.

Aristotle's ontology of substance, a combination of matter and form, restricted the extent of *efficient* cause (V2, Metaphysics VII 1028a 30-35, p 1623). In his view there could only be relationships of whatever kind as a result of substances which can be related and there can only be events because of substances behaving in a certain way (V1 Categories 1b 24, p 4ff).

Aristotle believed that the form of a substance established limits or boundaries on the extent of possible change. For example, form may not establish the functionality of a particular person, but form may establish what a particular person cannot do – such as breath underwater or fly – although other forms might.

The concepts of spontaneity and chance also were forms of *efficient c*ausation (V1, Physics II 195b 31, p 334.). Yet chance was as we might perceive coincidence – the accidental alignment of events – resulting from an intention to achieve an effect in one manner, but achieving it by another. Since chance required deliberation and choice, it only applied to (adult) human beings as '*only that which was capable of action could do things by chance*' (Physics II, 197b 6, p 337). An example of this might be an engineer wishing to replace a particular component but chances upon a substitute and (being deliberative) chooses it in preference to the original intention. Spontaneity applied to beasts and inanimates where it described either habitual behaviour or natural laws (e.g. the tendency for things to fall downwards).

Hulswit (2002) observed that it was unclear whether Aristotle aligned with the modern idea that there is a necessary relationship between causes and their effects. Given a certain effect there must be some factors that brought it about, but Aristotle did not infer that the existence of certain conditions necessitated some effect.

A final important concept in Aristotle's epistemology of causation is the distinction between 'being' and 'essence'. The fact that something 'is' (exists), differs from 'what it is' (essence). This distinction operates plainly in both organisational and non-organisational cases. For example, we know 'what a Dragon is' (its essence), but of

course they do not exist. Furthermore, we also know that many sub-atomic particles exist, but their essence is unclear.

3.2.3 Summary of Ancient Causation.

The relevance of Aristotle's theories of causation to benchmarking lies in the four distinctions *formal*, *final*, *efficient* and *material*. In keeping with Aristotle's chain of enquiry, we might think of benchmarking in terms of *final* cause – contributing to the *sake for which something is done*, and we might also regard it as an *efficient* cause for changing organisational substances (events and practices) to achieve the effect *caused* by the adoption of some better knowledge. We might even recognise that not all forms or material substances are identical and serve to differentiate possibilities for change on that basis. Indeed, Aristotle was most particular in observing that form was '*preserved*' in change insofar as formal attributes could not be inconsistent with the transformation. Recall that the form of a man might not determine all that a man can do, but it determines what he cannot. Might this concept apply to benchmarking one organisation against another?

However, Aristotle's causation precluded a *necessary* relationship between a set of certain conditions and an effect. In fact, Aristotle's preclusion should signal a sound warning to an anomalar. If an exemplar displayed a particular and desirable state, clearly it was caused by something; however, whatever that might be did not necessitate its achievement elsewhere. The concept of necessity was not settled easily and remains a vexed question even today as it is by no means easy to isolate all causal relata and thereby duplicate them in another situation.

The distinction between 'being' and 'essence' is developed further by middle age philosophers but Aristotle's dichotomy is important to the benchmarking process as observing a desirable (being) organisational state does not necessarily convey any knowledge of exactly what it is (essence). This is *ex post facto efficient* causation and is not entirely without merit in organisational contexts where the cause of a particular organisational state may be either unknown or unobservable (Smallman, 2007).

The *formal* and *material* causes also relate to benchmarking as they offer a mechanism to refine relevance and determinacy (supervenience) between the exemplar and anomalar.

As shall be evident from the mechanics of causation, not all exemplars are relevant to anomalars.

To summarise, Aristotle's answer to the question 'why?' resulted in four streams of enquiry underpinned by his substance ontology and an empirical teleology. *Efficient* cause does not satisfactorily encompass causation in respect of benchmarking as there are evidently other components such as *form, structure and sake* that certainly appear to motivate and be associated with purposeful organisational change. Yet the uncertainties of causal necessity, implausibility of chance and a substance ontology limited by undiscovered knowledge compels the historical review of causation to continue onto the Middle Ages.

3.3 Middle Ages Causation.

Aristotle's epistemology of causation appeared complete and when integrated into the Stoical theses of fate and regularity provided a unified theory of nature to pre-CE Greek and post-CE Roman scholars. Hulswit (2000, p 6) and Bunge (1959, p 99,134) summarise the relationship between causation and Stoic philosophy as systematic – linear chains of successive cause and effect, fate-driven (necessity) and universal. Nothing (substance or events) could occur without *cause*, otherwise the universe lacked coherence and *effects* were a necessary and exceptionless outcome of a particular set of circumstances. Not without argument or debate, the Stoic philosophy endured for centuries until Christian theology addressed the origin of fate and regularity.

3.3.1 Thomistic Causation.

Thomas Aquinas (1225? – 74) extended Aristotle's concepts of matter and form and extended *efficient* cause into the concepts of primary and secondary categories and internal and external causes. Concerned with the application of an Aristotelian and Socratic scientific methodology to theology, Aquinas' causality was ontological and hierarchical: *efficient* cause was a way of 'proving' the existence of God – the exemplar (Montagnes, 2001, p 38) – from the perspective of necessity (everything must start somewhere) (Montagnes, 2001, p 31), agency (someone must start something) and finality (there must be a sake for which things are done).

Refinement of causality into a three-level hierarchy was a consequence of the distinction between the natural and divine forms of reality (Montagnes, 2001, p 36; M. Hulswit, 2002, p 9). These levels are 'the generic, the formal and virtual' divisions or, from a substance ontology, the 'material, spiritual and divine'. Recalling Aristotle's substance ontology involved 'being and essence' of material things, Aquinas's substance ontology also extends to the divine where, in that case, essence and existence are one.

Leaving aside the arguments relating to the divine, there are others that provide insight into generic or material components of causation. For example, Aquinas offered a clearer relationship between *efficient* and *final* causation than Aristotle and held that there was a relationship between cause and effect – a natural necessity rather than spontaneity (Aquinas, II: 35.4). This extended the concept of causality and, according to Hulswit (2002, p 15), the conception of *efficient* cause as a means to an end influenced the modern view of causation considerably.

Aquinas, however, stood firmly with Aristotle on the phenomenon of chance. The universe was ordered, its order known by God (Kenny, 2002) and chance was precluded. Where other phenomena such as decay or deterioration occurred in nature, Aquinas held these to be manifestations of Aristotle's *formal* cause – as all natural substances inherently decay.

The ontological nature of Aquinas' philosophy constrained the need for explanation or inquiry beyond a certain point. Davis (1993) quotes Aquinas scholar Herbert McCabe (1987) on Thomistic explanation: 'once you have found the cause there is no further question about why this cause should produce this effect, to understand the cause is to just understand that it naturally produces this effect'. Whilst Aristotle's mode of enquiry appeared limitless, 'what comes to be after what', it presupposed finality. Aquinas held that the finality of causal sequence was the exemplar, God.

A weakness of Aristotle's relationship between cause and effect was addressed by Aquinas and provides insight into the role of a benchmark. Aristotle and Aquinas viewed the exhaustion of causal sequence slightly differently. Aquinas held God to be the exemplar and first mover and Aristotle held that there was some beginning to *efficient* cause, but their differences lay in the journey. The Thomistic concept of necessity, evinced by laws of a natural kind operating to translate a particular situation

into an effect should have suited Aristotle, except perhaps at the start of the journey. The concept of circumstances necessarily determining effects was an advance on Ancient concepts of causality – even though explanation via 'laws of a natural kind' remained elusive.

However the concept of necessity operates reliably within a modern organisation's systems and processes. Indeed, most commercial software systems, and all financial systems, depend on the principle that identical (necessary) outcomes always result from identical inputs (circumstances). Aristotle and Aquinas held that certainty rather than probability governed the outcome of an assembly of criteria (substances and events), and in many cases this translates favourably into a *singular* benchmark paradigm where *identical* states of affairs and governing systems exist between organisations. In such a specific situations the correct application of an exemplar's circumstances will *necessarily* transform an anomalar to a better state of affairs.

The next developments in causation focus onto the concepts of necessity, sufficiency and sequence which were developed from a renaissance of scientific insight during the 16^{th} - 20^{th} centuries.

3.4 Renaissance Causation.

Up until the era of profound scientific discovery and explanation lead by luminaries such as Galileo Galilei (1564 –1642), Newton (1642-1727) and Leibniz (1646-1716) and philosophers such as Hobbes (1588-1679), Descartes (1596-1650), Spinoza (1632-1677), Locke (1632-1704), Leibniz (1646-1716), Hume (1711-1776), Kant (1724-1804) and Mill (1806-1873), the theory of causation was adequately articulated by Aristotle and Aquinas. Subsequent to Aquinas, progress was gradual and involved a complex interweaving of theological and scientific views. The ground won by Aquinas for the primacy of God was not to be yielded easily as both Galileo and Newton were aware that their explanations of natural phenomena using mathematics were incomplete.

3.4.1 Galileo's Definition of Causation.

Galileo offered one of the clearest definitions of cause so far: the concept of the necessary and sufficient condition relating cause and effect – 'that and no other is to be called cause, at the presence of which the effect always follows, and at whose removal,

the effect disappears'. In his definition, it is left unsaid as to what constitutes 'that'. The essence of cause is unclear, but its existence is confirmed by observance of effect. Observance was an Aristotelian fundamental, but Galileo believed that it was necessary to describe an observance, compare it with other observances and to use the language of mathematics as the mechanism for doing so (Pearl, 2000; Hodgson, 2003).

The final end of this 'scientific approach' was the codification of cause via mathematical descriptions that not only accounted for the observed effect, but predicted others under varying circumstances. In other words, Aristotelian logic – driven by human perception, was necessary, but not sufficient. According to Galileo the combination of logical perception and mathematics provided both necessary and sufficient conditions for describing nature.

3.4.2 Thomas Hobbes and Causal Necessity.

Hobbes espoused a deterministic view, firmly cemented in theology that, 'God's decree was the foundation of all cause' (Hobbes, 1688/2000). The reciprocity between God's knowledge and decrees sufficed to confer rigid necessity on nature. The effect of this was a renunciation of scientific influence on the epistemology of causation. The root of Hobbes' determinism was theologically based rather than a carefully weighted argument against empiricism (Hulswit p 19). Hobbes was captured by the omnipotence and omniscience of God and saw no escape from the conjecture that experience, whatever it was, had to be consistent with theology in the long run. Investigation, consistent with God's plan, was of course still possible because even though the 'whole cause' was the work of God; parts of the cause were clearly not. In his 'Treatise on Liberty' (TL) Hobbes notes – 'the influence of the stars is but a small part of the whole cause, consisting of the concourse of all agents', and further, 'nor does the concourse of all causes make one simple chain or concatenation, but an innumerable number of chains joined together, not in all parts, but in the first link God Almighty; and consequently the whole cause of an event does not always depend on one single chain, but on many together' (Chappell, 1999). His example of the 'last feather that breaks the horse's back' (Chappell, p 34, TL §23) also re-affirms his distinction between whole, partial, and even the last component of cause.

Causal necessity and sufficiency were also addressed with some clarity: 'Seventhly, I hold that to be a sufficient cause to which nothing is wanting that is needful to the

producing of the effect. The same also is a necessary cause' (Chappell, p 72, TL §31). This extends to the conjecture that all events have necessary causes and that consequently that cause is sufficient (Chappell, p 74 TL §34). For a singular causal relatum, Hobbes' argument holds well, but for complex relata (as Mill will later describe) a distinction exists between sufficiency and necessity.

It is tempting to interpret Hobbes' causation as conjecture that some '*final* cause' and the first causal link were the handiwork of God, but thereafter other agencies were involved with partial causality which was limited to *efficient* cause (movement or events). Further, his uses of necessary and sufficient conditions clearly identify with singular causal relatum and his references to chains of singular causes admit complexity but through a series of causal conjunctions as shown in the following example using Hobbes' horse.

If A is the feather and B the Horse (plus any prior feathers on its back) then C is the effect of their combined weight. This sequence is repeated until the horse's back fails. A and B are conjoined to cause C and thereafter, C is conjoined with D, another feather, to effect E, greater weight, repetition of which eventually increases weight to X which is a sufficient condition for the last cause of the effect F – to break the horse's back.

$$\text{E.g. } A \wedge B \Leftarrow C,\ C \wedge D \Leftarrow E,\ \ldots.\ , \Leftarrow X \Leftarrow F.$$ (see Glossary)

Hobbes singular causal chains of conjunctions provide a basis for modelling both simple and complex systems and processes with the application of propositional logic. If a singular conjunction is combined with negation a complete logical set including disjunction is available from De Morgan's laws of propositional logic.

$$\text{E.g. } \neg(A \wedge B) = (\neg A) \vee (\neg B),$$
$$(\neg A) \wedge (\neg B) = \neg(A \vee B).$$

Modern digital technology is based principally on logical (Boolean) causation where functions are performed by arrays of Boolean processors and storage elements.

3.5 The New Metaphysicists.

The Metaphysics of René Descartes, Benedict (Baruch) Spinoza and Gottfried Leibnitz founded a post-Aristotelian school of philosophy to which historians credit Descartes as its father (Woolhouse, 1993). *Efficient* causation, manifest in moving bodies and

observable through the effects of their collision provided this school with a mechanical model for their philosophy (Woolhouse, p 134).

3.5.1 Cartesian Explanation.

René Descartes firmly believed in a mechanistic view of nature and explanation. He held that kinematical transfers or motion was a force within matter that passed from one part of it to another to effect change. His ontology admitted two aspects of substance: substance itself and its associated mode. If two substances interacted, their modes might alter – evidenced by changes in speed, etc, but the shape, say, of the substance was unable to change and if necessary transferred to the other substance in its entirety (Gaukroger, 2002b). Leibniz later criticised this view on the ground that it appeared inconsistent with the conservation of energy.

Descartes also broke with the four-part Aristotelian causal model holding that *efficient* causation was the only basis for explanation (Woolhouse, p 194). This parting of the ways was due to the prevailing Thomistic influence that accorded to God the positions of *final* cause and creator of material and form (Hulswit, p 18). Even so, Descartes distinguished between God's universal cause and particular (secondary) causes that were the subset of causes God allowed to creatures and accounted for the diverse movements evident in individual bodies (Gaukroger, 2002a, p 132).

Descartes also commented on semiotics, the study of signs and symbols. As his causation was mechanical and evident through motion, other phenomena, such as human behaviour, were ineligible as causal relata. In his Principia, (Part IV, Living Things), Descartes instances 'tears as a natural sign of sadness and laughter as a natural sign of joy' yet admits no causal link between each of these emotions (effects) and their physiological antecedent. Causes are held to be otherwise since we are not required to recognise them in order that they might occur whereas signs are the recognition and interpretation of innate capacities endowed by God (Gaukroger, 2002b).

The weaknesses of the mechanical model were quickly exposed by Leibniz and others and semiotic causal relata were addressed by Charles Peirce two hundred years later in the formation of semiotic causation theory.

3.5.2 Spinoza's Logical Necessity.

Benedict Spinoza's contribution to causation was contained in his work on Ethics. He was said to display a somewhat anthropomorphic epistemology perhaps evidenced by his forthright dismissal of the concept of *final* cause as 'nothing but human fictions' (Woolhouse, p 193, Hulswit, p 24), as nature has no set end or purpose to it. As with Descartes distinction between overall cause and particular causes, Spinoza also segmented cause into two categories: free and necessary. These, he explained differed in that free causes were established by the 'necessity of their own nature', whereas necessary causes are serially related to free causes in a chain-like manner (Hulswit, p 23).

Spinoza's approach to causality was empirical and singular - 'knowledge of the effect is nothing but acquiring a more perfect knowledge of its cause' (Nadler, 2002). Delahunty's (1985/1999) example illustrates both the singular and empirical: 'if you hear barking, expect a dog'.

His dependence on logical necessity was resolute: nature was driven by necessity and God was the embodiment of nature - the immanent cause (Delahunty, p 131). Inspired by Euclid's geometry (Spinoza was a lens-maker), axioms prefaced the important parts of his work. Some were matters of fact whilst others, self-evident to Spinoza, were simply stated as such. Unlike Descartes who prevailed upon God for veracity of thought and knowledge (Nadler, p 48), Spinoza simply called it as he saw it. Spinoza attributed two particular axioms to causation:

> 'From a given determinate cause the effect follows necessarily; and conversely, if there is no determinate cause, it is impossible for an effect to follow',
>
> 'The knowledge of an effect depends on, and involves, the knowledge of its cause' (Nadler, p 59).

The first causal axiom establishes causal necessity and the second, causal rationality. For Spinoza, causal necessity was what Mill would later describe in terms of necessity and sufficiency – an invariant logical relationship where the determinants were sufficient to generate the effect. No spontaneity or chance components of causation

were possible since there was a universal relationship between causal determinates and the effect.

Spinoza's second causal axiom was the statement of a fundamental epistemological tenet. He was saying that to know the truth about something, to understand all there is about it, was sufficient to understand its causality.

The first axiom and the second – taken to the extreme - have counterfactual overtones. Spinoza claims absence of a particular effect disassociates its determinants from being a cause of anything else. However this leap of logic must also be linked to Spinoza's substance ontology which was also singular in nature.

Another of his axioms held that nature has nothing but 'substance and their affections'. By substance Spinoza meant what Aristotle would call its *material* cause: what something is made of, its attributes and its very essence. Substance need not be immediately tangible as it included thoughts about something: tangible but not present. Affections were similar to Aristotle's *formal* cause – modes, or perceptions that we have about substance. Spinoza held that through its attributes, substances were unique: no two substances could have the same attributes (otherwise they were identical) and thus had nothing in common (Nadler, p 64).

Spinoza's second definition of causation recommends itself to the process of benchmarking where it is common for effects to be more obvious than their cause. Knowledge of the 'truth' about an exemplar's processes and systems, whether technical or otherwise, is essential to understanding the cause of its success. Although modern science refutes Spinoza's notion of a total lack of commonality between substances he would doubtless have continued to argue that organisations displaying different attributes were indeed unique and their degree of commonality questionable. The process of benchmarking involves the anomalar striving to attain the attributes of an exemplar, but Spinoza's metaphysics precludes imitation and suggests that a takeover is the way to guarantee success.

3.5.3 Leibniz and Mechanistic Causation.

Gottfried Leibniz held that natural processes must be explained by the aid of unalterable laws inherent in the existing nature of things – i.e. mechanically, and that these laws

must constitute the essence and very nature of the things themselves (Merz, 1884/2001). Through reasoning alone knowledge of the existing nature of things would reveal their causes (Hulswit, p 27). He also believed that scientific and philosophical treatments could be different, particularly as the scientific approach strove towards the observable and measurable. Encumbering scientific progress with metaphysical notions (or for that matter legal or political interference) was an entirely unnecessary activity (Merz, p 45). Even so, Leibniz could not dismiss out of hand the Thomistic fundamental that God was the primary cause, but he reasoned that if this were so, there were secondary forms that caused motion both internally and externally to substances (Lodge, 2004).

This 'reasonable' approach to the determination of cause was a modification of the antecedent epistemology holding cause and effect to be binaries. Insistence that 'a reason' be associated with effect introduced the concept that cause and effect were bound by natural laws, rather than Spinoza's necessity or Aquinas's finality, differentiated Leibniz and added strength to the scientific process of explanation.

3.5.4 Locke's Empiricism.

John Locke in his Essay on Human Understanding (EHU) demonstrated familiarity with the works of Galileo evinced by his questioning of the experiment that claimed equal swing periods for a given pendulum length. He observed that 'if anyone should be asked how he certainly knows that the two successive swings of a pendulum are equal, it would be very hard to satisfy him that they are infallibly so; since we cannot be sure that the cause of that motion, which is unknown to us, shall always operate equally' (Locke, 1690/2001). This example is one of many where Locke's appreciation of recent empirical developments intersected with an appreciation that contemporary knowledge had limitations – he knew what was unknown! His reference to 'cause' also indicated that the concept of cause and effect also contained another component – 'reason' In this example substitution of 'reason for' for 'cause of' would have more suitably illustrated Locke's work to isolate the metaphysical from the empirical via the mechanism of generating 'ideas' from observation. Cause and effect were simply ideas.

Eschewing brevity, Locke noted in the Introduction to EHU that understanding our human capacity is to be likened to 'a sailor who makes good use of a sounding line which, although short and cannot fathom all oceans, is long enough to avoid perilous shoals' (Locke, EHU, Vol. 1, Introduction, p 23).

This hierarchical approach to explanation underpins the importance of striving to understand but to only lengthen the sounding line when matters are proven. Although William of Occam's 'razor' is not mentioned, Locke's writings suggest that reflexive adoption of metaphysical necessity in the absence of natural explanation is to be resisted.

On causality, he notes that change is associated with 'power' since 'the sun has the power to melt wax', (Locke, II, XXI, 19, p 191) and that human 'will' was also 'power'. Locke observed these to be demonstrations of causality as a consequence of power in action. Power held two forms; active and passive – the former able to make change and the latter able to receive it (Locke, II, XXI, 2, p 184). Specifically: 'cause is that which produces any simple or complex idea and effect is that which is produced' (Locke, p 258). The melting wax analogy combined these concepts and denoted fluidity as the effect of the cause, heat, where heat was power.

But what of Locke's view of causal necessity and the concept of chance? There appears to be no direct reflection on causal necessity – apart from equating it to power. When addressing the issue of human knowledge, Locke did so from the perspective of 'ideas born out of observation', and held that knowledge extended no further than the perception of agreement or disagreement (Locke IV, III, 2, p 455). The implication of 'this is that' if we have 'no idea', then we cannot have knowledge of the relationships between cause and effect and the concept of necessity is latent. Chance was also as Aristotle perceived it - a haphazard coincidence of circumstances which if unfathomable by reason, was at least the outcome of God's plan (Locke, IV, XVII 2, 24 p 575; XX, 3, p 592).

Locke's purpose was to establish an epistemology of inquiry into human knowledge and established boundaries surrounding its extent, dissent, fact and opinion. His contribution to causality was the concept of 'power at work' without significant challenge to prevailing Aristotelian beliefs. However, his use of 'reason' as a causal relatum is significant, as it hints at an alignment of *efficient* and *final* causation.

3.5.5 Newton's Rationality.

Isaac Newton's Principia Mathematica, written almost 20 years after his insight into the behaviour of light and moving bodies provided the tools for a rational and mechanistic

approach to causation. This work established laws of physics in mathematical terms and thus established many causal relationships anticipated by Galileo. Newton also accelerated the reduction of causation to a single, *efficient* form – leaving superfluous or at least questionable, the notion of purpose or *final* cause. The codification of an already suspected mechanistic world prompted a zero-based, rationalist approach to causation. Newton's compulsion - or law-like behaviour was to the rationalist all that was necessary since (in a closed system) 'knowledge of the initial state is sufficient for the prediction of its state at another later time' (Bunge, 1959). This view is also held by Collingwood (1938) who claimed that Newton eschewed a 'law of universal causation' as for example in his first law of motion where movement is uncaused.

But according to Gabbey (2002), Newton understood that the new physics promised more than it could deliver as its explanatory simplicities still did not unlock the understanding of all natural phenomena. In fact Newton observed that it was unnecessary to question beyond the existence of something, such as gravity, given that its laws described its behaviour (Gabbey, p 58). In that sense, Newton and Aquinas were at one with each other.

Newton also addressed Locke's observation on the pendulum – that whilst behaviour is observable, cause might not be known. Hulswit (p 31) takes these observations to mean that Newton distinguished between causation and law-like behaviour. However Gabbey (p 332) suggests that Newton's position on metaphysics was more ambiguous and cites Newton's Opticks (Query 28) otherwise ':whereas the main Business of natural Philosophy is to argue from phaenomena without feigning Hypotheses and deduce Causes from effects, til we come to the very first Cause, which certainly is not mechanical'. Echoes of Locke are evident here.

The legacy of Newton was the elevation of scientific enquiry to a height unimagined by any before him through fundamental scientific explanation and the successful attribution of law-like behaviour to common phenomena.

In organisational benchmarking, exemplar behaviour may arise from better parameterisation of law-like behaviour. For example, the laws governing the transmission of electricity along conductors are now well known, but influenced by characteristics and parameters wholly within the power of alteration by engineers. The

advantage of an exemplar enjoying low transmission loss might be equalled by an anomalar who learns to align transmission parameters in the light of knowledge of the exemplar's achievements.

3.5.6 Humean Causation.

David Hume's treatise on Human Nature (1739/1999) wasted no time in addressing the importance of causation. He saw the potential for the human imagination's free reign towards unaccountability checked by some universal principles which he described as being relationships that exist between things in all times and places. Causation, contiguity and resemblance were the seemingly self-evident qualities that related these ideas or relationships (Hume, p 15) where causation was the most extensive. Hume was an empiricist who saw a history of impressions forming the basis of knowledge and understanding where these ideas encompassed relations, modes and substances. His view on causation gave rise to Kant's subsequent revelations on the distinctions between the empiricist and the rationalist. Hume saw the nature of cause and effect as being entirely experiential, and challenged alternatives:

'The existence, therefore, of any being can only be proved by arguments from its cause or its effect; and these arguments are founded entirely on experience. If we reason *a priori*, anything may appear able to produce anything. The falling of a pebble may, for aught we know, extinguish the sun; or the wish of a man control the planets in their orbits. It is only experience, which teaches us the nature and bounds of cause and effect, and enables us to infer the existence of one object from that of another' (Hume, 1748/2000).

Causation is an idea that involves tracing any idea to its point of origin and examining the primary impression from which it arises (Hume, 1739, p 57). This leads to the concept that causation has two properties: contiguity – connections via chains of causes and priority of time – cause precedes effect (Hume, 1739, p 58, 59). But Hume's Enquiry Concerning Human Understanding addresses his sceptics and advises them to bow to more profound research that shows causation to be evident, subject to the fallibility of human nature,

'.. from the relation of cause and effect; that we have no other idea of this relation than that of two objects, which have been frequently conjoined together; that we have no

argument to convince us, that objects, which have, in our experience, been frequently conjoined, will likewise, in other instances, be conjoined in the same manner; and that nothing leads us to this inference but custom or a certain instinct of our nature; which it is indeed difficult to resist, but which, like other instincts, may be fallacious and deceitful.' (Hume, 1748/2000, p 109, 110)

In other words, cause is the experiential inference of one object followed by another. Hume goes further: experience, not argument, infers that objects similar to the first to be followed by objects similar to the second at some other time. This is a clearer basis for counterfactual argument than was suggested by Spinoza: cause makes a difference, and the difference it makes is different from what would have happened without it (D. Lewis, 1973).

Hume's position on necessity as a causal relatum was more obscure. He identified and refuted some conventional arguments that advanced necessary relationships between cause and effect. Hume was dissatisfied with these conventional arguments. 'It is absurd to hold that nothing can be the cause of something; Everything must have a cause since, if a thing wanted a cause it would produce it itself and thence by priority of time, exist prior to its existence'. He noted that whilst these arguments were false, he could not advance any himself and wrote that it was 'neither intuitively nor demonstratively certain' that every object that begins to exist does so owing to a cause (Hume, 1739, p 62).

Hume's concept of contiguity and time priority built on earlier ideas and his causal succession (chains of causes) remain relevant in modern applications. Contiguity is a generic description of commonplace organisational process models; the process of benchmarking is chain of causal processes matching exemplar and anomalar effects with the objective of creating improvements.

3.5.7 Immanuel Kant.

Whereas Hume expressed doubts about the extent to which causality could be grasped by reason, Kant had no such difficulty and extended the epistemology to include substances and their causal interaction (Kant, 1783/2004). Kant also addressed issues such as reciprocity, simultaneity and necessity in association with causality. Acknowledging Hume as the spark, Kant explained that the uncertainty of existence

being ‘owed’ to a ‘cause’ was not the only concept through which the understanding thinks connections of things *a priori*; rather, metaphysics consists wholly of such concepts (Kant, 1783/2004, p 10; Wallace, 2000). Hume, the empiricist, believed that reason compelled *a posteriori* knowledge, but Kant, the rationalist held that ‘an inner truth independent of all experience’ accounted for what we observe in nature. He went on to further qualify the empiricist approach in rationalist terms:

‘Empirical judgments, insofar as they have objective validity, are judgments of experience; those, however, that are only subjectively valid I call mere judgments of perception. The latter do not require a pure concept of the understanding, but only the logical connection of perceptions in a thinking subject. But the former always demand, in addition to the representations of sensory intuition, special concepts originally generated in the understanding, which are precisely what make the judgment of experience objectively valid’(Kant, p 50).

Such advice does not go astray in organisations where ‘management by fact’ is one of the Malcolm Baldrige National Quality Award’s core values and is a fundamental benchmarking concept (NIST, 2007).

Kant also described his realisation of *a priori* knowledge in words that equated to the modern ‘paradigm shift’. He explained that metaphysics itself was of no help to him in arriving at this conclusion (Kant, p 50). His realisation that the concept of *a priori* was a consequence of nature’s necessary compliance with universal laws resulted in this articulation of causality:

'For this concept [of causality] makes strict demand that something, A, should be such that something else, B, follows from it necessarily and in accordance with an absolutely universal rule' (Adorno, 2001).

But this also extended to include precedence as well. The sequence of cause preceding effect in time was not merely a perceptive phenomenon but a principle (Kant, p 185). Furthermore, there was a lapse of time taken to complete an effect, whether long or too short to be perceived.

Reciprocity was implicit: if B was the effect of A, then the existence of B implied that A had caused it, hence A and B are reciprocally (necessarily) related. Hulswit (2000, p 38) cites Kant's use of examples as evidence for holding substance and events to be causal relata.

Kant's *a priori* metaphysics introduced a rationalist approach to causality and established the functional concepts of necessity and sequence. The empiricist approach was useful and relevant in that it was often based on an unrecognised universal *a priori* principle that eluded observers although their perceptions remained real. This nuance is of use to benchmarking as it identifies circumstances where an exemplar's perceived performance is imitated without success. Kant's message is that there is a reciprocal (necessary) relationship between cause and effect if there is an exact imitation of situations; otherwise another universal rule is operating rather than the one expected!

3.5.8 Mill's Unconditional Causal Antecedence.

John Stuart Mill was one of the most influential thinkers of the 19th century and his power was at its peak during the period 1860-1870. He is regarded as an empiricist and naturalist; 'the human being is the natural entity and beliefs, purposes and sentiments are its genuine properties and nothing beyond them is required' (Skorupski, 1998) - and his contribution to causality introduces the modern concepts of logical analysis.

In his System of Logic he re-specified Hume's view that cause was 'the universal law of successive phenomena is the law that every consequent has an invariable antecedent' (Mill, 1872/1973). Its importance is highlighted by the next sentence: 'Cause is the root of the whole theory of induction and its idea should be fixed and determined with the utmost practical precision'. In this definition, Mill establishes the Humean concept of chains of antecedence and adds the Kantian concept of necessity. However, not only were there chains of causation, but generally multiple antecedents as well. Singling out only one of the antecedents as the basis for causal determination was generally impossible as the assembly of antecedents generally provided the conditions, all of which were necessary, to produce the effect. He gives the example of a man dying after eating a particular dish and notes that the dish need not be the cause of death, as there may be no invariable relationship between the eating of the dish and death, yet there is certainly some combination of antecedents that constitutes this cause.

The ‘real cause’ is the whole of the antecedents (Mill, p 328). He redefines this by stating that the ‘cause is not the invariable antecedent, but the unconditional invariable antecedent’ (Mill, p 338). This may be expressed logically:

$$A \Leftarrow B, \text{ iff } A \rightarrow B \text{ and } \{A\} \rightarrow \{B\}$$

A is the cause of B and B is the effect of A, if and only if A is immediately followed by B and things similar to A are always followed by things similar to B.

To exemplify this, Mill noted that we do not believe that night is always followed by day since it is so only provided that the sun rises above the horizon. However, we do not say night is the cause of day, or even a condition of it since either perpetual night or day are possibilities admitted by the laws of nature. It is the unconditional nature of the relationship that Mill emphasises, which is generally to be found in sets of conditions, each of which are necessary, but none of which are sufficient to cause the effect.

Interestingly, Mill also noted that ‘it was a law that every event depends on some law: it is a law that there is a law for everything’ (Mill, p 325) and consequently that every event has a cause, that if there is a cause, then there are necessary conditions for that event and that there is an unconditional sufficient condition for that event (the conjunction of those necessary conditions is the sufficient condition).

But there might well be individual sufficient conditions or insufficient parts of the sufficient conditions. For example, if a fire started, the necessary conditions for that effect include oxygen, heat and flammable material. Heat is a necessary condition, but a sufficient condition to generate heat might be a spark. Hence a spark is insufficient for fire, but part of a necessary condition, heat, which is now rendered unnecessary and in the presence of oxygen and flammable material a spark becomes sufficient. This is known as INUS (Insufficient, Necessary parts of Unnecessary Sufficient conditions) later described by Mackie (1975, p 62).

Suppose Heat (H), Flammable Material (M) and Oxygen (O) are causal antecedents of Fire (F). If there is a Fire, it is the effect of a necessary conjunction of antecedents that forms the sufficient condition:

$$\{H\} \wedge \{M\} \wedge O \Leftarrow F;$$

If {C} is defined as a complex condition,

$$\{C\} = \{H\} \wedge \{M\} \wedge O, \text{ then } \{C\} \Leftarrow F.$$

But if Heat is also set of independent producing conditions, Spark, (S) or electricity (E) then:

$$S \vee E \rightarrow \{H\}.$$

Then S (or E) is INUS of F is written as

$$S \Leftarrow \{H\};\ \{H\} \neg\Leftarrow \{C\} \Leftarrow F, \text{ hence } S \neg\Leftarrow F.$$

This says that for complex conditions, there might be more than one way of achieving the effect.

The concept of reciprocity is also invalid if causation is viewed strictly mathematically. Using the above example, if, $\{C\} \Leftarrow F$ we are not entitled to say $F \Leftarrow \{C\}$ as this is clearly ridiculous. The presence of Fire might imply that a sufficient condition occurred, but it cannot be disaggregated further; we cannot say that the Fire caused either the Spark or the Electricity.

INUS is perhaps more easily recognised in Boolean expressions. Taking the above antecedents as Boolean (a mathematical system of logic that constrains arguments to be either true or false), we have:

$$F = H \bullet M \bullet O$$

$$\text{But if } H = (E+S) \text{ then } F = (E+S) \bullet M \bullet O.$$

Thus S, alone is insufficient to cause a fire (M and O must also be true), but it is part of a necessary condition (H), S is also an unnecessary part of a sufficient condition (if M and O were true), since E could equally suffice.

For Mill, the teleology of science is the establishment of laws giving the necessary and sufficient conditions for all events (Skorupski, p 246). But if this were so, Mill sees this teleology not as *final* cause but science's role in the refinement of *efficient* cause. Yet by identifying cause as an unconditional and invariant conjunction of antecedents Mill's seems to have abandoned the empiricist epistemology to join the rationalists. Hulswit (p

220) likens Mill to Laplace who held that knowledge of initial conditions was sufficient to determine nature through mechanistic principles.

Mill's commentary on necessary and sufficient conditions, the existence of complex conditions, and the concept of INUS is relevant to benchmarking. Clarity of effect does not imply clarity of cause, especially when there are numerous necessary conditions. Causal reciprocity is clearly wrong and signals that observations of exemplar behaviour (effects) may simply confirm little more than its existence and leave the anomalar in no wiser state. Even knowledge of necessary conditions may not help the anomalar, unless they are determined to be sufficient, and for reliance, also invariant.

However, there is much unsaid in the work of Mill. There is no consideration of cause other than the Aristotelian concept of *efficient* cause. Mill's substance ontology is bound in his observation that there is 'a law for everything' – substances and events included, but he goes no further. Chance and spontaneity are rationalised likewise. We will see that chance, or probability, plays a role in law-like behaviour and limits the certainty of necessary and sufficient conditions in causal relationships. These issues are of interest and importance to benchmarking as causal relationships in organisational environments appear broader than those explained by Mill.

3.6 Summary of the Ancient to the Modern.

The trajectory of causation theory from the ancients to the modern era has oscillated between the epistemologies of the natural and divine. Aristotle's natural empiricism augmented with Stoical theses of fate and determinism was strengthened by Aquinas whose solution to the problem of human reasoning explaining causal extremities lay elsewhere: they were vested in the divine Creator.

Even so, by the start of the seventeenth century the foundations of causal theory were well established and the 'solutions' gained by the assumptions of Aquinas were steadily lost to the 'solutions' of the rationalists and natural empiricists that followed him. Throughout, Aristotle's taxonomy of cause was steadily whittled to embrace only *efficient* cause with natural science offering an appealing encompassment of *formal* and *material* causes.

Final cause was deemed to be at best situational – serving only immediate ends or an intrinsic part of *efficient* cause that appeared to explain either purposeful or spontaneous effects. Better knowledge of the structure of matter confirmed Spinoza's second axiom of cause, Kant's *a priori* approach avoided fundamental reliance on the Almighty to achieve explanation and Hume, a devout empiricist, rationalised causal chains and the time priority of cause and effect.

The concept of necessity as a causal relatum was accidentally unlocked by Mill by his acknowledgement that causal sufficiency consisted of complex combinations of necessary conditions. This logic, empirical in nature, also implied the existence of insufficient necessary parts of unnecessary sufficient conditions (INUS), demonstrating that universality was more of a concept than a fact once the actual details of necessary conditions were known. Counterfactual analysis relied upon universal regularity of the relationship between cause and effect, but this may fail to fully satisfy causal relationships as there might be INUS antecedents.

Issues such as chance and likelihood were too difficult and were excluded from causal theory either on the basis of outright impossibility or acknowledgement that insufficient knowledge existed to fathom the relationship between cause and effect.

Yet so far as forming a basis for the construction of a theory of benchmarking, the period spanned by Aristotle to Mill delivers a rich source of causation theory that is worthy of attention. If causal extremities are provisionally set aside, we avoid ontological issues surrounding first cause and admit all theories for consideration in a modern organisational environment.

In order of immediacy, Aristotle's *efficient* and *final* causes are relevant. Spinoza would have no difficulty in observing that action and purpose are identifiable and self-evident organisational characteristics. For the purpose of benchmarking, *formal* and *material* cause signals that there are issues of substance and governance that need to be considered. Aristotle's *formal* cause indicates the want of knowledge on the laws or principles governing an organisational state of affairs and it may assist if the historical practice of subsuming this into *efficient* cause is resisted. Moreover it is not possible to avoid Aristotle's scientific principle of answering the question 'what comes to be after what, and so on at each step of the series'. Finally, Aristotle's substance ontology prompts a

distinction between being and essence: knowledge of organisational states may not imply knowledge of their essence, however compelling it may be to presume otherwise. This distinction has been recognised and Ghoshal (2005) summarises Andrews' (1980) observation 'that replacing human intentionality with a firm belief in causal determinism' is but a 'pretence of knowledge' and serves a warning that causation is not merely *efficient* cause.

After Aristotle, it might be argued, that little real progress on causation occurred until Galileo asserted that nature's handiwork also needed to be recorded in mathematical language. This challenged the empiricist and gave rise to rationalist explanation, yet the two approaches are not mutually exclusive.

Developments from Galileo to Mill, although they included Newton, Leibniz and Descartes, amounted to better descriptions of causal necessity and sufficiency. But these developments were not without significant reinforcement; Kant's rationalism provided 'management by fact', which, when coupled with Mill's complex causal antecedents, offers a model for singular *efficient c*ausation. However, Hume's counterfactual arguments fail and Mill's INUS warn of the dangers of causal necessity and sufficiency if there are complex causal relata about which there is incomplete knowledge. This is of great importance to benchmarking as it invites logical errors of cause from reliable observations of effects.

3.7 Contemporary developments.

Several unanswered questions still command attention. Linkage between causation and the probability of antecedent conjunction was not been addressed by Mill or his predecessors. Also, the focus on *efficient* cause, prompted by greater understanding of the 'laws' of nature, still does not provide a satisfactory basis for causal behaviour, especially organisational behaviour. Aristotle's 'for the sake of' question lingers and needs development to address causation in situations that are clearly purposeful. It may also be useful to consider whether Aristotle's substance ontology concerning the preservation of 'form' remains valid in organisational settings where benchmarking occurs. Form might not encompass all that an organisation can do, but it might determine what it may not do.

The next Sections examine causal explanation from the perspectives of counterfactuals, statistical explanation, singular causation and regularity, causal laws and relations,

analytical reductionism, causal relata and Peircean causation. These perspectives address considerations that arise in organisational benchmarking and complete the framework that provides a basis for developing a theory of benchmarking based on causation.

3.7.1 Counterfactual Dependency.

The basis of counterfactual dependency is the existence of necessary conditions that provide a non-stochastic relationship between cause and effect. 'Cause makes a difference and the difference it makes is different from what would have happened without it. Had the cause been absent, then its effects, some of them at least, and usually all - would be absent as well' (Lewis, 1973, p 557). Criticisms of counterfactual dependence as a basis for causal explanation are both broad and narrow.

Broad arguments include multiple simultaneous actions: two simultaneous actions, each of which is sufficient to cause change, leave the counterfactual argument unable to delineate between the two and is said to be a causal over-determination as either condition suffices.

Narrow examples include counterfactual dependence that is unnecessary for causation. Kim (1973) and Psillos (2002) elaborate by providing numerous examples of the following kind: 'if I had not spoken, I would not have been heard', 'if yesterday had not been Monday, today would not be Tuesday'. These are not universal causal relationships as, for example, Tuesday is not caused by Monday, they are simply logically sequential and speaking does not always result in one being heard, and one can also be heard without speaking. These are examples of causal under-determination, where care is needed to delineate between causal and non-causal counterfactuals.

In its contribution to a benchmarking theory, counterfactual dependency raises difficulties can only be addressed through careful delineation between the apparent and real causes of an effect. To conclude that the absence of an effect implies the absence of a cause based on a counterfactual argument also requires proof that a singular, invariant relationship exists so as to avoid the error of over or under determination of causal relata.

3.7.2 Statistical Explanation.

The existence of necessary and sufficient conditions that cause an effect is appealing but not always true. Quantum physics, for example, is incompatible with this view and leads

to the consideration of a probabilistic model of causation. From observation of weather events we know that certain patterns precede fine weather or storms and prompt the view that not all events cause the same effect each time. If an event A is associated with an effect B, but not always, there is a probabilistic relationship. We might say that the probability of effect B, given that A occurs exceeds the probability of effect B given that A does not occur; i.e.

$$p(B|A) > p(B|\neg A)$$

(Tooley, 1987; M. Hulswit, 2002, p 52; Psillos, 2002, p 243)

This leads to one of the basic notions of probabilistic causation: a cause should *increase* the probability of the effect. However Pearl (2000, p 254) cites the example of a vaccine, generally held to lower the probability of contracting a disease, actually being the cause of it.

The issue with probabilistic causation is that it might statistically associate cause with effect in the long run, but in a single instance provides no certainty as to outcome. To say that 'smoking' causes lung cancer is a broad statement that asserts that the probability of contracting this disease is higher if one smokes. But if a smoker contracts lung cancer, smoking did not enhance the probability of contracting the disease. As in Pearl's earlier vaccine example, it actually caused it! Psillos (2002, p 247) reflects on the difficulties associated with characterising probabilities objectively and suggests that rather than interpret p(B|A) as the frequency with which the effect B occurs, given A, it should be interpreted as the objective chance of A being B. This toggles the argument from the general to the particular as it makes more sense to speak about the chances of an event occurring rather than to speak of their long-run frequencies.

There are many organisational processes governed by statistical dependence. For example, tourism is replete and agriculture likewise with weather-dependent activities that cannot be sold in adverse conditions, but over a longer timeframe. Successful organisations balance product portfolios to hedge against the possibility that on any day some products might be unviable leaving others to be offered as substitutes. Unavoidable effects that are statistically explainable (e.g. 'about' 5% of Ferry services are cancelled in any calendar year caused by adverse sea states) are either opportunities or constraints that exemplar organisations learn to manage. An anomalar seeking to imitate such an exemplar would need to both recognise and understand the manner in

which these unavoidable effects influence its superior state of affairs, how they are addressed and finally, whether the same statistical explanations are even relevant to its desired state of affairs.

3.7.3 Singular causation and regularity.

Hume's definition of the meaning causation was crisply based on contiguity and time priority and Mill further strengthened this by asserting that it was the unconditional invariable antecedent. Ducasse (1926) refuted these approaches and claimed that even a *singular* occurrence was sufficient for a causal relationship thus constancy was unnecessary.

Ducasse addressed the latter by providing an example which is relevant to the study of benchmarking as it demonstrates that co-incidence or correlation is not causation. 'A person sequestered from childhood hears two clocks striking the hours. One clock is slightly faster than the other so that it chimes a fraction earlier'. According to Mill and Hume, the person would be entitled to claim that the earlier clock causes the sound of the other (Psillos, p 57). In this example, Ducasse separates correlation from cause and further, separates correlated events from common cause (the two clocks have nothing to do with each other). Constant conjunction is not necessary for causation. The counter-argument is that Hume was really referring to *genuine* causal relationships – though how one tells the difference is unclear. Mill's unconditionality refers to situations where there is a causal relation if correlations cannot be screened off (Psillos, 2002, p 62). In counterfactual terms, 'if x had not happened, then y would still have happened'.

Ducasse (1926, p 58) stressed that only events could be considered as causes – i.e. 'and by an event is to be understood either a change or an absence of change (whether qualitative or relational) of an object. On the other hand, objects themselves (in the sense of substances, e.g., gold; or things, e.g., a tree) never can properly be spoken of as causes or effects but only as agents or patients, as components or compounds, as parts or wholes'.

And Ducasse's definition of cause has shifted to include a broader scope of consideration and is quoted in full:

‘Considering two changes, C and K (which may be either of the same or of different objects), the change C is said to have been sufficient to, i.e., to have caused, the change K, if:

1. The change C occurred during a time and through a space terminating at the instant I at the surface S
2. The change K occurred during a time and through a space beginning at the instant I at the surface S.
3. No change other than C occurred during the time and through the space of C, and no change other than K during the time and through the space of K.

More roughly, but in briefer and more easily intuited terms, we may say that the cause of the particular change K was such particular change C as alone occurred in the immediate environment of K immediately before’. Ducasse (p 59)

This definition is important for two reasons. Firstly it introduces the notion of singular causation (as there is no assumption of a regular relationship between C and K, a single occurrence of C suffices. Ducasse also implies that there is no such thing as a ‘causal law’ – only that there are causal relations). Secondly it introduces the environment as a component of the causal relationship: the environment of an object, some change in the object and some change in the environment thus forming a triadic relationship (Ducasse, 1926, p 59).

There are criticisms of the singularist view – from Ducasse himself and from others. The distinction between accident and cause or sequence and cause may be difficult to resolve. Psillos (p 69) provides an example of accidental or co-incidental action being indistinguishable from intended action (or the cause). Ducasse (1969) admitted that a sequential action meeting the criteria of beginning at instant, I, and terminating at the surface, S, was indistinguishable from a purposeful one. These two criticisms are similar and illustrate that this theory of causation also has its problems – particularly in situations where it is not possible to reliably decompose change and be assured that the relationship is genuinely causal.

As a contribution towards a theory of benchmarking, Ducasse’s definition is appealing as it encompasses the environment associated with those changes (events, objects) the anomalar seeks to understand and imitate and further admits that they might also have

been singular in nature. Discerning the difference between genuine causation and accidental (or sequential) events remains problematic, but in theory, exact knowledge of the circumstances would resolve this. Such a pragmatic approach might address organisational benchmarking relationships, but would not satisfy philosophers. For them, the inclusion of the environment adds extraneous considerations that must be reconciled. For instance, if there was a third event Z, adjacent to C and K, how can it be reconciled into the causal equation? The answer, Ducasse claims, is arrived at by reduction. If Z was apparently insufficient to cause the event and what was apparently sufficient to cause the event would have remained so even when Z was removed, then Z was not really part of the environment.

This approach is of relevance to a tactical application of a benchmarking theory as it highlights the difficulties of establishing causal relationships in the presence of other proximal events in the environment. Yet Ducasse advances one element of Hume's and Mill's theories by identifying that cause might not be regular, but singular. This is of importance to organisational benchmarking. If an anomalar seeks to establish a successful improvement strategy based solely on observations of '*regular*' (unconditional, invariant antecedents) changes in an exemplar's events they may omit prior changes that were singular in nature but were necessary for the exemplary state of affairs under observation. Identification of all of the environmental variables and their relationships (whether purposeful or not) is clearly essential to the establishment of causal relationships that meet Hume's and Mill's and definitions of regularity as well as Ducasse's definitions that admit singularity.

3.7.4 Other Contemporary Issues on Causation.

Since Aristotle's original taxonomy of causation, scholarship has concentrated principally on *efficient* causation – the cause-effect transition. Benchmarking organisational processes is analogous to *efficient* causation, but there is more to benchmarking, and to causation, than the element of *efficient* causation. Scholars have examined the mechanisms surrounding the establishment of an effect and whilst all theories are perfectly suited to some circumstances, they all experience difficulties as complexity increases – either in the nature of the causal relata or the environment in which change occurs. Before examining a theory of causation that is broader than *efficient* causation, it is appropriate to examine contemporary views on some of these

unresolved problems as well as the nature of causal relata – the building blocks of causation.

Sosa and Tooley (1993) identified a number of themes of enquiry that expose residual issues about *efficient* causation. Two in particular will be examined in the context of benchmarking.

1. Are there causal laws or causal relationships? Is the singularist view correct, or is the Humean/Mill view correct? Does this distinction matter so far as either view's contribution to benchmarking is concerned?
2. Causal Analysability: Are causal concepts axiomatic or do they stand in need of analysis?

These are discussed in the next sections.

3.7.5 Causal Laws and Causal Relations.

It quickly becomes clear that a distinction must be made between causes and the features or explanations describing them (Davidson, 1993). For instance the Humean definition of cause: 'an object, followed by another, and where objects similar to the first are followed by objects similar to the second' calls for explicit understanding of the relationship between objects – airy explanations prove inadequate. For instance, in an organisational benchmark setting the statement:

'benchmarking caused us to improve',

is clearly problematic since not all instances of benchmarking always result in improvement. However, if the statement broken down into a set of facts it might be more accurate – for example:

> 'The fact that we engaged in a benchmarking exercise demonstrated that one of our processes was inefficient and remedying this deficiency caused our improvement'.

This is a Humean statement where object of the inefficient kind followed by objects of the remedial kind are causal. Davidson (1993, p 82) further observes that the fuller we make the description of the cause, the better our chances of demonstrating that it was sufficient to produce the effect (and the worse our chances of demonstrating that it was necessary), whereas the fuller we make the description of the effect, the better our

chances of demonstrating that the described cause was necessary (and the worse our chances of demonstrating that it was sufficient). This observation directs analysis to deeper levels to establish causal necessity and sufficiency. However even sufficiency might not result in causality if there were in fact other contributory objects. The statement: 'remedying a deficiency is followed by an improvement' demonstrates sufficiency, but if an improvement is observed, it is possible that many deficiencies were remedied not just that a single deficiency was remedied thus the cause of this effect was the result of both necessary *and* sufficient conditions!

In summary, the distinction between causal laws and causal relations is found in the manner of explanation: casual laws that are analogously Newtonian or compulsive (e.g. gravitational attraction) operate fundamentally between objects, but causal relationships (causal explanations) describe the events in which objects are observed to participate. The fuller the description of the transition between objects in question, *the more likely it is* that light will be shed on whether or not there is a genuine causal relationship, what causal laws are in effect and whether the cause is regular or singular.

3.7.6 Casual Analysability.

According to Tooley (1993), there are two basic approaches to causation: reductionism or realism. The reductionist holds that causal laws are supervenient upon the totality of history whereas the realist holds that they are not. The arguments that surface between these two views illuminates some of the difficulties associated with benchmarking practices and the conclusions anomalars might reach from a benchmarking exercise.

Reductionists address the question as to whether facts about causal relations between events are reducible, as a matter of logical necessity, to facts about other states of affairs (Tooley, 1993, p 173). This is a question of supervenience in which causal relations are logically supervenient on non-causal relations.

Concepts of over and under-determination of causal relata were raised in the treatment of Counterfactual Dependency. The same concept applies to reductionist arguments that suffer the inability to discern between laws and accidental regularities. If it is genuinely accidental that object A is associated with object B, how could this be known, and how would observers avoid the assumption of a law that associates them, even if the 'accident' occurs more than once?

This notion of accidents occurring more than once is exactly what might occur under a probabilistic law. If a fair coin is tossed an infinite number of times, the probability of it falling 'heads' is equal to that of it falling 'tails', i.e. 50%. However, if a fair coin is tossed 100 times, it does not follow that 50% of its falls will be 'heads' as it could in fact be any proportion whatsoever. Thus one could conclude that there is a law that does not entail any restrictions on the proportion of events with the property, "coin tossing", with the property of "equal probability of outcome".

Whereas non-probabilistic laws impose absolute constraints on the history of the world, probabilistic laws impose no constraints whatsoever. Tooley (1993, p 178) contends that this defies reductionism. Moreover, he continues the attack with another argument that carries both epistemological and ontological connotations: can a finite body of evidence justify that some law obtains? What to make of a law that has yet to be instantiated?

If reductionism is to be abandoned, is realism its successor? Tooley (1993, p 190), a realist, explains that realism is based on two views: causal relationships are observable and causal concepts are theoretical in that they satisfy some appropriate theory. Of course this approach aligns with the principles of scientific enquiry where theory is paradigmatic (Kuhn, 1996) where paradigms rise or fall, depending on their ability to satisfactorily explain both instantiated and hitherto un-instantiated events. Thus a realist holds that causal concepts are theoretical relationships that require analytical rather than reductionist explanation.

Benchmarking aligns more naturally with the realist perspective in that the relationships are intrinsically observable (through the benchmarking process) and that the explanation of *effect* is consistent with normal science which form the prevailing paradigms.

However, there is a fine distinction between the approaches of the reductionist and the realist. Each seeks to analyse the causal relationship but the path diverges in that reductionism seeks explanation based on supervenience of existing knowledge whereas the realist seeks explanation based on theories that simply offer analytical consistency.

Organisational benchmarking explanations, via a reductionist approach, could become immensely complex since every element contributing to the superior state of affairs

would need to be isolated and reduced to elemental explanation in terms of existing laws. Realism offers a more pragmatic approach: whilst there is no avoidance of analysis, such analysis is within the umbrella of an appropriate theory.

These two approaches converge where an 'appropriate theory' is none other than a system of assembled facts addressing instantiated events. Once again, the distinction between accidental and genuine causal relationships re-appears and establishes itself as a factor essential to successful benchmarking.

3.7.7 Causal Relata.

A distinction must be made between the entities standing in a causal relationship and the relationship of causation. Hulswit (2002, p 67) observes that what we take causation to be depends on what we take causes and effects to be, and vice versa. Modern causation treats causation as a relationship between events – a change in the states of substances. Benchmarking is a process that involves observing and comparing different organisational states of affairs where substances differ to some degree. Benchmarking itself does not cause the advancement of an anomalar's particular state of affairs, rather it is the implementation of particular relationship between substances that does so. Benchmarking seeks to identify causal relationships. This raises the question as to whether the benchmarking decision maker (the agent) is the event that actually occurs in order to alter the states of substances.

Hulswit (2000, p 68) also identifies 'facts' as causal relata. Facts can be expressed in conjunctive relationships such as 'Our organisation improved *because* we remedied an inefficient process'. The two clauses of this sentence are discrete facts within a conjunctive relationship. Facts may also imply counterfactually dependent events – 'had we not remedied an inefficient process, our organisation would not have improved'.

Ducasse (1926, p 58) clearly identifies that only events can stand in causal relationships and that objects or substances cannot be considered causes or effects but rather are agents or components of the change process. Facts, he claims are generalisations of causal relationships: 'The causal relation is essentially a relation between concrete individual events; and it is only so far as these events exhibit likeness to others, and can therefore be grouped with them into kinds, that it is possible to pass from individual causal facts to causal laws' (Ducasse, 1926, p 61).

Davidson (1980) holds that events are unrepeatable and dated particulars (i.e. concrete substances rather than abstractions) that can be truly described. The inclusion of time in this explanation means that events are unrepeatable, although, clearly, they may be similar. An anomalar certainly seeks to enjoy the event once held by the exemplar in a concrete manner, though it will be the anomalar's event once it occurs.

Davidson's (1980) observation threatens the anomalar with the prospect of a truly unrepeatable event. Does the passage of time, or the particular time itself permit the future attainment of a similar state of affairs to that obtained by the exemplar? Clearly this may not be so. Suppose an event conjoined in time to confer a desirable state of affairs on an exemplar (perhaps the temporary relaxation of some legal constraint, or some short-lived global event). Unless this event's particulars may also be obtained at such later time as the anomalar addresses changes arising from benchmarking, there is little prospect of the two states of affairs being comparable. This particular argument has *formal* causation overtones as a shift of governing laws was seen by Aristotle as a *formal* rather than *efficient* cause.

Overall, the philosophical debate between events and fact theories of causation reduce to whether one has primacy over the other Hulswit (2002, p 71).

3.7.8 Summary of Contemporary Developments.

One inescapable observation from this examination of contemporary thought on causation is whether or not there is a 'normal' theory of causation. There are many theories responding to issues that remain problematic. This highlights that there is a plurality of theories and prompts the question as to whether there is no theory at all! Cartwright (2007) anguishes about this and reminds us that just as physics is rife with causal claims (e.g. gravitation forces are collinear with action, electromagnetic forces are orthogonal to action) perhaps we should take our lead from the domain approach of science where specific domains have their own theories.

Causation theories examined so far add considerably to an understanding of benchmarking. They establish numerous criteria that enable greater certainty to accrue to an exercise where the favourable states of affairs of one organisation can be identified and replicated, either in whole or in part, by another.

Counterfactual Dependency reminds us that causal relationships must be explicit and neither over nor under-determined. An inability to explicitly establish causal relata may lead to false or accidental claims of causation. Ducasse (1926) identifies singular causation as fundamental and claims that Humean regularity is but a more general case of a singular occurrence. This observation also provokes greater introspection during an organisational benchmarking process: is the observed state of affairs regular, where 'objects of one kind are always followed by objects of another kind' or is it a composite of many states of affairs – some regular and some singular?

Ducasse (1926) also identifies an important triadic relationship: the environment of an object, some change in the object and some change in the environment. This is an important observation in organisational benchmarking where dissimilar environments may not obtain similar outcomes when objects change in the same manner.

Probabilistic causation is considered from two perspectives. A probabilistic relationship may not be observable as such but as some altogether different relationship. Moreover are also mathematical methods for estimating the statistical confidence (likelihood rather than certainty) with which 'objects of one kind are always followed by objects of another kind'.

Davidson's (1993) discussion on causal laws (Chapter 3.7.5) provides benchmarking with sound advice: detailed knowledge of effects assists determination of antecedent necessities.

Facts and events are the building blocks of causation and whether one has priority over the other is not immediately important to the development of a theory of benchmarking. What is important is that they be ascertained.

Overall, there is an encouraging body of scholarly endeavour that adds insight to the proposition that benchmarking is a product of causation. However, the earlier observation that benchmarking is also a teleological phenomenon has yet to be addressed. There is no doubt that benchmarking is done for the sake of some greater organisational purpose and this will now be addressed via the work of C.S. Peirce.

3.8 Peircean Causation.

Charles Sanders Peirce (1839-1914), an American philosopher, logician and mathematician born in Cambridge, Massachusetts is often cited as the 'founder of pragmatism'. He wrote prolifically on philosophical topics, including causation. Publications containing Peirce's works include some papers and lectures delivered between 1867 and 1909, but the majority of his work was published from collections of his papers compiled long after his death. Much of his work is accessible through literature published by those having access to his manuscripts. There are also electronic resources available, such as the Digital Encyclopaedia of C S Peirce.

3.8.1 Teleological Causation.

Peirce developed an original view that acts of causation involved three components: *efficient*, *final* and *chance*. Efficient component refers, very broadly, to the *efficient* cause espoused by Aristotle and signifies that each event has its predecessor. Aristotle's *efficient* cause reflected the effects on things or substances by some agent – causing them to behave in a certain way. Peirce saw *efficient* cause relating to events or facts linked by previous events or facts – in similar vein to Hobbs' causal chains described earlier. Peirce's *final* component refers to the teleological purpose of the event – the sake for which the cause occurs and the *chance* component is that each event has some aspect that is determined neither by the *efficient* nor the *final* cause (Peirce, 1998, EP 2 #9).

Since Aristotle, it was generally held that *formal*, *material and final* causes were simply different manifestations of *efficient* cause – the mechanism describing transitions between events or the progression of facts. However Peirce held that *final* causes were teleological processes of a general nature that tended to be realised through determining processes of *efficient* causation. Hulswit (2000 III.1) interpreted this to mean that *final* causes were akin to habits – whether like human habits or habits of nature because they tended towards an end state (e.g. nature's inherent minimisation of energy).

This view is reminiscent of Aristotle's example of a person undertaking exercise, not merely to change location, but for the sake of their health (i.e. longevity, survival). In keeping with this theme, *final* causes might not be static, but evolve over time, to develop from, perhaps, intermediate states to another, and so on. In his description of

personality, Peirce recognised that this characteristic was more than a purposive pursuit of a predetermined end, it was developmental (*a developmental teleology*) and was influenced at later times by what was currently not conscious (Peirce, 1935, VI.156).

A teleological view of causation accords well with organisational behaviour since changes (whether purposeful or not) that do not contribute to the overarching *sake* of survival or, at more elemental levels, to improvements in current states of affairs, are at best inefficient or at worst terminal.

3.8.2 Attributes of Peircean Causation.

There were also other characteristics of Peircean *final* causes; they are not necessarily end states, but physical possibilities. Moreover, there might be many different ways and different timescales to progress towards these possibilities but the process directing this progression is irreversible. Irreversibility was the pragmatic outcome of the inexorable progression of time: 'boys grow into men, but not men into boys' (Peirce, 1935, VI.554).

Efficient causation is a secondary form of causation and stands beneath *final* causes, which are fundamental (Peirce, 1935, VI.101.f). Peirce thus states that causation is a triadic, hierarchical phenomenon that is depicted in Figure 3-1.

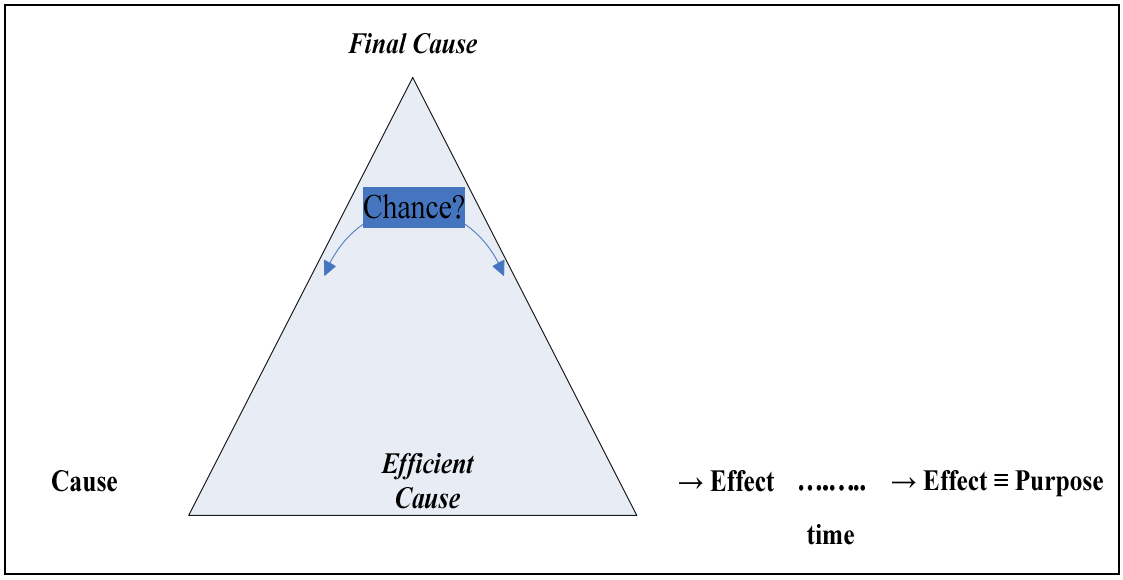

Figure 3-1. Peircean Causation.

Adapted from Hulswit (2000, 2002, p 81)

This diagram depicts *efficient* causation as a single event which gives rise to an effect

which may or may not be directed towards a particular purpose. The effect may arise as Peirce explains from blind compulsion or a mechanistic law, but none-the-less initiates change (Peirce, 1935 VI.217 & 454).

Efficient cause is dyadic and relates two events but also contains an element of chance. Peirce saw chance as a spontaneity or characteristic of the universe where it acts always and everywhere, though restrained within narrow bounds by law, 'producing infinitesimal departures from law continually and great departures with infinite infrequency' (Peirce, 1998 1 #308). This is not inconsistent with Darwin's theory of natural selection where rare and random, but helpful, changes in fauna and flora abet the survival (sake) of a species. Peirce clearly recognises that there may be unapparent complexities in the cause-effect transition which might sometimes deliver other than what an observer ordinarily expects. This pragmatic observation addressed some of the difficulties experienced by earlier causation theories – how to deal with Humean regularity and reduction of observations into deterministic components? More fundamentally, without chance, there is no teleology as pure *efficient* causation would be simple determinism.

Final causation is triadic and serves to mediate between concrete *efficient* cause and some general *final* cause, the exact nature of which might not even exist at that time, although the possibility of it does exist. There is no reverse causation: the final purpose cannot determine the event, but it does mediate what types of *efficient* cause might occur to attain the possibility of the overall purpose.

Hulswit (2002, p 85) provides an example that has many parallels. Suppose a sportsman shoots a rifle at a target. The target (*final* cause) will not be hit without the rifle being shot (*efficient* cause), however, the rifle will not be shot without the sportsman purposefully loading, aiming and firing the weapon at the target – thus there will be no *efficient* cause without *final* cause. The laws governing the trajectory of the bullet are subject to general laws of physics. Even so, there may be infinitesimal or chance-like perturbations (wind shifts, frictions, barrel resonances, etc) that might spoil the directed path of the bullet to the target and result in a miss. In summary, the general intention of the sportsman is to hit the target, this intention is unrealised (but its possibility exists) when the rifle is fired and only realised when any chance-like perturbations fail to influence what was otherwise a perfect aim.

Peirce provided another analogy to amplify the relationship between *efficient* and *final* causation. He observed that a Court of Justice could not be imagined without Police (the Sheriff), for although the Police would have their power without a Court, they would lack efficiency. Thus *final* causation without *efficient* causation is helpless. *Efficient* causation without *final* causation he claims is worse than helpless – 'it is mere chaos' (Peirce 1935 I.220).

In summary, *final* causation is a teleological phenomenon which implies that there is a triadic relationship between individual *efficient* causes, un-realised but generally conceivable physical possibilities and individual effects that are subject to the possibility of chance. *Final* cause might also be *developmental* where it evolves from one concept to another over time. *Final* cause, whether conceptual or concrete, influences the sequences or processes of *efficient* causes that have the intention of achieving its purpose.

Van de Ven and Poole (1995) also cite *final* causation as an explanation for organisational development – one of their many 'motors' that drive organisational change. These authors highlight *equifinality* (Van de Ven and Poole, 1995, p 516) as the achievement of an organisational goal through the adoption of one of many equally effective ways. This is completely consistent with Peircean *final* causation (*developmental teleology). Efficient* causation (compulsive or natural laws) is employed to abet this developmental teleology which continues to progress rather than simply stabilise. Organisational progression under a teleological theory is - in many ways similar to Darwin's natural selection and as Peirce observed, a variety of actions derived from a set of possibilities to form an *a posteriori* developmental trajectory (Van de Ven and Poole, 1995, p 517).

3.8.3 Relating Peircean Causation to Benchmarking.

The erosion of Aristotle's taxonomy of causation to the sole category of *efficient causation* was arrested by Peirce who understood that anthropogenic activity is teleological and that causation was little different. Peirce also considered that even nature's habits, expressed in terms of compulsive laws (such as Newton's laws of motion), reflect concepts such as energy minimisation. Modern physics would hold this to

be either a contentious concept or one patently at odds with currently accepted theory where there is no need to introduce the notion of *final* or intentional causal explanation. Even though a ball will travel down a staircase, falling at each step (*efficient* cause), and is compelled to repeatedly do so in a sequence or process that only ends when there is no lower energy state that can naturally be obtained, one struggles to say that this is a teleological phenomenon. But one can say that in the totality of all human empiricism, perpetual motion has never been identified.

Peirce, the pragmatist, held that in the natural world it was inevitable that a sequence of *efficient* causes reached some final state. Peirce acknowledged that the word *teleology* might be too strong a word to apply to these natural laws and suggested the invention of a word such as '*finious*' to express the empirical tendency of such natural laws toward a final state (Peirce, 1935: VII.471). This may seem no more than a semantic as it does not alter Peirce's hypothesis that the absence of *final* cause makes *efficient* cause 'chaotic'.

The question must now be asked – does Peirce advance a theory of causation that is free from challenge? Hulswit (2000, IV) identifies several metaphysical issues that require further inquiry – such as the relationship between time and causation as well as reconciling Peircean causation to that of Hume (necessity) and others. Whilst answers to these issues are outside of the scope of this work, there is a pragmatic appeal arising from Peircean causation that assists our understanding of modern organisational practices such as benchmarking.

Peircean causation advances our understanding of benchmarking by addressing what earlier theories did not – the purpose (both fixed and developmental) and chance elements of change. However earlier epistemologies on *efficient* causation also advance our understanding of benchmarking and are not swept away by Peirce, quite the opposite. Peirce places *efficient* causation within a purposeful context that had, since Aristotle, either been attributed to God (theological) or, since Descartes, substantially disregarded as an unnecessary causal relatum in a deterministic world. By addressing this and by including the notion of chance, explanations that previously struggled now enjoy better illumination. For example, Hume would have immense difficulty in accepting that Peirce's bullet failed to hit the target, despite being perfectly aimed, except for an exhaustive post-mortem on the myriad of infinitesimal perturbations associated with the cartridge, primer, propellant, bullet, chamber, rifling, barrel and atmospherics that

impinged on this perfection. Darwin would have recognised Peirce's teleological umbrella over *efficient* cause and organisations will recognise a developmental teleology where their survival is a journey replete with successive milestones – each deemed necessary at the time, but none ever sufficient.

The combination of modern theories on *efficient* causation and the triadic framework of Peirce establish a robust platform to construct a theory of benchmarking. In particular, the inclusion of *final* cause establishes a pathway to include modern welfare theory as a basis for *developmental organisational teleology*– survival based on progression towards future wealth-generating possibilities.

3.9 A Summary of Causation.

Benchmarking has the purpose of effecting a favourable transition from one organisational state of affairs to another. Whilst this might only involve the attainment of a single and simple transition of some isolated organisational element, the epistemology of benchmarking holds otherwise. Current practices identifying the characteristics of what are certain to be a complex transitions and applying such knowledge effectively was shown in Chapter 2 to be *a-theoretical*. This summary addresses this difficulty by drawing upon established causation theories to identify how practitioners should approach the task of organisational benchmarking.

The history of causation from ancient to modern times suggests a circular epistemology where the Aristotelian four-part taxonomy collapsed to a single category only to be restored in modern times by Peirce. Not all of Aristotle's taxonomy was addressed by Peirce: specifically the *formal* and *material* causes, yet these remain important issues so far as organisational benchmarking is concerned and will be addressed via supervenience and entailment. The following sections crystallise the theories of causation into a practical framework that relates to organisational benchmarking. Peircean causation will establish the overall fabric of this framework, but insight from the epistemology of *efficient* causation and the ontology of substance will be included.

3.9.1 *Formal and Material* Causation.

Aristotle's examples of *material* and *formal* cause retain relevance in an organisational setting. History has not invalidated the need to ask 'why is this so', and those embarking on an organisational benchmarking exercise with the expectation of distinguishing between effective and ineffective efforts require knowledge of the composition of both the exemplar and anomalar organisations – their *material* cause. There is a corresponding need to also understand their culture, customs and practices that operate at both formal and informal levels to govern what they do– their *formal* cause. As noted by Davidson, causal analysability ('The fuller the description of the transition between objects in question, the more likely it is that light will be shed on whether or not there is a genuine causal relationship, what causal laws are in effect and whether the cause is regular or singular'.) is not simply applicable to *efficient* cause, but also to an understanding of the very nature of the exemplar and anomalar organisations themselves.

Identification of *formal* cause from the perspective of what might not be done when two organisations are compared is instructive. For example, tourism organisations seeking to benchmark against each other may not have *formal* alignment. The laws or rules governing employment, taxation, environmental compliance and liabilities may be significantly different. Whilst it is unlikely that their *material* alignment will be an issue, it is clear that even small changes in local conditions such as taxation (e.g. selective local visitor taxes) and environmental compliances, may make the imitation of certain states of affairs unachievable. Aristotle's ontology of substance highlighted the preservation of form in a transition, where form might not identify all that was possible, but would indicate what was not possible.

So benchmarking transitions where the anomalar seeks to capture all the effects of the exemplar are only possible where there is *formal* alignment between the respective organisations. Practitioners must undertake a systematic examination of rules or laws governing both anomalar and exemplar behaviour beforehand. This does not necessarily preclude benchmarking, but it establishes the limits of achievement, *a priori*, that might be expected from the exercise. *Formal* alignment is a necessary condition for organisational benchmarking.

3.9.2 *Efficient* Causation.

Causation theories stress the existence of conditions for change. Hume held that there were necessary and sufficient conditions, causal relata, that invariably produced an effect simultaneous with their being obtained. Ducasse criticised this by holding that singular transitions were possible and that Hume only explained the special case where the causal relata themselves were repeated. Critics of probabilistic causation identified the paradox of distinguishing statistical from deterministic behaviour prompting Psillos to urge that rather than speak of the *frequency* with which $A \Leftarrow B$, we should speak of the *objective chance* of $B \Rightarrow A$.

Mill's causation: $A \Leftarrow B$, iff $A \rightarrow B$ and $\{A\} \rightarrow \{B\}$, is the basis of conventional (Zairi's (1994) taxonomy) benchmarking described in Chapter 2. Anomalars observing exemplar behaviour and analysing their processes that place reliance only on the first part of Mill's claim, namely $A \Leftarrow B$, iff $A \rightarrow B$, must be mindful that this is only part of the proposition. Transferring the knowledge $A \Leftarrow B$, iff $A \rightarrow B$ without establishing that $\{A\} \rightarrow \{B\}$ is also true may mislead. In other words, there must be consistent causal relata between anomalar and exemplar.

The problems of necessary and sufficient conditions and the INUS also pose challenges for benchmarking practitioners. Determination of the set of necessary conditions, given Ducasse's singular causation may prove difficult in complex cases. For example, if a desired organisational state of affairs is found to be entirely dependent on a simple manufacturing process, it will be straightforward to establish the necessary conditions that obtain that process elsewhere. Benchmarking is straightforward in such cases. However, if the desired state of affairs has been achieved through a history of cultural, technological and marketplace changes, isolation of the necessary conditions and replicating them elsewhere is less straightforward and may even prove impossible.

A tourism example illustrates this - inter-country benchmarking. Attempting to benchmark New Zealand's tourism strategy against, say, that of Ireland is fraught with difficulty. Whilst there are many elements of comparison that are simple reflections of one another (e.g. promotion effectiveness or price competitiveness), the cultural, employment and economic rules that operate in each country differ considerably – thus making the benchmarking exercise a more limited affair. In such cases, it would be

optimistic to expect such a benchmarking exercise to identify more than a subset of the necessary conditions that might advance an anomalar's current state of affairs to that of the exemplar. However it would be quite realistic to harbour an expectation that benchmarking *could* result in the derivation of *some* advantage or contribute to planning processes via stimulation of fresh ideas. Estimations as to the degree of additional welfare that might arise under such circumstances will be considered in Chapter 5.

To summarise: benchmarking practitioners may draw the following additional conclusions from *efficient* causation theory:

- For each state of affairs, B, under analysis, necessary conditions must be established such that:

$$A \Leftarrow B, \text{ iff } A \rightarrow B$$

  Where A are the causal relata of the effect B, and to apply this knowledge to the anomalar's state of affairs where the more general relationship,

$$\{A\} \rightarrow \{B\} \text{ must also hold.}$$

- Invariance of causal relata must be established separately. Singular causes must be identified and included in the benchmarking exercise.

- Probabilistic causes must be established in terms of the objective chance of obtaining B given the occurrence of A, and further, that this applies in the case of $\{A\} \rightarrow \{B\}$.

3.9.3 Peircean Causation.

Organisations wishing to benchmark against others may need to consider that results arising from *formal* and *efficient* causation theory may be insufficient to achieve parity. Peircean causation theory holds that *efficient* causes are influenced by teleological processes that form a triadic contribution to the state of affairs under observation. This raises the question as to the degree to which complex organisational states of affairs can be benchmarked. Processes within this state of affairs that exhibit what Peirce called 'compulsive behaviour' present few problems not already encompassed by *efficient* causation. However the presence of significant teleological factors, such as an

organisational culture, that impinge on the state of affairs under observation can constrain the degree to which benchmarking can be effective.

An example that might serve to illustrate this might arise if Microsoft Corporation benchmarked significant states of affairs against Apple Corporation. Compulsive or natural processes in manufacturing, distribution and logistics may benchmark without significant constraint, but the teleological processes embodied in different organisational cultures may constrain the degree to which an overall benchmark could be applied.

Peircean causation also offers explanation of the difficulties experienced in organisational mergers and takeovers. Again, the integration of compulsive processes is relatively straightforward, but conflicting teleologies represent a different level of challenge. What is done is often as important as how it is done and why it is done, and an inability to align developmental teleologies would, according to Peircean causation, account for lower overall welfare than anticipated.

Benchmarking practitioners may draw the conclusion that the efficacy of an exemplar's organisational processes (*efficient* causation) may be inextricably linked to its teleological processes, which, if misunderstood or ignored, reduce the value available to the anomalar's organisation. As teleological processes are only evident *a posteriori*, historical benchmarking may benefit new organisations wishing to establish the initial conditions or a prevailing culture as they commence trading.

In Chapter 2, a powerful criticism was raised against benchmarking as it is retrospective. Peircean causation holds that events are irreversible as they are functions of time and chance. Of course similar results may occur at later times and compulsive or natural processes will invariably behave according to Humean or Mill causation, but the inevitability of an anomalar always lagging an exemplar is a theoretical feature of benchmarking.

3.10 Conclusion.

This chapter has reviewed the history of causation and concluded that the application of modern causation theories within the umbrella of Peircean Causation provides a theoretical basis for benchmarking. Causation theory addresses many of the theoretical criticisms levelled against benchmarking in Chapter 2.

As an example, Zairi's taxonomy is an effective practitioner tool that helps to identify various types of benchmarking. These types relate to various organisational situations. Causation theory is the relationship between antecedent relata and subsequent effects. For benchmarking, the locus of exemplary antecedent relata is immaterial since the (exemplary) effect is evidence that these antecedents have already been obtained. It remains to be discovered whether the conditions giving rise to these exemplary antecedents can be also obtained within the anomalar's organisation and if so, whether they will also improve matters. Causation theory guides the practitioner further – it identifies a *formal, efficient* and *final* (teleological) causation filter to establish these conditions and thus the effectiveness of the benchmarking process should be able to be gauged in advance of its implementation.

However, the extent to which a benchmarking process can deliver benefit to an anomalar remains unanswered. In cases where there are simple invariant processes, anomalars should expect improvements that equate to those of the exemplar. In other cases, this may not be so and it is necessary to extend Peircean causation by including economic welfare theory to quantify the potential gains available to an anomalar. The contribution of economic welfare will be addressed in Chapter 5.

Although causation and economic welfare appear to be empirical contributors to explanations of benchmarking processes, they have not yet been validated as theoretical contributors. The next chapter examines the constitution of theoretical constructions and the degree to which empirical (or other) contributors are either necessary or sufficient.

Chapter 4. Establishing a Valid Theoretical Framework: What is Theory; how do we know?

4 Introduction.

What constitutes theory in organisational studies? How do we know whether a construction is theoretical? This chapter examines theory construction and concludes that a given construction can claim to be potentially theoretical if it is validated against three criteria: consistency with, and extension of, presumed realities - epistemological validation; logical or causal validation and associative validation within or in extension of current organisational paradigms.

4.1 An Epistemological Approach to Theory Building.

Sumatra Ghoshal's (2005) posthumous paper criticised management theories that ran counter to good management practices. He cited Boyer's (1990) claim that scholarship consisted of four equally important facets: research, synthesis, application and pedagogy. Undue focus on only one of these facets – in particular a confinement of effort into research or discovery, simply created imbalances that, un-corrected, resulted in catastrophes. 'A theory must illuminate and explain and, if it cannot do those things, it is not a theory — neither good nor bad. Wishes and hopes are not theory. Sermons and preaching are not theory either' (Ghoshal 2005, p 86). An important distinction was also drawn between theories describing behaviours referred to in Chapter 3 by Peirce as 'compulsive' (natural laws) and those Ghoshal refers to as 'management' (organisational behaviour).

Theories encompassing compulsive or natural behaviour have the fortunate advantage that reality remains unaltered by the formulation and application of bad theory. For example, Ptolemaic theory concerning the motion of the heavenly bodies was geocentric, yet this theoretical error had no effect on the heliocentric nature of planetary motion. Errors were reduced by the theoretical perspectives of Copernicus (circular), Kepler (elliptical), Newton (gravitational) and Einstein (relativistic).

Management theories are quite the opposite. Ghoshal observes that a bad management theory may cause managerial behaviour to alter in accordance with it. This affects reality if bad theory becomes established practice.

The first example cited by Ghoshal is ‘liberalism’, expounded by Friedman (2002), that ‘freedom is the ultimate goal and the individual is the ultimate societal entity’ (Ghoshal, p 77) thereby leading to the placement of ethical issues at the feet of individuals rather than recognising them to be the relata of social theory. His second example is the organisational theory of performance centred upon the goal of ‘maximising shareholder value’. Ghoshal claims that indistinction between shareholder rights to receive economic surpluses and their rights to the assets and businesses of a company (itself a legal identity) has lead to a particularly narrow theoretical perspective of company performance – the application of which creates distortions that may not maximise the value of the company (Ghoshal, p 80). In contrast, theories acknowledging the reality of a broader spectrum of stakeholders and beneficiaries offer a more credible theory of organisational performance.

The lesson is that management theories (including those relating to organisational benchmarking) must contain more than ‘pretence of knowledge’ and possess common sense (Ghoshal, p 81).

Elster (1983) claimed that when applying theory, to natural sciences or humanities, there was ‘no difference between either the method of enquiry or the interests they served; there were only differences in the mode of explanation’. Three elements of theoretical explanation were identified by Elster: ‘Causal, Functional and Intentional’. Causal relates to compulsive or natural laws, ‘Functional’ relates to benefits, evolution or progress and ‘Intentional’ is synonymous with purpose (or teleology).

Physics, for example, is solely dependent on causal explanation. According to Elster (p 18) concepts such as benefits or teleology play no part in such theories. Peirce might not have wholly concurred with this distinction because in his view, explained in Chapter 3, even compulsive causal explanation (*efficient* cause) was also subject to ‘*finious’* purpose (a weaker concept than teleology, perhaps reflected by energy minimisation or entropy maximisation).

Elster claims Social Sciences are subject to intentional explanation (teleology), but functional and causal theories are relatively rare because in his view there is ‘no general law in social sciences comparable to that of natural selection in biology’ (p 20). Again, Peirce would not entirely agree as he identifies *chance* as a causal relatum which is, he says, ‘Omni-present, acting infinitesimally always and significantly, infinitely rarely’. In a

Peircean world, there is *always* a causal explanation comprised of *efficient* (including chance) causes and a *final* cause that might also be developmental.

Ghoshal claims that current management theories are overwhelmingly causal (*efficient* cause) or functional (Ghoshal, p 79), yet benchmarking is clearly intentional or teleological as it is performed *for the sake of* organisational improvement. Perhaps it is the failure of management theory to accommodate the three Elsterian distinctions (nominally Peircean causation) that leads to Ghoshal's criticisms: that bad or unethical purpose is often the outcome of applying these theories, as exemplified by welfare loss for many other important organisational participants aside from shareholders (i.e. no overall Pareto improvement[2]).

Bourgeois (1979) also lamented the absence of models that aid construction of good theory. He described a somewhat linear seven-stage process that addressed what he called 'middle-theory', citing Merton (1968) as 'theory that lies between day-to-day working hypotheses and unified theory'. These stages were categorised as follows: (1) partitions the topic under investigation, (2) method of theory construction, (3) review of literature, (4) construction of theory (induction from empirical base), (5) extension of theory (deductions into propositions), (6) metaphysical elaboration and (7) conclusion.

This sequence is indicative of normal research, except perhaps for stage (6). By metaphysical elaboration, Bourgeois meant 'philosophising': to express 'ideas and deductions that cannot, because of their speculative or perhaps un-testable nature, be properly subjected to the rigour of the analysis that middle-range theory requires' (Bourgeois, p 445). Even though the approach appears linear, it is still a combination of deductive and inductive reasoning (Bourgeois, 1979, p 446).

Popper (1987) assumes a broader perspective: the correct approach to any nascent theory is *a posteriori*; the acceptability of its results to those dependent upon them – presumed to be its ability to explain what is known and to create new knowledge.

[2] Given a set of alternative allocations of, say, goods or income for a set of individuals, a movement from one allocation to another that can make at least one individual better off without making any other individual worse off is called a **Pareto improvement.**

Weick (1989) also examined a range of approaches and distinguished between treatments of theory as linear exercises versus a more evolutionary approach he termed 'sensemaking' (Weick, 1989, p 519) – 'the ordering of relationships that constitute the theorist's focus of attention in the real world':

'Good theory is plausible, and more plausible if it is interesting (versus obvious, irrelevant or absurd), obvious in novel ways, a source of unexpected connections, high in narrative rationality, aesthetically pleasing or correspondent with presumed realities' (Weick, 1989, p 517).

Llewelyn (2003) claims that theorising can occur at 5 levels – ranging from the use of metaphors to differentiation between forms, conceptualisation, context-bound theories and general theories. The first three of these present some difficulties as their contribution to theory had been challenged earlier by Pinder and Bourgeois (1982) and Morgan (1983) as being based upon deficient causal relata: i.e. generally unreflective of *formal*, *material* and *efficient* cause. These deficiencies are easiest to see in the metaphor, but even conceptualisation suffers. Indistinction between *seems* and *is* threatens the role of conceptualisation as a theoretical relatum in much the same manner as it does to the metaphor. Recall Hamlet's response: 'Seems, madam! Nay, it is. I know not 'seems'' (Shakespeare's Hamlet, I, ii, 76). This leaves context-bound and generic theories as more robust examples to strive for.

Weick's 'sensemaking' and Poppers' acceptability of results are two tests that will be applied to a theoretical construction of benchmarking.

4.2 A Logical Approach to Theory Building.

Theory has also been addressed counterfactually, from the perspective of what it is not. Sutton and Staw (1995) isolated '*a-theoretical*' constructions from analysis of scholarly submissions to the *Administrative Science Quarterly* journal. They identified a lack of consensus as to what the word 'theory' actually meant, but described five appealing but *a-theoretical* approaches (or errors) to theory-building. These were 'References, Data, Variables, Diagrams and Hypotheses'. Weick (1995) quickly criticised the isolation of these five approaches by noting that their 'mere presence' was insufficient for condemnation as they were also to be found in the process of 'theorising'. The difference, according to Weick (1995, p 390), is an unconscious expectation that these

approaches constitute theory rather than a conscious explanation that they are part of the struggle that accompanies the establishment of theory – a process that refines approximations to advance scholarship (Weick, 1995, p 385).

With this criticism in mind, it is useful to review these five approaches, appraise them in terms of causation theory and then apply them as a filter in the development of a theory on benchmarking. Recognising that the process of theorising is developmental, Sutton, Staw and Weick conclude that undue reliance on the following approaches will likely result in *a-theoretical* constructions.

References. Citation of existing theories may be necessary but is an insufficient basis for theory construction. It is a necessary condition that any streams of logic underpinning existing theories are able to be explicated and demonstrated to be causally similar to the matter at hand. This is a re-statement of Humean-Mill causation where:

$$A \Leftarrow B, \text{ iff } A \rightarrow B \text{ and } \{A\} \rightarrow \{B\}, \text{ (Chapter 3.5.8).}$$

Data. Empirical evidence is a necessary but insufficient condition to revise, discredit or confirm existing or developing theory. It is a further necessary condition that any data also supports the second part of Humean causation; $A \Leftarrow B$ iff $\{A\} \rightarrow \{B\}$ rather than data that solely supports the first part: $A \Leftarrow B$, iff $A \rightarrow B$.

Variables and Constructs. Identification of variables and their definitions is an important and therefore necessary part of any theory, but it is a further necessary condition that these variables and constructions be shown to be connected to the theory via its hypotheses. This is also a re-statement of Occam's razor: a theory should be obtained from the essential minimum number of variables and constructs.

Diagrams. Depictions of causal processes are neither necessary nor sufficient conditions for theory construction but may assist with explanation of causal relationships and address what Ghoshal termed the pedagogical element of scholarship (Ghoshal, p 82).

Hypotheses. Concise statements about what is expected to occur rather than why it does occur. Hypotheses are necessary but insufficient conditions for theory construction to explicate the presence and operationalisation of variables and their constructions. Sutton

and Staw (1995) concur with Dubin's (1976) claim that 'a theoretical model is not simply a statement of hypothesis'.

This list of *a-theoretical* approaches has also exposed a further set of necessary conditions whose conjunction might now obtain a sufficient condition for potentially theoretical constructions.

If a potential theory is the result of remedying what is held to be *a-theoretical*, Sutton and Staw's observations reduce to three sets of conditions: two of which are principal and one consequential;

1. Hypotheses that explicate the presence and operationalisation of variables and their constructions form a principal condition (variables entail hypotheses),
2. Causal logic that associates the effects of the theory with its antecedent hypotheses and variables also form a principal condition,
3. Supervenience of hypotheses on variables and constructs forms a consequential condition.

Supervenience of hypotheses upon variables and constructs is a statement about the nomological relationships between the environment, what is expected to occur and, when combined with entailment, reflects *formal* cause – preservation of properties and laws in causal relata (see Chapter 3 and Appendix 2). Hypotheses should establish the relationships or concurrences between data, as expressed generally in terms of relevant environmental variables and outcomes (events). The properties of any variable within a hypothesis should be the same as its properties outside of it (entailment, *formal* cause). Whilst causal logic is a necessary condition it is insufficient as 'relationships' might still be circumstantial or ambiguous. Elimination of these possibilities falls to the third condition, supervenience.

4.2.1 Deriving Necessary Conditions for Theory.

A formal crystallisation of Sutton and Staw's objections may now be made.

Suppose that a certain theory T is conjectured. T necessarily characterises the set of Effects {E} arising from the conjunction of a set of observable environmental Variables {V} and a set of associated Hypotheses {H}.

Suppose that {H} is comprised of *M* elements that portend the occurrence of a certain set of Effects, {E}. Let {V} be comprised of *N* Environmental Variables that are associated with {H}.

For Individual environmental variables V_x where $x \in N$ and individual hypotheses H_y where $y \in M$, then every V_x having a property, F within {V} there is also at least one property G in {H} such that if another variable V_y has the property G (necessarily) everything having G also has F. This is to say that {V} strongly (necessarily, $\Diamond$) supervenes on {H} if:

$$\Diamond \forall x \, \forall F \in V \, [F_x \rightarrow \exists G \in H \, (G_x \, \& \, \Diamond \, \forall y \, (G_y \rightarrow F_y))].$$

If this were otherwise, the behaviour of variables outside of any hypothesis could differ from their behaviour inside the hypothesis. Clearly all variables must be constructed consistently within the hypotheses and that all hypotheses must be constructed from consistent variables.

If the conjunction of {V} and {H} is the antecedent of the effect {E}, then for theory T to characterise the causal linkages of {V} and {H} to {E}, the Humean *efficient* causation identity must be obtained: i.e.

$$\{V\} \wedge \{H\} \Leftarrow \{E\}, \text{ iff } \{V\} \wedge \{H\} \rightarrow \{E\} \text{ and } \{\{V\} \wedge \{H\}\} \rightarrow \{\{E\}\}$$

This says that the Effects are *caused* by the conjunction of environmental variables having certain properties and acting in certain ways (supervenience of hypotheses) provided this not only occurs in specific cases but also in the general sense where 'things like {V} acting in a manner similar to {H} also produce effects like {E}' as well. This latter condition imposes invariance on the relationships and serves to eliminate ambiguities or accidents (but not singular causation).

The combination of supervenience and the *efficient* causation identity converts Sutton and Staw's objections into a set of conditions that indicate *potentially theoretical* constructions.

Casual relationships are not necessarily Humean. According to Peirce, Humean conditions are insufficient to generally explicate effects – even though they may likely

suffice in many situations. Application of the Peircean Causation model (Chapter 3. 8.2) in organisational contexts may prove challenging as any element of *efficient c*ausation (singular, probabilistic, INUS, etc) is contextual – as it also entails purpose and may also be subject to chance. Hence knowledge of *efficient* cause without knowledge of its contextual purpose or the limits of probabilistic variance is, for organisations, incomplete. Unfortunately elements of *final* causation and chance may either be unknown or unobservable.

Similarly, any singular relationships between the effects and the conjunction of the environmental variables and hypotheses may also prove difficult to obtain. Thus theory risks explanatory gaps where it does not encompass contributions arising from Teleology, Chance and perhaps Singularity.

4.3 Theory-building from Association with Multiple Paradigms.

Benchmarking is a teleological organisational phenomenon. However there appear to be many rather than just a single or overarching paradigm that explains organisational phenomena. A robust theory of benchmarking should be encompassed (entailed) by any organisational paradigm that admits (supervenes upon) 'improvement' as a valid teleological construct.

Gioia and Pitre (1990) summarised the issues arising from paradigmatic scholarship into three categories that reflected fundamental but differing organisational assumptions. These were 'ontology' (nature of organisational phenomena), 'epistemology' (nature of organisational knowledge) and 'methodology' (nature of ways of studying organisational phenomena). Gioia and Pitre (p 585) cited Burrell and Morgan's (1979) taxonomy of organisational paradigms to identify four research paradigms, shown in Figure 4-1, that encompassed these different organisational assumptions.

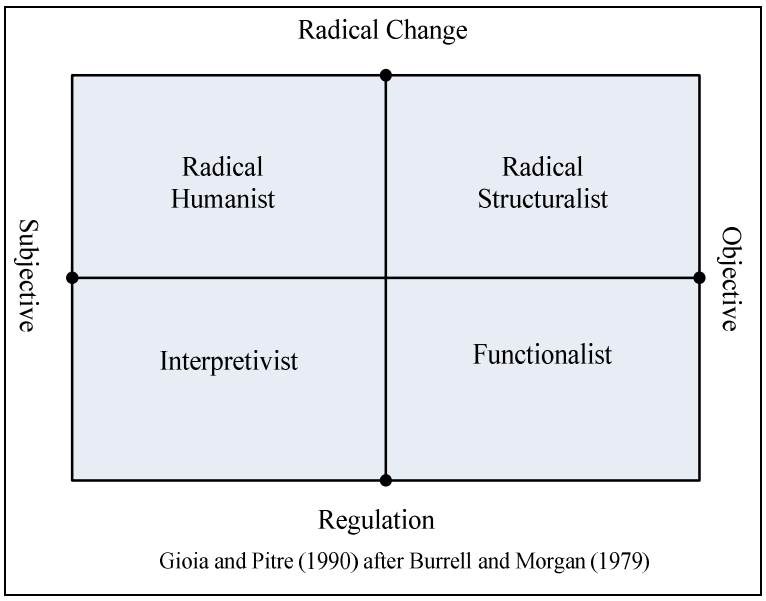

Figure 4-1. Four Organisational Paradigms or Perspectives.

These four paradigms reflect the different ways of addressing organisational behaviour and the important observation for this study is their overall incommensurability – even though there may be indistinct boundaries between each of them. The question is how can theory be built within the constraint of multiple paradigms when there is no unifying thread linking them?

Willmott (1993) criticised the notion of incommensurability and reflected on Kuhn's contention that there was continuity as well as incommensurability in the revolutionary process of scientific discovery. Popper (1996) also noted that incommensurability involved *formal* rather than *material* relata and did not preclude discussions or comparisons between these 'so-called' paradigms. Indeed, the use of the term '*paradigm*' by Burrell and Morgan and its re-emphasis by Gioia and Pitre appears different (softer) to Kuhn's stricter usage of that term. Willmott (p 686) repeated Kuhn's emphasis on the nature of knowledge transition: 'that there will be large but never complete overlap between the problems that can be solved by the old and by the new paradigm'. Willmott proceeded to question whether social science is pre-paradigmatic as the mature development of an accepted, single paradigm has yet to occur (Willmott, p 687). Even so, it is useful to examine benchmarking within, what might less controversially be called, these 'different organisational perspectives' to gauge its contribution to each of them.

4.3.1 Functionalist and Interpretive Perspectives.

The Functionalist perspective, reflects a deductive methodology that embraces what Van de Ven and Poole (2005) would characterise as a '*Variance'* methodology (See Table 4-1). Organisations under this perspective are shaped in 'fairly deterministic ways' (Gioia and Pitre, p 590). Hypotheses are based on the identification of antecedent relata that give rise to identified effects. *Efficient* causation would describe the basis of 'Functionalist' hypothetical development with its accompanying causal attributes as described in Chapter 3: immediateness, existence of necessary and sufficient causal relata and invariance of outcomes once conditions are obtained.

Variance Approach	**Process Approach**
Fixed entities with varying attributes	Entities participate in events and may alter over time
Explanation based on necessary and sufficient causal relata	Explanation based on necessary causal relata
Explanation based on *efficient* causation	Explanation based on *final*, *formal* and *efficient* causation
Generality depends on uniformity across contexts	Generality depends on versatility across cases
Time ordering among independent variables is immaterial	Time ordering of independent events is critical
Emphasis on immediate effects	Explanation is layered and incorporates both immediate and distal effects
Attributes have a single meaning over time	Entities, attributes & events may change in meaning over time

Table 4-1. Process and Variance characteristics of organisational change.

Modified from Van de Ven and Poole (2005, p 1382)

An 'Interpretive' perspective, according to Morgan (1980) and Gioia and Pitre (1990), reflects a social world with a precarious ontology where participants construct and sustain their own organisational realities. Rule–based processes determine the outcome of structures and the stance towards theory building is one of becoming part of the evolving events experienced – a perspective that reflects those within the organisation experiencing the structuring processes. Whereas the functionalist perspective sees order or determinism, the interpretive theorist is closer to what Van de Ven and Poole (2005, p 1382) might describe as adherents to a '*Process'* methodology. A '*Process'* methodology espouses versatility, changes in the meaning of things and evinces what has already been described in Chapter 3 as Peircean Causation – the triadic nature of *efficient* and *final* cause as well as their subsequent effects. Although Van de Ven and Poole (2005) cite Peircean Pragmatism (*not* Peircean Causation), the teleological, or

more particularly the *developmental teleological,* nature of their *Process* approaches to explanations of organisational change align within Burrell and Morgan's (1979) Interpretive perspective and Peircean causation.

Both the Functionalist and Interpretive perspectives admit the utility of benchmarking – either as a means of objectively improving an existing organisational attribute or as an aspirational trigger to develop a better reality. For continuity, the remaining two perspectives will also be examined briefly and the challenge of establishing a benchmarking theory within this framework will be revisited.

4.3.2 Radical Humanist and Radical Structuralist Perspectives.

These perspectives reflect some of the attributes of their axial partners but each introduces the notion that there are higher extremes that govern the scope of organisational progress. To the Radical Humanist, this scope is not bounded by current structures but rather examines their legitimacy and replacement as a central ideology (Gioia and Pitre 1990, p 589). On the other hand, the Radical Structuralist holds that there is a concrete reality independent of current perception and practice. The tensions and contradictions that constrain attainment of this reality are to be understood and altered through either 'radical action or conflict' (G. Morgan, 1980, p 609).

In the case of Radical Structuralist the evolution of anomalars provides evidence of processes, underpinned by organisational purpose, that have lead to unsatisfactory states of affairs. Again, benchmarking operates at anomalar level to identify the relationships between processes and purpose so as to expose opportunities (potential pathways) leading towards a more desirable state of affairs whose purpose also evolves beyond what is current. Whilst the radical humanist perspective eschews reliance on a developing purpose, the radical structuralist perspective relies on the espousal of purpose and adjusts processes (radically if necessary) so as to achieve it.

4.4 Using Meta-paradigmatic Approaches to Theory Building.

How do such diverse perspectives, apparently incommensurate in nature, affect benchmarking theory? Poole and Van de Ven (1989, p 567) eschew the need to integrate theories or paradoxes. They emphasise the importance of accounting for these perspectives and the use of concepts such as triangulation (meta-triangulation) to achieve what might graphically be depicted as a mezzanine-floor or meta-paradigm perspective of

them. A theory that co-exists within these four perspectives should accommodate rather than embrace their key attributes even though that theory might be rooted in any particular one of them.

Lewis and Grimes (1999) commented on the need to integrate Weick's (1989) 'disciplined imagination' and 'sensemaking' with 'metaconjectures' – the development and testing of questions that might be explained through alternative lenses. They illustrated two techniques for this: conjecture inversion and conversations. Conjecture inversion involves re-framing a broad question within multiple paradigms and examining how it is explained. Conversations probe paradigm debates to discover creative means of identifying and justifying contradictions. These two techniques provide a basis for a prima facie consideration of the application of benchmarking to Burrell and Morgan's organisational perspectives.

4.4.1 A Multi-Paradigm perspective of Benchmarking.

The first task in addressing the application of a theory of benchmarking that embraces multiple organisational perspectives is to establish its natural locus. Causation theory, particularly Peirce's triadic theory explains benchmarking's role as a purposeful phenomenon that helps an organisation to advance to a superior state of affairs. If this is the case, it should exist within each of Burrell and Morgan's four organisational perspectives, Figure 4-1, as none of them eschew the objective of improvement, only the mechanism(s) for achieving such improvement. The locus of a general organisational theory within Burrell and Morgan's perspectives is depicted in Figure 4-2

A particular theoretical perspective, represented by the fluid meta-level theoretical platform, might be located above a particular paradigm to indicate its roots, but is cognisant of the 'structures' of the others and whilst it may not or cannot integrate them, can offer explanations as to disparity or congruence. Benchmarking assumes only the existence of anomalars and exemplars rather than their paradigmatic locus. As shall be enlarged upon, not all of these organisational perspectives necessarily possess exemplars, although all possess anomalars.

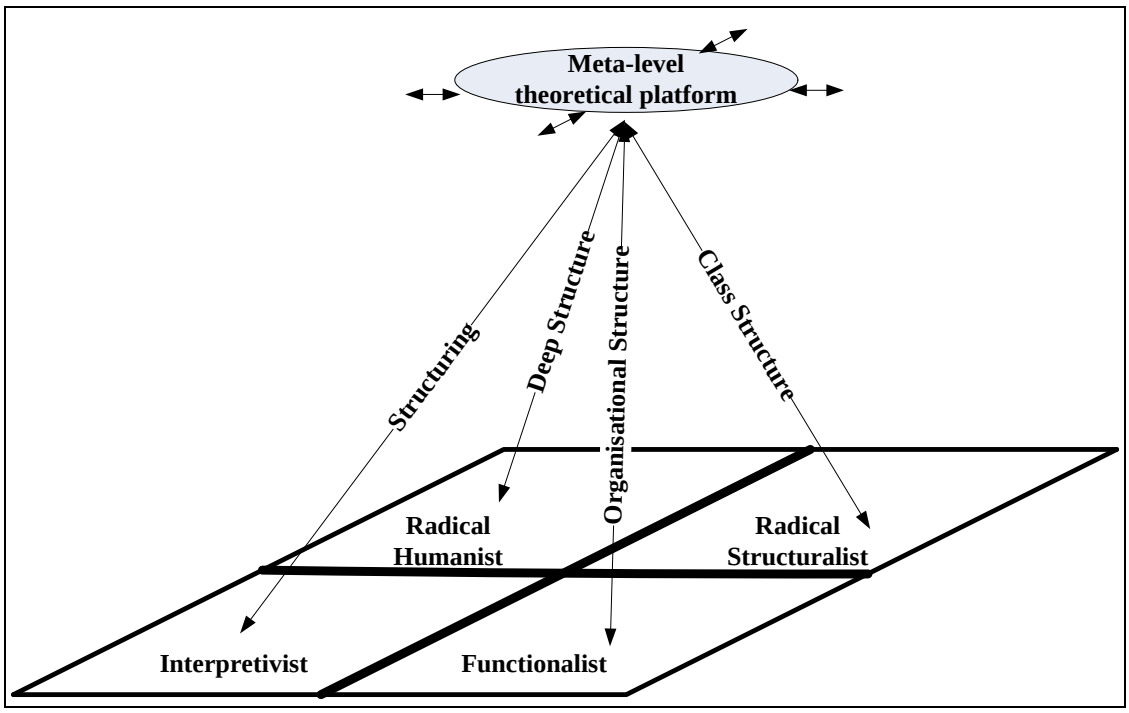

Figure 4-2. Locus of Meta-level Theories.

After Burrell and Morgan (1979), and Gioia and Pitre (1990, p 597)

Benchmarking, consistent with Peircean causation, has a triadic nature that accommodates the functionalist and interpretive paradigms: *efficient c*ausation coupled with developmental teleology addresses both variance and process approaches to Organisational Change described by Van de Ven and Poole and illustrated in Table 4-1.

Functionalist paradigms may also be viewed at a more fundamental level through the lens of *efficient* causation where behaviours satisfy Humean-Mill causation conditions.

In the case of the Radical Humanist perspective, current structures are necessarily reflective of unsatisfactory states of affairs. Benchmarking may form a trigger to pursue a more desirable state of affairs based not so much on the existence of exemplars, but rather on the preponderance of anomalars and their observed performance (variance) rather than the combination of their performance within a purposeful framework. To the radical humanist, there is a rational variance between the current and desired (hypothetical) state of affairs, and this variance can be eliminated through radical change. Benchmarking assembles the causal relata of anomalous states of affairs and in doing so may serve to identify the relata needed (or not needed) for their replacement. The issue with such an approach is that a causation-based benchmarking theory could

only establish a set of necessary, but insufficient conditions for the attainment of some new, proposed state of affairs since such a state is hypothetical.

Radical paradigms (perspectives) might also draw upon benchmarking, although in these cases to establish the current unsatisfactory states of affairs. Benchmarking under both radical perspectives can be described as the case where anomalars appear everywhere and exemplars nowhere! The trajectories of both Humanist and Structuralist perspectives are driven by the objective (teleology) of attaining radically different exemplary states of affairs, altogether devoid of anomalar inadequacy.

For both radical perspectives (Humanist and Structuralist), the conclusion is that the full power of benchmarking is available at anomalar levels where anomalars are transformed into exemplars of inadequacy.

Radical humanist benchmarking is Humean-Mill in nature where '*efficient'* causal relata serve to identify how anomalous variance occurred and provide a basis for their inclusion or exclusion in the attainment of an idealised exemplary state of affairs. The concept of supervenience may be set aside as new (radical) environmental variables and rules may be required to eliminate this variance.

Radical structuralist use of benchmarking is Peircean – the triadic nature of currently anomalous states of affairs, is examined with respect to purpose and process. Again supervenience may be set aside as the new, driving teleology could disestablish anomalous processes altogether. Implementing Peircean benchmarking within a radical structuralist context poses challenges as the relationships between '*efficient'* and '*final'* cause may not be evident. There may also be the added complication of isolating any dynamic developmental teleologies contributing to observed states of affairs. However Peircean benchmarking conceptually explains the nature of a process-orientated approach and supports organisational change driven by either a fixed or developmental teleology to change away from current structures based on knowledge of their causal relata.

These observations suggest that the locus of any benchmarking theory is a central feature within the Gioia and Pitre (1990) diagram depicted in Figure 4-3.

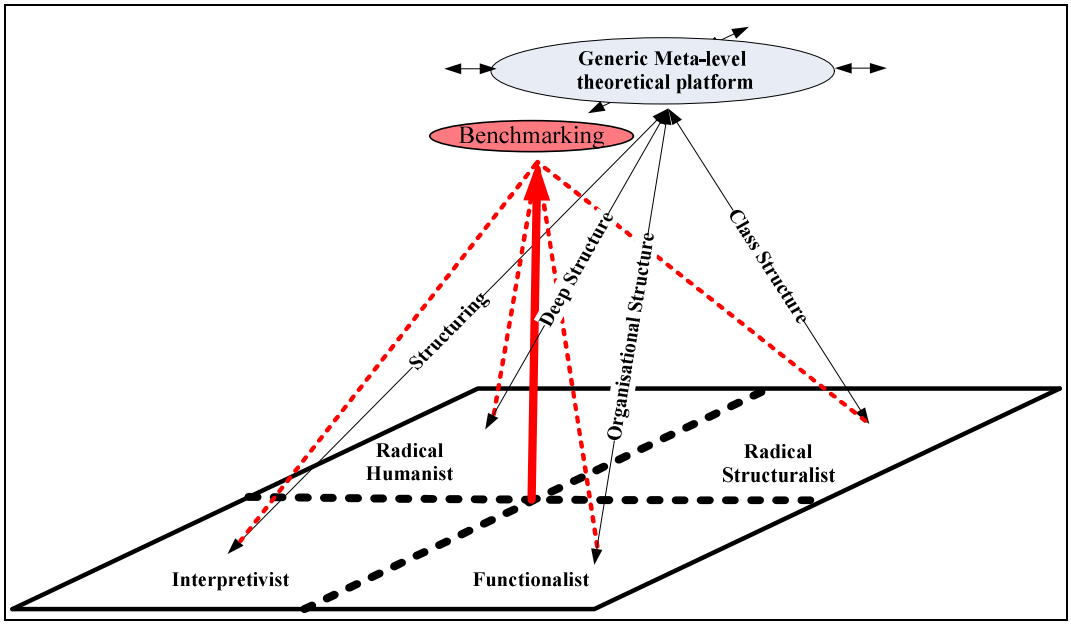

Figure 4-3. Meta-paradigm Locus of Benchmarking.
After Burrell and Morgan (1979), and Gioia and Pitre (1990, p 597)

Such a placement indicates that benchmarking theory is *a-paradigmatic*, and can explain or abet the purpose of organisational improvement by reference to exemplary states of affairs, or, alternatively, a purpose for organisational improvement by reference to anomalous states of affairs. Benchmarking can exemplify what *should* be as well as what *should not* be part of the trajectory towards organisational improvement.

4.5 Summary of Theory-building Relata.

Three approaches have been examined to answer the question: what is theory? The first approach considered theory-building from the empirical observations of scholars and practitioners, the second examined theory-building counterfactually - from the perspective of what it was *not* thereby constructing a set of necessary logical or causal conditions to establish a closer approximation to what it potentially *is*. The final approach considered theory-building as an associative opus where co-existence with and explanation of its locus amongst prevailing organisational paradigms (perspectives) was also a necessary condition. None of these approaches has isolated sufficient conditions for theory-building. This is not unexpected since history demonstrates that theory is seldom static and its trajectory towards perfection is driven by both evolution and revolution.

To progress from an empirical and a theoretical encompassment of benchmarking to a theory of benchmarking the following validations, filters or conditions will need to be met

to concord with the conjunction of all identified necessary conditions for theory building or validation of a nascent theory.

4.5.1 Empirical Conditions.

The concept of 'sensemaking' espoused by Weick is an obvious starting point. Elster strengthens this by the addition of three theoretical relata: cause, function and intention. Two of these, cause and intention, may be further related to Aristotle's causal theory of explanation as respectively: *efficient* and *final* cause. However Elster's 'function' theoretical relatum refers to benefits, evolution, adaptation and selection – most evidently seen as a biological phenomenon. Peirce also recognised this relatum and it formed part of his triadic theory of causation as the element of 'chance'. Ghoshal referred to the effects of bad theory and lamented tendencies to quarantine explanations to selected circumstances whilst ignoring evident issues arising from the application of such a theory in practice – i.e. disregarding Hume and Mill's causal invariance. The systematic and structural approach of Bourgeois approached theory in a project management-like manner but cautioned against a wholly linear approach as evidence of inductive and deductive reasoning was also necessary.

The combination of these observations provides a set of necessary conditions against which any new theory should be gauged. For a new theory, these relata represent the 'data' against which the theory is to be validated:

- **Sensemaking**: generally correspondent with presumed realities.

- **Causal**: relationships between all theoretical relata are explained in terms of *efficient*, *formal*, chance and *final* (teleological) cause. *Material* cause is presumed consistent amongst organisational environmental variables.

- **Complete**: peripheral issues are identified and quarantined rather than ignored. This recognises the broader condition within Humean-Mill causation where it is insufficient to hinge relationships on the particular; they must also apply in general – i.e. 'things *similar* to X must also produce effects *similar* to Y'.

- **Structured**: Theory construction is consistent with the methods and principles of normal science.

4.5.2 Logical Conditions.

Critical examination of *a-theoretical* constructions establishes a set of necessary but insufficient logical conditions that apply both generally and particularly to the environmental variables and their associated hypotheses (relationships). One methodological condition and two logical relationships and were established in Section 4.1.1. These established three necessary logical conditions:

- **Hypothetical Construction**: a methodology explaining a set of effects {E} formed from the conjunction of a set of hypothetical propositions {H} and a set of environmental variables {V}. Environmental variables entail hypotheses.

- **Supervenience**: If there are two arbitrary properties F and G associated with *x* elements of {V} and *y* elements of {H} then in accordance with *formal cause*, {V} supervenes on {H};

$$\Diamond \forall x \, \forall F \in V \, [F_x \rightarrow \exists G \in H \, (G_x \, \& \, (\Diamond) \, \forall y (G_y \rightarrow F_y))].$$

- **Causation**: Whether a Humean-Mill or Peircean causation model is used, theory should relate causal relata ({V} and {H}) to both observed and predicted effects ({E}). The following logical relationship should hold:

$$\{V\} \wedge \{H\} \Leftarrow \{E\}, \text{ iff } \{V\} \wedge \{H\} \rightarrow \{E\} \text{ and } \{\{V\} \wedge \{H\}\} \rightarrow \{\{E\}\}$$

There is a consistency between the logical and empirical sets of conditions for theory-building. The first logical condition is consistent with the empirical criterion 'Structured' and the second and third logical conditions are also reflected in the 'Causal' and 'Completeness' criteria cited in the preceding section. The empirical concept of 'Sensemaking' is reflected in the second part of the third logical condition where sense is made of general similarities, rather than only particular situations. The consistency criterion addresses Ghoshal's concern where unsatisfactory side-effects of a particulars-based theory have been disregarded rather than viewed as a violation of necessary causal conditions for related and general circumstances affected by such theory building. In essence, the logical conditions underpin the empirical conditions and provide a firmer platform for theory-building.

4.5.3 An Associative Condition.

There is one further condition that amplifies and extends the concept of 'sensemaking' rather than establishes an altogether new condition. Unlike the field of Mechanics, where there has been a long-standing dominant paradigm that successfully explains all but the most insignificant elements of observed physical behaviours, organisational behaviour is encompassed by numerous perspectives (as noted before 'perspectives' might be a more appropriate term than paradigm), but none of them dominant. Currently accepted organisational paradigms, such as those of Burrell and Morgan, structure organisational behaviour within orthogonal continuums (objective to subjective; regulated to radical) and provide platforms for 'sensemaking'. As these are the best perspectives available for organisational explanation it is appropriate to project emerging theories of organisational phenomena onto them to identify areas of concordance or conflict. However this raises a number of issues requiring reconciliation;

- Incommensurability: perspectives held to be incommensurate pose problems for theories addressing pan-organisational phenomena
- Paradigmatic evolution: the evolutionary struggle between emergent organisational perspectives imposes a similar dynamic on pan-organisational theories.

Meta-level approaches have been described as a mechanism for reconciling the co-existence of pan-organisational theory within orthogonal frameworks. The ideal locus of a proposed pan-organisational theory is where it is biased above its parent perspective yet its effects may be cast upon all other perspectives. Where a pan-organisational theory is deeply biased towards one of the currently accepted perspectives it must also tolerate evolutionary struggle as challenges over dominance mount. If a theory is radical, it should still offer explanation within existing perspectives provided such explanations are scientifically based (consistent with observed reality and natural laws) but it may also project beyond them to obtain explanations not previously encompassed or do so more elegantly.

A pan-organisational theory that can explain and predict ('sensemaking') successfully in all currently accepted perspectives is stronger than one that cannot. This establishes an additional condition for theory-building:

- It is a necessary condition that pan-organisational theories exhibit 'sensemaking' within currently accepted organisational perspectives.

This amplifies the first empirical condition and is consistent with supervenience and causation.

4.6 A Filter for the Validation of Organisational Theories.

The empirical, logical and associative conditions developed for theory building represent different components of endeavour all seeking to establish the same purpose. The concept of 'sensemaking' is broad and overarches theory building, but is imprecise and benefits from triangulation. Examination of three different perspectives of theory-building triangulates to a consistent set of necessary conditions (or validations) that are linked to the evolutionary advancement of knowledge. Proposed or nascent organisational theories are filtered (or validated) against these necessary conditions and either advance as potentially *theoretic* or are rejected as *a-theoretic* if necessary conditions are not obtained. This process is shown diagrammatically in Figure 4-4.

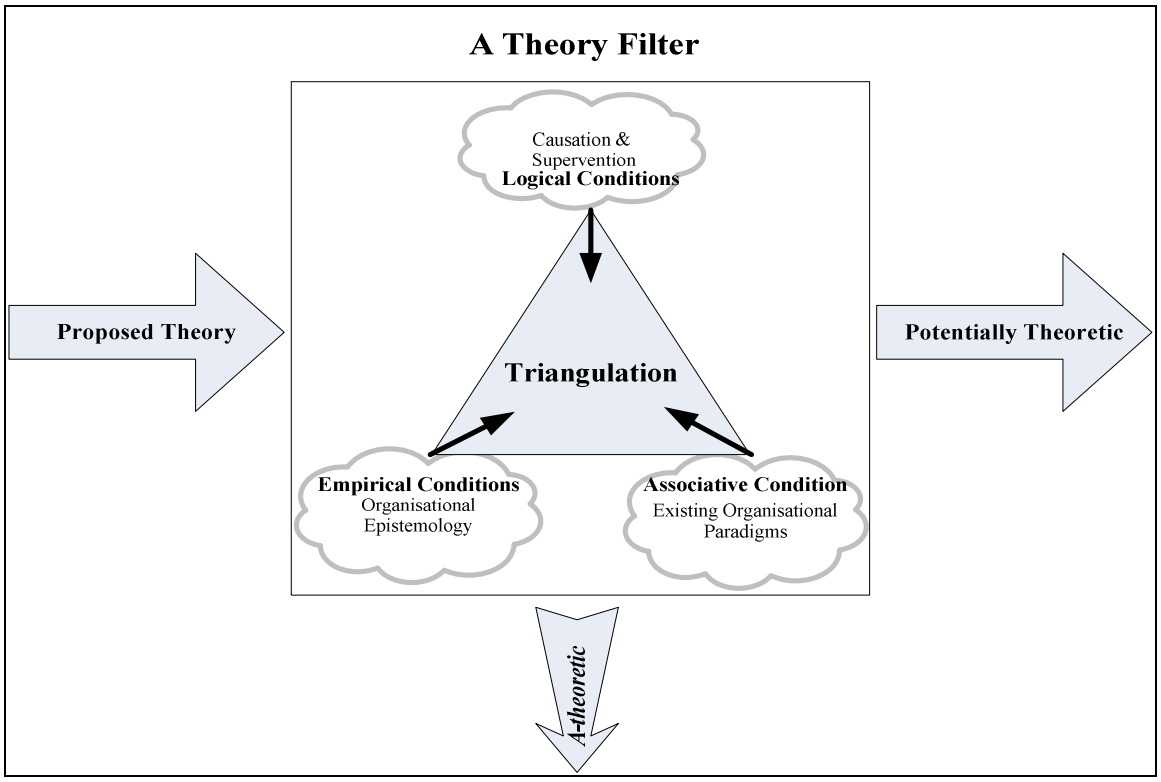

Figure 4-4. A Filter (or validator) for the Advancement of Organisational Theories.

It has not been shown whether these conditions form a closed set. No sufficient conditions have been found and it is ventured that there are none as knowledge itself is evolutionary

and necessary conditions are based on current scholarship.

All that can be said with confidence is that the set of necessary conditions established earlier provide a filter or validator that rejects *a-theoretical* constructions and passes *potentially theoretical* constructions consistent with current organisational perspectives whose standing is based on Popper's 'acceptable outcomes'. Advancement of a nascent theory from the status of *potentially theoretical* to *theoretical* is based solely on its ability to satisfy whatever demands may be made of it by future research and practice.

The next Chapter establishes and examines benchmarking's relata with the objective of satisfying the necessary requirements indicated by this Theory Filter.

Chapter 5. Developing a Theory of Benchmarking.

5 Introducing the theoretical relata.

Earlier chapters have examined the epistemology of Benchmarking and Causation with the latter drawing the conclusion that Causation Theory offers explanations as to the driving forces within a benchmarking exercise. However the important question as to what practitioners might expect to accomplish from a benchmarking exercise extends beyond knowledge of these driving forces. Addressing this question requires two further theoretical contributions: definition and quantification of prevailing anomalar and exemplar states of affairs as well as their congruence with organisational purpose.

The previous chapter examined the nature and composition of a theoretical framework. It concluded that three structural elements were necessary for the elimination of *a-theoretical* constructions: logical, epistemological and associative, but also concluded that none of them were sufficient. Even so, the most sensible starting point in the development of any theory is the establishment of logical or causal relationships between its relata.

Benchmarking relies upon the identification of exemplary performance and distillation of the differences that characterise it from anomalous performance to establish improvement strategies. Establishing consistent logical relationships between the relata of the exemplar and anomalar elevates benchmarking from a pragmatic to a theoretical plane. Doing so in a manner that makes sense within a broad range of organisational contexts addresses the other structural elements of theory-building and reduces the likelihood of producing *a-theoretical* explanation.

This chapter applies the findings of these earlier chapters with the objective of establishing logical relationships that describe benchmarking within an organisational framework whilst retaining coherence with established theories of Causation and Economic Welfare.

5.1 Organisational States of Affairs.

In the process of developing a theory of benchmarking, two key elements have arisen: the nature of purposeful change and the concept of a state of affairs. Anomalar behaviour is reflected in disparities between the nature and deployment of its own resources and those

of an exemplar. The term ‘state of affairs’ has been used to identify the behaviour and disposition of particular organisational resources and it is important that this term be defined within a context of purposeful change.

5.1.1 Defining a State of Affairs.

The ‘sake’ for which benchmarking is performed is purposeful organisational improvement. But why improve? Benchmarking is associated with purposeful organisational improvement, yet it cannot be said to represent a *final* cause. Purposeful improvement may be obtained without benchmarking and achieved in so many different ways. Improvements represent streams of activities or developing teleologies that are, as Peirce observes (1935, VI.156) directed towards some more fundamental or explicit *final* purpose.

Survival is an inescapable *final* purpose of any organisation; it is a *sine qua non* relatum in any organisational teleology since ‘its pursuit is continuous and its attainment never automatic’ (Pfeffer, 1997). Peirce observed that processes behind this pursuit might not even be conscious of it (1935, VI.156) and their trajectories might only involve physical possibilities rather than physical certainties. Survival is a journey. Organisations, unlike people, are not pre-destined to mortality: they could feasibly exist as long as society itself exists. Perhaps ‘*Sustainability*’ is a better synonym for the concept of the ‘teleological trajectory’ of an organisation towards its *final* purpose.

Pfeffer and Salancik (2003) equate organisational survival with environmental dependence coupled with the ability to acquire and maintain resources. An effective organisation ‘satisfies the demands of those in its environment from whom it requires support for its continued existence’ (Pfeffer *et al.*, p 242). Of course there are numerous viable organisational competencies that contribute to satisfying these demands: legitimacy (Dowling & Pfeffer, 1975; Singh *et al.*, 1986), ability to learn (Levinthal, 1992) and innovation (Han *et al.*, 1998; Baumol, 2002; Cefis & Marsili, 2005) being examples.

Howsoever these competencies are prioritised or deployed, they reduce to a contribution towards the continued satisfaction of demands of those from whom resources are acquired or by whom they are maintained so that the organisation continues to exist. Benchmarking has the objective of recognising exemplary competencies, understanding

them and gauging their merit so as to secure better teleological trajectories elsewhere. This suggests the following definition of states of affairs:

‘States of affairs represent the status of organisational competencies at some point in time that can be gauged according to some consistent metric and establish the teleological trajectory or sustainability of an organisation.’

A benchmarking process must identify organisational competencies, gauge their value or impact according to some consistent metric (cardinal, real, monetary, etc.) and also establish how these competencies contribute to the sustainability of the exemplar organisation. It also follows that an anomalar must do likewise – isolate internal competencies and gauge their value or impact on its own sustainability otherwise the objective of organisational improvement through benchmarking may fail due to mismatched causal relata.

A state of affairs represents both *efficient* and *final* causal components. *Efficient* causal components are those giving rise to observable organisational effects (such as better market share, lower defect rate, lower prices, better staff retention, etc.). Since these effects are observable, the task of identifying their *efficient* causal relata is one of good process: this is made easier if the exemplar and anomalar co-operate.

The *final* components might be difficult to establish - other than to the obvious degree of “survival”. Many different situations are plausible. Exemplars and anomalars might operate in a quiescent environment where their intentions and relationships with their marketplaces are openly transparent. In such cases it would be pragmatic to reduce benchmarking to the resolution of exemplary *efficient* causes. However current sustainability might also be caused by a conscious pursuit of developmental teleologies where current states of affairs reflect intermediate competencies. For example, an organisation might consciously pursue an intermediate technology not for the sake of its perfection and consequential market gain, but for the sake of establishing a pool of resources (organisational learning) that might *possibly* do novel things at a later time. Including such an exemplar in a benchmarking process may not deliver benefits to the anomalar as both *efficient* and *final* causes reflect a transient state of affairs. Even if the exemplar’s developmental trajectory was discerned, it might not be sustainable within the anomalar’s organisation.

Tools for establishing metrics that reflect the teleological trajectory of organisational states of affairs might conceivably involve the establishment of sets of scenarios that trace out possible futures, within which, corresponding *efficient* causal components are quantified according to consistent (economic) criteria. This will be pursued in Chapter 9 within the context of extending this research.

The tools for establishing metrics that reflect the *efficient* causal components of an organisation's state of affairs will now be examined.

5.2 Measuring Organisational Improvement.

A necessary element of successful benchmarking requires that anomalars gauge the status of their current states of affairs, the status of exemplary states of affairs elsewhere, and discern the difference. A methodology for quantifying the status of a state of affairs is of paramount importance as there is little purpose in embarking on purposeful change if organisational improvement either cannot be gauged, or if gauged lacks merit on account of inefficiency.

Quantifying the status of a state of affairs not only involves quantifying the effects of its *efficient* causal relata but also those of its teleological and chance relata. These additional causal relata may influence the manner in which a state of affairs came into being, where its future trajectory may lie and the contributions of any probabilistic dynamics.

In normal competitive-market environments the ability to continuously acquire and maintain resources is predicated on economic advantage arising from the effective behaviour of organisational states of affairs. In competitive-market environments, inefficient or wasteful behaviour will not remain effective as competitors making more efficient use of resources will prevail (Lancaster, 1966).

In non competitive-market environments, such as those often found within the public sector, other factors might influence sustainability. However in Chapter 2 the arguments of van Helden and Tillema (2005) were presented to advance benchmarking as a substitute for the absence of competitive-market forces. In these environments they argued that the inevitable pressures arising from inefficient organisational performance, whether real or perceived, can attract remedial action from higher authorities (e.g.

government). This may threaten the survival of errant organisations through dissolution or devolution into more efficient entities. Since public sector organisations are resourced from imposts or legislative monopolies on the activities of others, a representative democracy will inevitably require evidence of their economic efficiency should suspicion to the contrary arise.

The concept of economic efficiency is not intended to exclude phenomena that might have other than monetary or economic metrics. Whilst many organisational elements are measured in non-monetary terms (customer satisfaction, staff retention, workplace ambience, market-share, recognition, etc.), their impacts are ultimately economic. For example, remediation of adverse economic impacts arising from unsatisfactory customer satisfaction (reduced revenues) or staff retention (loss of intellectual capital and increased recruitment costs) might require the enhancement of soft resources (e.g. social systems and processes) rather than simply the acquisition or maintenance of physical resources. In such cases, the distinction between *efficient* and *final* cause is clear. Social or other non-physical change may be desirable for the sake of improved economic performance which serves the purpose of economic sufficiency to acquire the resources necessary for survival.

5.3 Efficiency of States of Affairs.

The extent to which outputs are produced by inputs reflects the efficiency of a process, system of processes or state of affairs. If a state of affairs is represented by *n* outputs, each denoted by *y*, and *m* inputs, each denoted by *x*, its efficiency, shown in Equation 5-1, is the quotient of accumulated outputs and inputs where both have common dimensions:

$$\text{Efficiency} = \frac{\sum_{r}^{n} y_r}{\sum_{i}^{m} x_i}$$

Equation 5-1. Efficiency

Efficiency may be decomposed into its principal elements: technical and allocative efficiency. Technical efficiencies (TE) arise from the ability to apply better technologies or methodologies so as to improve outputs or production relative to the resources or inputs applied (Farrell, 1957). Allocative efficiencies (AE) arise from the application of better resources or inputs so as to minimise their use in relation to outputs or production

(Farrell, 1957). It should be noted that Farrell's description of 'price efficiency' is what is now known as allocative efficiency. The product of technical and allocative efficiency is overall economic efficiency (OE) and can be illustrated by way of a common benchmarking scenario.

Suppose two different organisations, A and B employ a common technology, ***α,*** but use their own processes to produce an identical product. This technology requires the consumption of only two inputs, Resource 1 and Resource 2. The relationship between these is depicted in Figure 5-1. Perfect (100% efficient) production using technology ***α*** is represented by the curve P-P'. This curve, known as the efficiency frontier, represents the maximum output for all different proportions of Resource 1 and Resource 2 – i.e. an isoquant.

If Organisation A produces perfect production and uses the proportion of resources consistent with position ***a*** and organisation B produces imperfect output in the same proportion as A but uses resources consistent with position ***b,*** then B is less efficient than A.

If anomalar, B, wishes to benchmark its production processes against exemplar A, a further examination of Figure 5-1 identifies the relative technical and allocative efficiency differences between the two organisations.

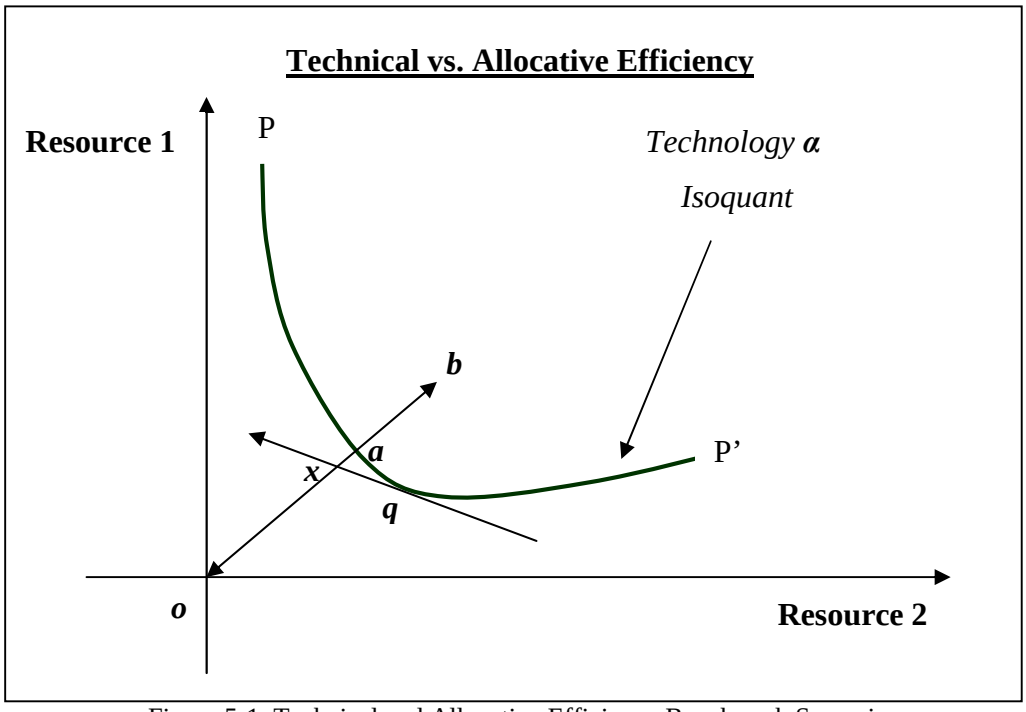

Figure 5-1. Technical and Allocative Efficiency Benchmark Scenario.

(After Farrell, 1957)

Organisations lying on the efficient frontier, P-P' exhibit technical efficiency. Of those, the organisation whose inputs reflect the minimum point on the efficient frontier is also allocatively efficient. It is necessary to establish the values (prices) of each input to establish the exemplar that is both technically (on the efficient frontier) and allocatively (using minimum resources) efficient.

The geometry of Figure 5-1 provides determination of these efficiencies. Table 5-1 shows how technical and allocative efficiency and their product, overall efficiency, are obtained.

	Technical Efficiency		Allocative Efficiency		Overall Efficiency (Technical*Allocative)	
Organisation A	$\frac{oa}{oa}$	100%	$\frac{ox}{oa}$	<100%	$\frac{ox}{oa}$	< 100%
Organisation B	$\frac{oa}{ob}$	< 100%	$\frac{ox}{oa}$	<100%	$\frac{ox}{ob}$	< 100%

Table 5-1. Technical and Allocative Efficiency Measures in a Benchmark Scenario.

This analysis also identifies two important benchmarking issues.

1. Organisation A, although an exemplar that exhibits 100% technical efficiency, may not exhibit allocative efficiency as it could have obtained equal output using fewer resources consistent with operating at point ***q***, the minimum-resource or optimum point of production.

2. Analytical determination of the production function may not be possible in situations where production processes are complex (Farrell, 1957). Referring to Figure 5-2, if P-P' is an estimate of the production function or perhaps only an estimate of organisation A's production function, then only the relative technical efficiency between itself and organisation B (or the other anomalars G and E) is obtained. In such cases, exemplars quantify relative rather than absolute improvements in resource usage. This suggests that the search for exemplars should be as broad as possible so as to provide anomalars (**G**, **B**, **E**) with the greatest likelihood of estimating the locus of the production function from the performance

of multiple exemplars (***C, D, A, F***) exhibiting 100% relative technical efficiency within the population.

A non-parametric analysis method such as Data Envelopment Analysis (Chapter 2.4.2) can be used to establish a piece-wise linear isoquant of a population of anomalars and exemplars via analysis of a set of their environmental variables (i.e. resources: inputs & outputs) - the segments C-D-A-F in Figure 5-2. Such analysis can also establish the overall efficiency (the product of technical and allocative efficiency) provided the values (prices) as well as the quantities of resources are known.

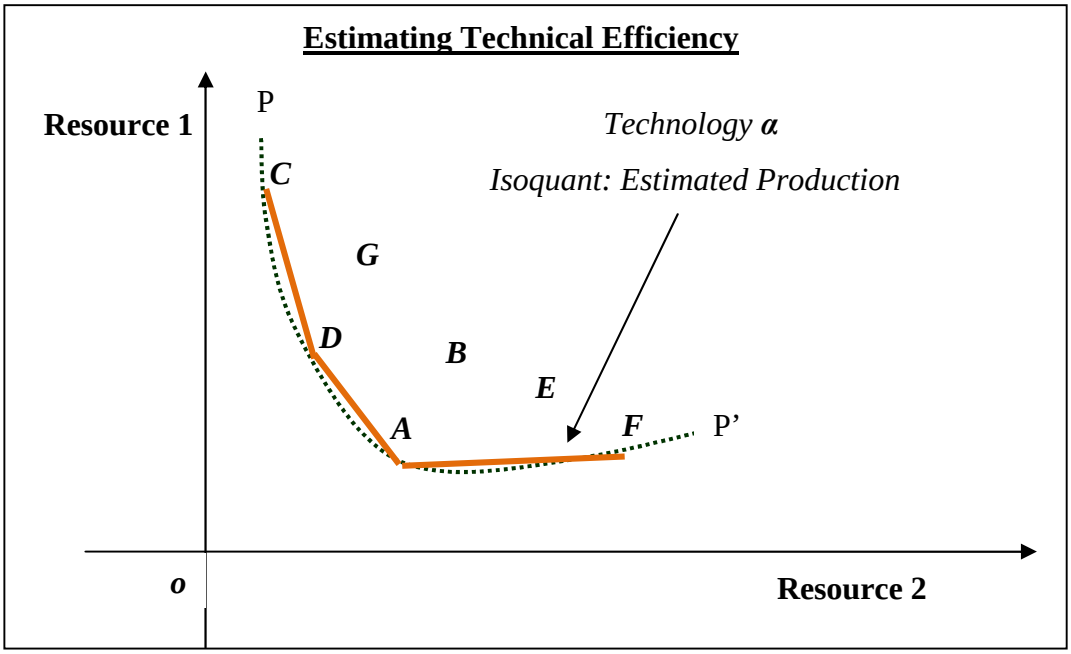

Figure 5-2. Piece-wise linear estimation of production functions from an isoquant.

DEA also requires other assumptions to be made. Resource behaviour under an assumed returns-to-scale condition is an important consideration as some resources may be constrained in the manner of their use – due to non-linear effects such as a capacity constraint, or perhaps an inability to be altered in any practical manner. For example, organisations may have substantial resources such as built infrastructure that cannot simply be scaled up or down in any practical manner or use utilities such as water, gas or electricity, the supply of which is capacity-limited by extra-organisational factors.

There are also other techniques for establishing an efficient frontier, the choice of which is determined by available information and computational ease, however they will not be considered at this stage.

In summary, a production function is a means of establishing economic metrics to quantify the overall status of a state of affairs as a function of its resources. Where a production function is unknown, tools such as DEA may be used to estimate one via the formation of an efficient frontier. This frontier represents the locus of exemplary technical efficiency relative to the population of organisations or states of affairs featuring in the benchmarking process. An algorithm for establishing a piece-wise frontier such as shown in Figure 5-2 is derived in Appendix 1.

5.4 Economic Welfare of States of Affairs.

Whilst it might suffice to simply benchmark one state of affairs against another, it may also be necessary to consider the overall effect of many states of affairs and further reflect their economic behaviour with respect to both time and chance – the other two causal factors in a benchmarking process governed by Peircean causation. This requires that the *teleological trajectory* or sustainability of an organisation be represented in economic terms so that benchmarking can distinguish effective from ineffective effort.

A mathematical statement of economic welfare may be derived from the following organisational factors (Stavins *et al.*, 2003; Arrow *et al.*, 2004):

- The timeframe over which resources are provided – denoted by the variable t,
- The organisation's aggregate consumption of resources over time – denoted as C(t), this is also synonymous with the concept of a production function.
- Utility or satisfaction over time provided by the consumption of these resources by those whose demand for which is necessary for organisational survival – denoted as an idealised utility function U(C(t)),
- An initial point in time denoted by the symbol τ,
- Reflection of the worth of present versus future utility denoted by a factor k giving a rate of decline in value from the initial point in time τ to a general point in time t – represented as an exponential $e^{-k(t-\tau)}$.

If the trajectory or sequence of states of affairs is known over a period of time, then the actual utility at some future time $\mathbf{t_f}$ shown in Equation 5.2 is the product of utility and its discount factor as illustrated in Figure 5-3,

$$U(C(t)).e^{-k(t_f-\tau)}$$

Equation 5-2. Utility of a State of Affairs

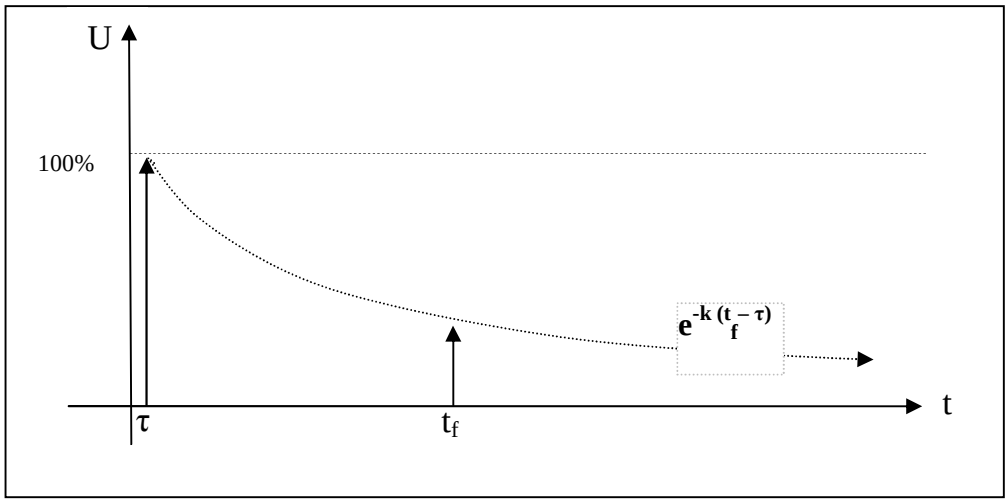

Figure 5-3. Utility Value of Aggregate Consumption over time.

Thus the total welfare, **W**, of the enterprise (or any 'state of affairs') under consideration is the accumulated discounted utility, which should also be positive to be sustainable:

$$W(t) = \int_{\tau}^{\infty} U(C(t)).e^{-k(t-\tau)}dt \geq 0$$

Equation 5-3. Welfare of a State of Affairs (Continuous)

In practical terms, welfare may be quantified from estimates or forecasts of future utility streams – i.e. Discounted economic surpluses, **S**, generated by the enterprise or state of affairs under consideration.

For trading enterprises economic surpluses are the periodic contribution to equity after all other consumption (e.g. labour, materials, services, capital and taxation expenditures) have been deducted from income. Enterprises generating negative economic surpluses fail to satisfy the demands of their resource providers (investors, staff, customers, suppliers and the community) and ultimately cease trading.

The discount factor in such cases is the prevailing risk adjusted market rate, **r**, for the opportunity cost of money. The Welfare, **W**, generated (or projected) by economic surpluses, **S**, over discrete (say, annual) periods of time, commencing at period **τ** and finishing in period **f** with an opportunity cost of money **r**, is given by the discrete summation:

$$W(t_f) = \sum_{j=\tau}^{f} S_j.(1+r)^{-(j-\tau)} \geq 0$$

Equation 5-4. Welfare of a State of Affairs (Discrete)

This is also the basis of enterprise economic value added; EVA® (Stewart, 1991) where an annual economic surplus reflects short-term performance and the present value of forecasted economic surpluses reflects its market value. Economic surplus was also used in the establishment of financial yield benchmarks for the NZ Tourism Sector to establish market performance criteria (J P Moriarty, 2007).

Intergenerational equity requires welfare to be both positive and non-decreasing. Periodic reductions in utility impose compensatory burdens in future time periods to preserve the present value of resource consumption. Over-consumption of resources at some period might also preclude the attainment of utility in future periods if there was 'an expectation of continuous resource availability' (Stavins *et al.*, 2003). If enterprises or their states of affairs are to be sustainable, welfare must be positive and monotonic increasing – i.e.

$$\frac{dW(t)}{dt} \geq 0 \text{ or } (\forall j \in [\tau, f], S_{j+1} - S_j \geq 0)$$

Equation 5-5. Sustainability Conditions

Welfare criteria for the sustainability of an enterprise may be specified but they do not carry many guarantees as to their attainment! Arrow *et al* (2004, p 150) also commented on a number of associated considerations that are important to organisations seeking to establish sustainable trajectories (business programmes):

- There is neither a unique nor optimal consumption path associated with conditions for sustainability. More specifically, inter-period consumption trajectories might not even be efficient,

- Because future consumption can only be forecast, there is no guarantee that either the conditions for sustainability will be met in future periods or that the utility in future periods will be as high as in prior periods,
- If exhaustible resources are essential to production and consumption then it is conceivable that there may be no sustainable trajectory for the organisation.

These caveats are important to benchmarking practices as they further emphasise the challenges associated with a snapshot of an exemplar's state of affairs in the absence of knowledge of its teleological trajectory – not all of which might be known or knowable. In comparison, the *efficient* causes that establish the exemplar's state of affairs are obtainable with fewer challenges, and should be seen as the initial objective of a benchmarking programme: - obtaining increased welfare. In conjunction with establishing and implementing strategies based on *efficient* cause, the anomalar will need to reflect on the nature of consequential developments based on its own teleological trajectory. These developments should also meet economic sustainability criteria to the greatest extent possible for as many future periods as remain within its influence.

In summary, the sustainability or teleological trajectory of an organisation or its states of affairs can be gauged provided the behaviour of associated resources can be monetised or quantified in some consistent manner. The accumulated utility or aggregate consumption of the organisation or state of affairs under consideration is a measure of its welfare and it is a condition of its sustainability that this be positive. If welfare is not only positive, but monotonic, there is an equitable relationship between intergenerational consumption. Future utility can be discounted to the present in order to quantify the opportunity gained or forgone by adopting a particular consumption trajectory. However, the conditions for economic sustainability cannot be assumed to imply optimality at any point on the trajectory. The implications for benchmarking follow:

- Sustainability, the teleological trajectory of an organisation, reflects both *efficient* and *final* cause, whereas efficiency only reflects *efficient* cause.
- Economic welfare can assist with the development of an anomalar's teleological trajectory by quantifying and prioritizing options that enable ongoing acquisition and maintenance of resources.

- Sustainable economic welfare implies the achievability rather than the optimality of ongoing resource consumption and maintenance - the *efficient* causal relata of states of affairs,

- Neither the feasibility nor optimality of a present state of affairs is a reflection of feasibility or optimality of a future state of affairs unless the disposition of all *efficient* causal relata (e.g. resources, inputs, outputs) constituting a future state of affairs are known *a priori*.

- The welfare disparity between exemplar and anomalar states of affairs is a gauge of the immediate rather than future opportunity that might be obtained through alignment of their *efficient* causal relata.

Welfare economics provides a mechanism for gauging the sustainability of organisational states of affairs, but there remains the question of the degree to which alignment is possible between the exemplar and anomalar. If states of affairs are identical in anomalar and exemplar organisations, alignment of one state of affairs upon another should be straightforward, but what if it is not?

5.5 Dependencies between of States of Affairs.

Benchmarking requires the identification of causal relata evincing exemplary behaviour to pave the way for improving the anomalar's state of affairs. An exemplary state of affairs is characterised by a set of relata behaving in such a manner as to obtain a set of observable effects. In a Humean-Mill *efficient* causal framework, these relata are unconditional antecedents of observed effects. In a Peircean causal framework, *efficient* cause is associated with both chance and *final* cause – denoted by a series of developing teleologies over time.

Causal relata are also characteristics of their environment having properties that vary in some fashion such that their interaction gives rise to these effects. If a state of affairs produces certain effects, we may say that its environmental variables *behaving in a certain fashion* at certain times *entail* these effects. If the environmental variables are known and their behaviours and circumstances are obtainable, then their effects are obtainable. This is the essence of benchmarking. Appendix 2 expands on the nature of environmental variables using the concepts of entailment (to describe their properties) and

supervenience (to describe their relationships). Terminology and findings from Appendix 2 will now be used.

Exemplary effects might also be caused by a variety of causal relata, since, in the absence of wholly sufficient conditions, there might be ambiguous relata that are associated with the effect (See Mill and Mackie, Chapter 3.5.8). Consider the following example.

> Suppose a profitable tourism organisation A observes another tourism organisation B to enjoy exemplary after-tax profitability, i.e. better performance. If these organisations are alike in all respects offering identical services and prices whilst enjoying identical patronage. However, A operates on one side of the street and B the other where the street is the boundary between two different jurisdictions. The set of environmental variables describing performance is {r, e, t} (the set of revenues, expenses and taxes) and similarly for B. Under prevailing taxation law, the unconditional antecedents of after-tax profit are these environmental variables behaving in a fixed manner. That is to say,
>
> $$\{r, e, t\} \Leftarrow \text{(after-tax profit)}.$$
>
> If Organisations A and B have identical revenues and expenses, or if the difference between their revenues and expenses are identical, then their taxes must differ since after-tax profitability logically entails these three environmental variables. On close inspection, Organisation A's taxes, t, are $\{t_l, t_n\}$ where t_l is a local tax and t_n is a national tax. Organisation B, on the other side of the street, is in a different tax jurisdiction and only pays t_n, hence its higher profitability. Thus
>
> $$\{EV_A\} = \{r,e,t_l,t_n\} \text{ and } \{EV_B\} = \{r,e,t_n\}.$$

This example raises two questions:

1. What can be said about their environmental variables?
2. To what extent can B be an exemplar to A?

We cannot say that A logically entails B, since A is a superset of B and all instances that make A true do not make any instances of B true (note that both anomalar and exemplar are profitable – i.e. both t_l and t_n exceed zero). Respective sets environmental variables are not equal although some individual variables are. Both organisations share similar dispositions and if these are examined they are probably the same – leaving open the

possibility of dispositional entailment. We may say that B weakly supervenes on A (as any variance in any $\{EV_B\}$ would result in a corresponding variance in $\{EV_A\}$, but the reverse is not necessarily true).

If this example is compared with the entailment and supervenience relationships presented in Appendix 2, Table A2-1 (replicated below), the anomalar is located in the horizontal category of 'potential teleological improvements'. Since there is neither logical nor probabilistic entailment but perhaps dispositional entailment, the scope of the benchmarking improvement opportunity is teleological.

Improvement Opportunity	**Humean-Mill/Peircean Model *Efficient***	**Peircean Model *Efficient***	**Peircean Model *Teleological***
Perfect Causal Improvements	$\{EV_A\} = \{EV_B\}$ $A \models B,\ B \blacktriangleright_{s,g} A$	$\{EV_A\} = \{EV_B\}$ $A \vdash B,\ B \blacktriangleright_{s,g} A$	$\{EV_A\} = \{EV_B\}$ $A \Vdash B,\ B \blacktriangleright_{s,g} A$
Potential Causal Improvements	$\{EV_A\} \subset \{EV_B\}$ $A \models B,\ B \blacktriangleright_{w} A$	$\{EV_A\} \subset \{EV_B\}$ $A \vdash B,\ B \blacktriangleright_{w} A$	$\{EV_A\} \subset \{EV_B\}$ $A \Vdash B,\ B \blacktriangleright_{w} A$
Potential Teleological Improvements	-		$\{EV_A\} \neq \{EV_B\}$ $A \Vdash B, B \blacktriangleright_{swg} A$

Table 5-2. Supervenience and Entailment Relationships in Benchmarking.

Supervenience, whether strong, weak or global, in the absence of logical or probabilistic entailment but evident dispositional entailment suggests that the anomalar's improvement strategy is to alter its teleological trajectory by simply moving across the street! Because there is no logical entailment, imitation is precluded. Of course if the local tax were avoidable there would also be logical entailment and the anomalar could fully imitate the exemplar. Alternatively, the anomalar might observe that since local tax is unavoidable the relevant environmental variables for the benchmarking exercise are those apart from it. If this were done, there would be both logical entailment and strong supervenience and the anomalar could avail itself of full causal improvement.

This example identifies two pragmatic options.

- Reducing the scope of a benchmarking exercise to include subsets of environmental variables might identify improvement opportunities that would otherwise be infeasible.

- Improvement need not be of an *efficient* causal nature – it might be teleological – driven by shared dispositions.

Provided the exemplar supervenes on the anomalar there may be scope to establish entailment within subsets of their respective environmental variables. Entailment at subset levels might identify localised improvement opportunities for the anomalar; however they may still be subject to broader causal constraints. E.g. if plant or structures needed to be expanded but were physically constrained from doing so in the anomalar's present circumstances, the degree to which such expansion could occur in the absence of a new teleological trajectory (i.e. relocation or reconstruction) would be a reasonable outcome of a benchmarking exercise.

5.6 Benchmarking States of Affairs.

States of affairs existing in the anomalar and exemplar organisations must be able to be observed and behaviours established in terms of sets of environmental variables. Within each state of affairs, causal relationships must exist between these variables to obtain their characteristic effects. Unless the environmental variables describing exemplary states of affairs are also supervenient upon the anomalar, other factors determine the effects of respective states of affairs. Supervenience together with some form of entailment establishes a reliable relationship between the exemplar and anomalar states of affairs and indicates that benchmarking efforts have the potential to be effective. Practitioners might benefit from observing exemplary state of affairs that supervene but dispositionally entail the anomalar if they stimulate reconsideration of current developmental trajectories or *final* cause. Improvements through *efficient* causal mechanisms necessarily require logical or probabilistic entailment and supervenience.

A final question is the representation of supervenience and entailment within a benchmarking programme. How can the process of benchmarking include these features and ascertain the degree to which improvement can be achieved?

5.7 Transitioning Between States of Affairs.

In the previous section the necessity of supervenience and logical entailment was shown to be a requirement for *equating* an anomalar's state of affairs to that of the exemplar. However, this might not always be possible, as there might be some factors that differ or

are absent between the two. The degree to which some improvement might be made can be represented by mapping the relationships between respective environmental variables. The following approach builds the pathways for achieving this.

5.7.1 A Transitional Mechanism between Benchmarked States of Affairs.

Suppose we have a simple situation where two states of affairs are identical, but their *effects* (performance) differ – i.e. identical environmental variable properties, but have different scalar values. Suppose the output of a general production process is reflected by three environmental variables,

$$EV = \{x, y, z\},$$

and that this production process has a utility function, exemplified by a simple linear expression:

$$U(x, y, z) = x + 2y + 3z.$$

Let EV_A be the set of 3 environmental variables associated with anomalar Organisation A such that $\{EV_A\} = \{a_1, a_2, a_3\}$. Let these have the values [2, 5, 4] to establish its state of affairs, S_A.

Let EV_B be the set of 3 environmental variables associated with exemplar Organisation A such that $\{EV_B\} = \{b_1, b_2, b_3\}$. Let these have the values [5, 5, 5] to establish its state of affairs, S_B.

The Utility of the anomalar's state of affairs, S_A is $a_1 + 2.a_2 + 3.a_3$, and that of the exemplar's state of affairs, S_B is $b_1 + 2.b_2 + 3.b_3$, then in Matrix form, the utility function, U, achieving this is the column matrix $[1, 2, 3]^T$. Thus the welfare (W) of each state of affairs is:

$$W_A = S_A .U = [a_1, a_2, a_3] [1, 2, 3]^T = a_1 + 2.a_2 + 3.a_3 = 24 \text{ units.}$$

And,

$$W_B = S_B. U = [b_1, b_2, b_3] [1, 2, 3]^T = b_1 + 2.b_2 + 3.b_3 = 30 \text{ units.}$$

Now since the anomalar logically entails the exemplar and the exemplar strongly supervenes on the anomalar, it is feasible for an exemplary value of, say, b_1 to transition to a_1 and deliver its effect to the anomalar. In this case, the same applies to the other two

environmental variable pairs, (a_2, b_2) and (a_3, b_3). In this case, there is no possibility of the value of b_2 being applied to a_1, or indeed to any other EV other than a_2, as they possess entirely different properties.

We may represent this by a Feasibility Matrix, F, identifying the feasible relationships between respective EV,

$$F = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

This diagonal matrix allows only mapping between EV with identical properties. If mapping were permitted between EV with different properties (as might be the case where substitutes are permitted – an example of which might be full time staff versus part-time staff contributing to a state of affairs), a non-zero value would be recorded. For example, suppose it was discovered in the benchmarking process that what is represented as a_3 by the anomalar is represented by the exemplar as b_3- ½b_4. Now as there is no such environmental variable as a_4 in the anomalar to correspond with b_4, this is recognised via the introduction of a null variable in the State of Affairs;

$$\text{i.e. } S_A = [a_1, a_2, a_3, 0] \text{ and } S_B = [b_1, b_2, b_3, b_4].$$

In such cases supervenience and entailment conditions between states of affairs still hold as changes in one state result in changes in the other and $S_A \subseteq S_B$. In this example the Feasibility Matrix would then be:

$$F = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -.5 & 0 \end{bmatrix}$$

So if $S_A \subseteq S_B$, linear combinations of multiple environmental variables can also be described by the Feasibility Matrix to accommodate evident relationships between the anomalar and the exemplar environmental variables sharing exactly the same properties. This matrix also expresses Aristotle's *formal* cause – the preservation of properties in a causal relationship.

In this example, the improvement required by the anomalar in order to equal the performance of the exemplar is simply given by the difference between the states of their feasible EV. If λ generally represents the difference, S_B - S_A, between feasible Welfare states – i.e. $\lambda_{1,\,1} = (b_1 - a_1)$, then the transition between all anomalar and exemplar Welfare states is given by the matrix:

$$\lambda.F = \left[\lambda_{1,1}, \lambda_{1,2}, \lambda_{1,3}\right]\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \left[(b_1 - a_1),(b_2 - a_2),(b_3 - a_3)\right],$$

So the improvement strategy for the anomalar is [3, 0, 1] for its respective EV.

Whilst this improvement strategy might have been obtained upon inspection in this particular example, the approach paves the way for generalisation and also for exploring situations where {EV} might not be identical, but might be substitutable or related in some fashion, or, indeed present according to some statistical relationship. In these cases the Feasibility Matrix might be populated with scalars or functions describing formal relationships between all possible combinations of EV sharing the same properties.

Alternatively, If the anomalar is "certain" that the exemplar's performance lies between particular bounds – not more than 'z', suspected to be 'y', but not less than 'x', the Feasibility Matrix might be populated with the appropriate (triangular in this case) probability functions applicable to *each* EV and a Monte Carlo method of estimation used to establish the 'most likely' improvement strategy.

5.8 Summary.

This chapter has addressed the nature of benchmarking by examining and defining relationships between the exemplar and the anomalar. The most pressing cause of any organisation is its survival – the ability to continuously satisfy those upon whom reliance is made for the supply and maintenance of resources. Organisations pursue this cause through the establishment of various states of affairs. These states of affairs consist of environmental variables within causal relationships that establish competencies whose objective is the satisfaction of organisational resource providers. Causal relationships between environmental variables include both *efficient* and *final* causation components. *Efficient* causation components reflect the observable and measurable relationships

between environmental variables whereas *final* causation components reflect the organisation's teleological trajectory or its strategic intent.

Benchmarking seeks to understand and benefit from the process of organisational survival by distilling exemplary behaviours into forms that can be used for improvement elsewhere.

Benchmarking based on *final* causation may prove difficult as an organisation's teleological trajectory might only reflect broad possibilities or transience en route to more desirable states. Benchmarking based on *efficient* causation was linked to efficiency improvement for the purpose of generating continuously increasing welfare streams that enables ongoing supply and maintenance of organisational resources. However, the degree to which efficiency improvement is possible not only depends on the depth of knowledge of exemplary production processes, but also the breadth of the benchmarking sample since improvements can only be relative to that sample.

If the objective of benchmarking is an optimal improvement in anomalar performance it is a necessary condition that the sampled exemplary production processes (technical efficiency) and their resources costs be known (allocative efficiency). If resource costs are unknown, improved relative technical efficiency is the best possible outcome of a benchmarking process.

This leads to the conclusion that optimality (the very 'best') is an unrealistic benchmarking objective as such an exercise would require knowledge of all possible exemplars.

The chapter also demonstrates mechanisms (see Appendix 1) for determining exemplary production functions where they are unknown or unavailable so that an anomalar can establish its own efficiency improvement strategy from a benchmarking exercise.

The objective of benchmarking is improved efficiency and enhancement of future welfare streams – the reflection of greater disposable wealth within the anomalar's sphere of interest. Increased welfare over time conveys no notion optimum efficiency – only the notion of inter-temporal improvement! Teleological trajectories only entail purpose, whereas *efficient* causal processes entail achievement and neither entails optimality,

ongoing feasibility or even ongoing technical or allocative efficiency. Benchmarking can be used to establish and project exemplary *efficient* causal processes into future timeframes by forecasting ongoing welfare streams within the teleological trajectories believed to prevail at such timeframes. Organisations risk failure if their welfare streams do not monotonically increase over time.

Causation theory also introduced two indicative concepts that assist with the distinction between successful and unsuccessful benchmarking effort. Appendix 2 examined the applicability of entailment and supervenience to the exemplar-anomalar relationship. Separation of objective benchmarking outcomes into Humean-Mill (strictly *efficient*) and Peircean (probabilistic *efficient* and *final*) causal components based on anomalar entailment of exemplar environmental variables and nomological supervenience of the exemplar upon the anomalar established a set of useful benchmarking expectations.

Ineffective benchmarking efforts arise from a lack of both entailment and supervenience whereas effective efforts – offering varying degrees of improvement – arise from combinations of entailment and supervenience.

Finally, the degree to which improvement is available from a benchmarking process is gauged by applying the concepts of efficiency, welfare, entailment and supervenience within a causal framework. A feasibility matrix maps the transition of appropriate improvement policies from an exemplar to an anomalar to exemplify the application of causation and economic theory in an effective benchmarking process.

The next Chapter will identify the necessary and sufficient conditions for effective benchmarking practices by combining the conclusions from this and previous Chapters into a statement of a theory of benchmarking.

Chapter 6. Articulating a Theory of Benchmarking.

6 Introduction to the theoretical framework.

Previous chapters have examined the organisational framework in which benchmarking is found, the nature of causal relationships, organisational purpose, methods for establishing metrics describing the status of organisational states of affairs and conditions necessary for constructing theory.

This chapter addresses the task of establishing benchmarking within a theoretical framework that is empirically robust, logical in its treatment of circumstances and entailed within commonly accepted organisational paradigms.

The theory filter or validator developed in Chapter 4, if updated to reflect the current task, does not constitute sufficiency for theory construction, but it does provide a triangulated set of conditions that necessarily avoid *a-theoretical* constructions.

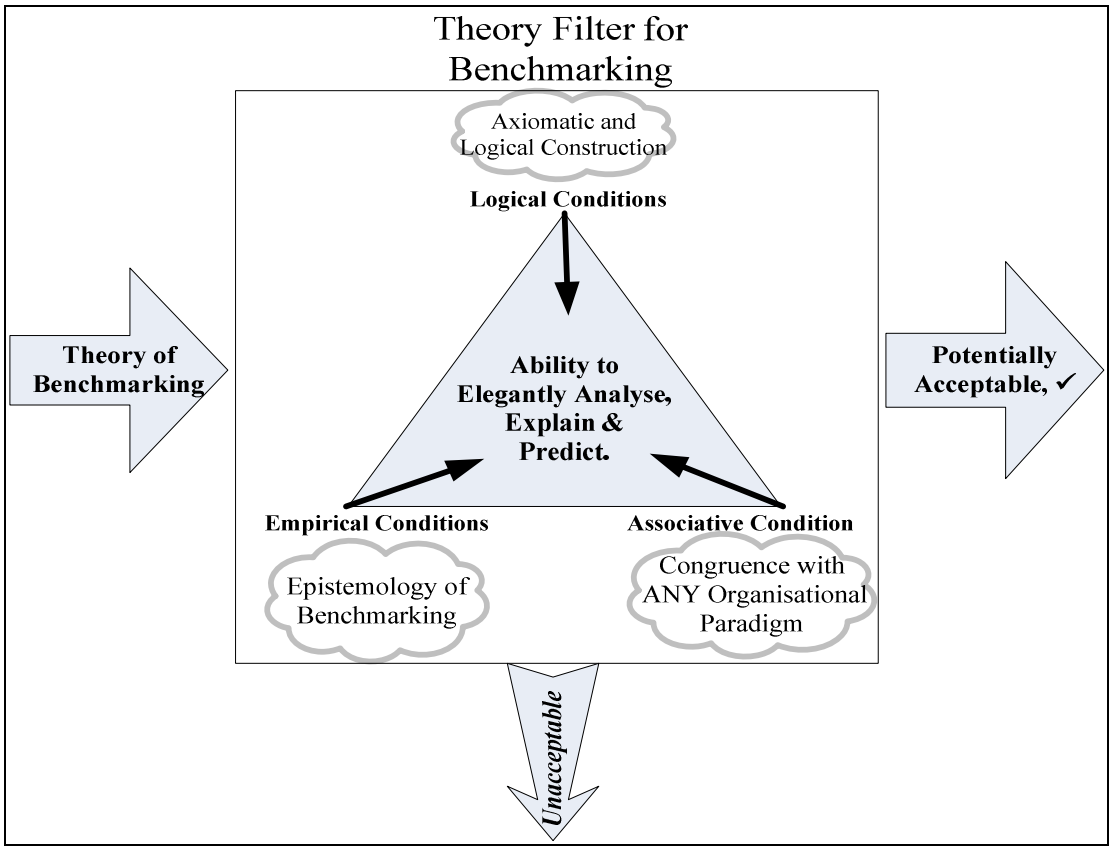

Figure 6-1. A Filter (or validator) for Benchmarking Theory.

Developed from Figure 4-4.

The epistemology of benchmarking is substantial. Chapter 2 identified general agreement that benchmarking is an important contributor to the imperative of organisational improvement – a necessary condition for survival in a competitive environment. However, as benchmarking is neither necessary nor sufficient for organisational improvement or survival, its role is that of a tool or mechanism that may add value to this purpose. A principal impediment to its efficacy has been its empiricism: practitioner efforts have not been grounded in either causal or paradigmatic theories but rather in an epistemology of procedures, models, case studies and good practices.

Likewise, current organisational paradigms do not deny the efficacy of 'improvement'. In Chapter 4, benchmarking was examined in the light of pan-organisational paradigms where the conclusion reached was that an organisational theory demonstrating 'sensemaking' within multiple organisational paradigms was stronger than one that did not. The locus of benchmarking is a stable meta-platform equally shared by the four incommensurable organisational paradigms of Burrell and Morgan (1979) and Gioia and Pitre (1990) as earlier illustrated (Chapter 4.3.1).

It was argued that Benchmarking not only exhibited 'sensemaking' within these four organisational paradigms but did so in any rational paradigm that admits 'improvement' as a survival-enhancing mechanism.

The remaining component in the 'theory filter' encompasses, perhaps, the strongest requirement for any nascent theory: general logical conditions that are based on the principle of causality. The remainder of this chapter will state and address the logical conditions for benchmarking in terms of two criteria: those of an effective benchmarking process as well as those of an effective benchmarking improvement.

6.1 Approach.

A definition of benchmarking was proposed in Chapter 2:

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

This definition addresses both the nature and purpose of benchmarking, but not how it is achieved. The approach of founding both the nature and purpose of benchmarking within established theories will be used to postulate a theory of benchmarking.

A primal axiom of organisational purpose, theories of causation and economic welfare together with the concepts of entailment and supervenience provide an established theoretical basis for describing the essence of benchmarking and *prior* determination of the potential for practitioner efforts to achieve their purpose.

The primal axiom relied upon for the foundation of a theory of benchmarking is that 'survival' is a *sine qua non* of the ontology of an organisation.

Benchmarking also requires the existence of two states: the first, an exemplar, evinces a desirable state of affairs and the second, an anomalar, evinces an existing state of affairs. Arising from this are a further series of requirements that identify necessary conditions for the establishment of a causal relationship between an anomalar and another desirable and evidently exemplary state of affairs, evidence of which exists. Desirability is a teleological phenomenon.

The concept of an exemplary state of affairs establishes a necessary condition of measurement. One state of affairs cannot be held superior to another unless gauged to be so by some relevant metric. The concept of 'welfare' is the gauge used to establish the status of a state of affairs. In most organisational circumstances practitioners will use some form of economic welfare metric as it reflects the satisfaction accorded by those upon whom competitive organisations depend for the supply and maintenance of resources necessary for survival.

Achieving the objective of intentional change for the sake of attaining a superior state of affairs (improvement) also requires knowledge of causal relationships. Establishing a causal relationship between two states of affairs necessarily requires determination of antecedents giving rise to effects evinced by the status of both an existing and superior state of affairs. Causal theories shown in previous chapters include the Aristotelian forms of *formal*, *material*, *efficient* and *final* causation; Humean-Mill *efficient* causation and Peircean *efficient* and *final* causation.

A Peircean causal framework holds that there can be no *efficient* cause without *final* cause and is a pragmatic reflection of organisational behaviour. Humean causal frameworks recognise only *efficient* cause and Mill's extension reflects the invariant behaviour of mechanistic processes within organisations.

For benchmarking to be a teleological process, it is necessary to demonstrate that its contribution to a superior state of affairs is consistent with organisational purpose and that such a purpose is consistent with organisational ontology. Peirce's framework embodying *efficient*, chance and *final* causal components provides explanation for the basis of states of affairs and constitutes the causal engine that is central to benchmarking: transferring power (Locke, Chapter 3.5.4), or welfare, from the exemplar's causal engine to the anomalar.

Here a distinction must be made between a benchmarking process and any consequential benchmarking improvement. A benchmarking process is effective if it identifies the potential for at least one improvement opportunity to arise. A benchmarking improvement necessarily increases the welfare of the anomalar's state of affairs as a result of an effective benchmarking process. Practitioners can only achieve benchmarking improvement via effective benchmarking processes.

There is a general caveat. Cognisance of, or exposure to, exemplary organisational practices of any kind may stimulate an observer to deduce or perceive possibilities other than those supported by evident nomological relationships. Whilst it is rational to acknowledge this phenomenon, it is not a rational to expect it to be *a priori* effective or an attribute of a process desiring intentional improvement. Stimulation of this sort is 'serendipity' not benchmarking. As with Aristotelian *formal* cause, the form (laws) of this theory states what benchmarking may achieve as well as stating what it may not.

6.2 The Theory of Benchmarking.

This Theory of Benchmarking rests on one Axiom and five logical conditions, Benchmarking Theory zero through five (BT0:BT5), that address the causal nature of relationships between exemplar and anomalar states of affairs to obtain effective an

effective benchmarking process and an effective benchmarking improvement. These conditions are stated in Table 6-1, based on reduction of table A2-1.

The Theory of Benchmarking	
BT0 Primal Axiom.	'To survive' is a *sine qua non* of organisational ontology.
BT1 Causal Engine.	Effective benchmarking processes necessarily entail Peircean Causation.
BT2 Effective Improvement.	Any effective benchmarking improvement necessarily requires an increase in anomalar welfare via the transformation of exemplary relata into feasible anomalar relata
(BT3 $\vee$ BT4 $\vee$ BT5) Effective Process.	An exemplary state of affairs is necessarily supervenient upon an anomalous state of affairs and an anomalous state of affairs necessarily entails an exemplary state of affairs.

Table 6-1. The Theory of Benchmarking.

The third, fourth and fifth logical conditions are represented as a single conjunction since a benchmarking process may be perfectly, potentially or teleologically effective. Each of these conditions will now be defended, warranted and qualified.

6.2.1 Primal Axiom: BT0.

BT0: 'To survive' is a *sine qua non* of organisational ontology.

Warrant: Actions that diminish welfare reduce dependent satisfactions, diminish the efficacy of resources and threaten survival – i.e. contravene the Primal Axiom.

Defence: 'To survive', an organisation evinces 'continued ability (increasing welfare) to satisfy that or those from which or whom there is dependence for the supply and maintenance of its resources.'

Qualifiers: There are none.

6.2.2 Benchmarking's Causal Engine: BT1.

BT1. Effective benchmarking processes necessarily entail Peircean Causation.

Warrants: The purpose of effective benchmarking process is to obtain an improvement in an anomalar's state of affairs – a *final* cause. The anomalar's purpose for

improvement is to increase its ability to survive: – axiom, *final* cause. An effective benchmarking process evinces the existence of antecedents that instantiate improvement – *efficient* cause. Effective benchmarking processes necessarily entail *efficient* and *final* cause. The conjunction of e*fficient* cause and *final* cause, subject to (Peircean) chance, necessarily entail Peircean Causation. Effective benchmarking processes necessarily entail Peircean causation.

Defence: Epistemology of benchmarking offers no purpose to the practice other than as a mechanism for organisational improvement – a *final* cause. A state of affairs that does not stand in need of improvement either dominates other observable and relevant states of affairs, or it does not. The determining process as to whether or not it dominates other observable and relevant states of affairs is the practice of benchmarking – *efficient* cause.

Qualifiers: Benchmarking is a purposeful process. Exemplar and anomalar states of affairs are only the causal relata of a purposeful process that entails determination of an improvement opportunity by comparing the causal relata of exemplary and anomalous states of affairs.

6.2.3 Effectiveness of the Benchmarking Improvement: BT2.

BT2. Any effective benchmarking improvement necessarily requires an increase in anomalar welfare via the transformation of exemplary relata into feasible anomalar relata.

Warrants: Substitution of exemplary environmental variables (relata) within an anomalar's state of affairs effects change, but not necessarily improvement. Such changes that, say, improve the magnitudes of any environmental variables but fail to increase anomalar welfare contradict the primal axiom and are ineffective.

Defence: Effective benchmarking reflects both change and welfare improvement. With the exception of a perfectly effective benchmarking process, conditions for effective benchmarking processes identify only potential opportunities as exemplary supervenience and entailment reflect a degree of encompassment rather than imitation of the exemplar's behaviour. Benchmarking becomes ineffective if the laws governing the anomalar's state of affairs do not confer or permit an increase in welfare when supplied with exemplary relata. Consistent with BT1, benchmarking improvements may apply to *efficient*

causal relata (of a tactical nature) or *final* causal relata (of a strategic nature) where their application preserves organisational consistency with the primal axiom.

Qualifiers: The condition is true, subject to the laws governing the anomalar's state of affairs remaining unaltered.

6.2.4 Effectiveness of the Benchmarking Process: BT3, BT4 and BT5.

There are three conditions that obtain an effective benchmarking process, depending on the type of entailment and whether the anomalar's environmental variables are an inclusive, proper or empty subset of the exemplar's environmental variables.

BT3. A perfectly effective benchmarking process necessarily requires an exemplary state of affairs to be strongly supervenient upon an anomalous state of affairs and necessarily requires the anomalous state of affairs to logically entail the exemplary state of affairs.

Warrants: Perfectly effective benchmarking entails identical nomological behaviour between the exemplar and anomalar states of affairs as well as the properties of their respective environmental variables. Strong supervenience entails mutual nomological behaviour over all mutually relevant states of their common environmental variables. An anomalar's state of affairs entailing an exemplar's state of affairs under nomological identity predicates variances in the state of any exemplary environmental variables necessarily predicates equal variances in the state of their respective anomalar environmental variables subject to Peircean chance. Adoption of the states of exemplar's environmental variables by the anomalar obtains a state of affairs equal to that of the exemplar. Perfect equality between exemplar and anomalar states of affairs is a sufficient condition for improvement.

Defence: Perfectly effective benchmarking is an expression of Humean-Mill-Peircean *efficient* causation; the unconditional antecedents of the effects of the exemplar are transferred to the anomalar.

Qualifiers: Nomological supervenience in perfectly effective benchmarking is subject to independence with respect to time and organisational teleology. Consequently, exemplary states of affairs obtained by singular causation, probabilistic causation or *final* causation do not entail Humean-Mill *efficient*

causation and are excluded from a statement that 'S_B is a perfectly effective exemplar for S_A'. Peircean chance is not excluded as it is elemental.

BT4. A potentially effective benchmarking process necessarily requires an exemplary state of affairs to be supervenient upon an anomalous state of affairs and requires a proper subset of an anomalous state of affairs to entail an exemplary state of affairs.

Warrants: Effective benchmarking entails nomological behaviour between the exemplar and anomalar states of affairs as well as the properties of their respective environmental variables. Weak supervenience entails mutual nomological behaviour over some relevant states of their common environmental variables. A subset of an anomalar's state of affairs entailing an exemplar state of affairs under weak supervenience predicates that variance in the state of at least one exemplary environmental variable necessarily predicates variance in the state of at least one anomalar environmental variable. Adoption of the state of at least one exemplary environmental variable by the anomalar might obtain an improved state of affairs.

Defence: Effective benchmarking is founded on Peircean causation; a less restrictive and more encompassing causal theory than Humean-Mill causation. Peircean causation admits teleological and chance causal phenomena in association with *efficient* causal relationships. Effective benchmarking arises from any supervenience of the exemplar on the anomalar that might obtain improvement based on either *efficient* or *final* cause.

Qualifiers: Whilst it is necessary for the exemplar to supervene upon anomalar states of affairs it is not sufficient condition for improvement. A subset of anomalar environmental variables entailing exemplar environmental variables is not a sufficient condition for improvement.

BT5. A potentially effective teleological benchmarking process necessarily requires an exemplar's state of affairs to be supervenient upon an anomalar's state of affairs.

Warrants: Supervenience in the absence of logical or probabilistic entailment of environmental variables necessarily entails common organisational dispositions. Solely dispositional relata cannot be *formal*, *material* or *efficient* causal relata they may only be *final* causal relata.

Defence: Dispositional entailment of environmental variables admits causal relata that only share their manifestations. Just as the manifestation of 'opacity' applies equally to chalk and gold – each different and un-substitutable, the manifestations of 'success or leadership' also apply to organisations where the antecedents of such behaviours may be quite different and un-substitutable. Relata under a supervenient relationship between exemplar and anomalar with only dispositionally entailed environmental variables cannot be *efficient* causal relata because they do not wholly satisfy the Humean-Mill causal identity:

$$A \Leftarrow B, \text{ iff } A \rightarrow B \text{ and } \{A\} \rightarrow \{B\}.$$

Even if $A \rightarrow B$, there is no expectation that $\{A\} \rightarrow \{B\}$ is always obtained. Dispositionally entailed relata cannot be Aristotelian *material* causal relata if they are different and un-substitutable. Differentiation between *formal* and *final* causal relata lies in distinction between laws that describe the nature of relata and the nature of their purpose. Where two relata only share dispositional properties but the laws describing their nature differ, they can only be the relata of some *final* cause. An organisational chain of *final* causes represents its teleological trajectory.

Qualifiers: A supervenient and dispositionally entailed relationship between exemplar and anomalar states of affairs is a necessary but insufficient condition for improvement.

6.3 Summary.

Unless a perfectly effective benchmarking process is obtained, conditions that confer *prior* effectiveness of a benchmarking process do not necessarily confer *prior* attainment of a benchmarking improvement because a supervenient relationship is a statement of a nomological relationship rather than equivalence. Improvement in the absence of nomological equivalence is based on the feasibility of transferring the status of exemplary environmental variables into an anomalar's state of affairs. The effect of this transfer, where it is feasible in respect of a state of affairs' *material* and *formal* cause, may still deny improvement if *efficient* or *final* causal criteria cannot be obtained. For example, it is easy to conceive an anomalar being unable to transfer exemplary environmental variables because of underlying resource capacity constraints. In some cases a perfectly acceptable benchmarking outcome might only be a limited degree of improvement such

as a small change in organisational purpose or the removal of one of many *efficient* causal constraints.

Wherever an effective benchmarking process is obtained, there is every reason to anticipate either a tangible improvement or a set of possible policies (*efficient* causal or teleological) that would result in improvement. However, where an effective benchmarking process is precluded by these theoretical conditions, there is no reason to anticipate improvement.

The definition of benchmarking offered in Chapter 2 may now be revisited in the light of these theoretical constructs by including the necessity of transforming feasible exemplary relata:

'Benchmarking is an exemplar-driven teleological process operating within an organisation with the objective of intentionally changing an anomalar's existing state of affairs into a superior state of affairs *via the transformation of feasible exemplary relata*'.

Extending the definition through the influence of BT2 *and* the disjunction of BT3, BT4 and BT5 (BT3∨BT4∨BT5), provides a definition that not only defines what benchmarking is, but also how it is to be achieved. Feasibility is obtained from the condition:

$$\mathbf{BT1 \wedge BT2 \wedge (BT3 \vee BT4 \vee BT5).}$$

Equation 6-1. Benchmarking Feasibility

If the influence of the primal axiom is also considered, benchmarking contributes to continuous improvement. Superiority is implicit in BT2: improvement in welfare is the only effective improvement recognised by this definition. Any other improvement – such as might be observed in changes to magnitudes of important organisational relata, are only effective if they also improve organisational welfare.

This Theory has many implications for practitioners. The next Chapter examines these implications for current benchmarking practices.

Chapter 7. Interpreting the Theory of Benchmarking.

7 Introducing the Theory to Practitioners.

The logical conditions of the Theory of Benchmarking predict certain approaches and circumstances will obtain effective outcomes whilst others will not. Translating this Theory into practitioner-amenable procedures addresses the practical implications of the research question by explaining how to distinguish between effective and ineffective efforts.

Practitioner experiences associate successful efforts with rather broad attributes or circumstances which have been included in popular multi-step procedures (Chapter 2). A demonstration of the insufficiency of these procedures is the extent to which organisational resources must be deployed to compensate for their imprecision. The first hurdle facing practitioners is how to identify what to benchmark and with whom. This consideration is quite apart from also justifying whether their choices will also contribute to future success. Although quantitative methods such as AHP, CFA, PCA and Priority Mapping (Chapter 2.3.2) have been suggested to assist with the task of selecting what to benchmark, literature has not been as specific on techniques for determining suitable exemplars or explaining interdependencies between these two tasks.

This chapter takes the Theory of Benchmarking's logical conditions, examines each of them within a selection of organisational contexts and collates them into a series of practitioner-amenable procedures. The original research question is further addressed through the creation of a theory-based process describing a sufficient set of practitioner actions that distinguish effective from ineffective benchmarking efforts.

The Theory Filter, Figure 6-1, identified three necessary criteria for theoretical constructions but the causal or logical criteria are of greatest importance as they determine what is or is not the case in detailed circumstances. Previous chapters have demonstrated the empirical and associative conditions of the Theory of Benchmarking to be robust yet they provide no substitute for demonstrating benchmarking's causal nature.

7.1 Implications of the Theory of Benchmarking.

The theory rests on five necessary causal or logical conditions (BT1: BT5) that may also be represented as practitioner-orientated criteria. Each of these conditions will now be examined within a selection of organisational contexts to emphasise or distinguish practitioner efforts that obtain effective benchmarking. The theory is generic. It is not possible to examine every conceivable organisational framework, nor is it possible to encompass innovations that might conceivably arise from any deliberative benchmarking exercise through what Aristotle would call 'chance' (Chapter 3) or in modern times, 'co-incidence or serendipity'.

7.1.1 BT1. Peircean Causation: Benchmarking's Causal Engine.

Peircean causation describes an indissoluble association of process and purpose. To be effective, benchmarking must not only confer changes onto anomalous organisational processes but also maintain alignment with exemplary organisational purpose. *Efficient* causation models of benchmarking are drawn to exemplary performance without regard to underlying teleology. In the most simple of situations this may hardly matter. If the attainment of a more efficient manufacturing process is all that an anomalar desires, *efficient* causation suffices, but desiring the attainment of a better complex state of affairs or organisation is another matter.

The implication of BT1 on exemplary performance is two-fold. A qualifying exemplar in an organisational benchmarking process is one whose production technologies are understood separately from their contribution to its teleological trajectory. If the latter is unknown, there is a risk that important, undisclosed relationships between the exemplar's production technologies and its strategic purpose may be ignored. If so, this would risk Peirce's criticism that *efficient* cause without *final* cause is 'mere chaos' (Chapter 3.8.3).

Information-orientated criticisms of benchmarking were outlined in Chapter 2.4.1. Anomalars risk uncompetitive imitation or fruitless gains unless exemplary production is also aligned with purpose. However, if anomalars do not fully understand the causal nature of their own states of affairs, pursuit of exemplary practices elsewhere would be premature and fuel earlier criticisms that benchmarking can distract or misdirect practitioner efforts.

Effective benchmarking practices necessarily require a full understanding of both anomalous and exemplary practices and their respective contributions to organisational purpose – i.e. Peircean causation. This level of understanding is also hierarchical. The remaining theoretical conditions require knowledge of the relationships between causation and welfare within an organisation's principal states of affairs.

7.1.2 BT2. Welfare Improvement.

Effective benchmarking requires an increase in anomalar welfare via transformation of exemplary relata into feasible anomalar relata. An understanding of the causal nature of benchmarking partners' processes is necessary; attainment of a superior state of affairs is central to the essence of benchmarking.

Apparently favourable changes in some organisational metrics, such as improved output volumes, are not measures of effectiveness unless they also improve overall welfare. This is easily demonstrated by way of example.

If an anomalar's current state of affairs exhibits decreasing returns to scale, a benchmarking improvement strategy that feasibly altered some of its input environmental variables to achieve greater output volumes might also reduce its welfare. Unless the anomalar improves both its technical and allocative efficiency, greater outputs may have been obtained at higher overall cost thus reducing utility and overall welfare. Figure 7-1 illustrates this.

A simple exemplary production system is modelled by a Cobb-Douglas function of the form: Output = A. Input $^{\alpha}$. This model exhibits decreasing returns to scale for $\alpha < 1$. The locus of exemplary production technology performance is the technical efficiency frontier. The allocative efficiency frontier is formed from the locus of production utility: in this case the production surplus calculated from the application of an input mark-up factor, (1+ r).

In this instance (A= 1, α = 0.7, r = 0.5, presenting no special case), production utility declines when inputs exceed about 1.2 units and turns negative at about 3.9 units – also delimiting the point of ongoing, positive welfare. Anomalar adoption of any exemplary

production strategy on the technical efficiency frontier is insufficient to guarantee a benchmarking advantage as the price of production also matters.

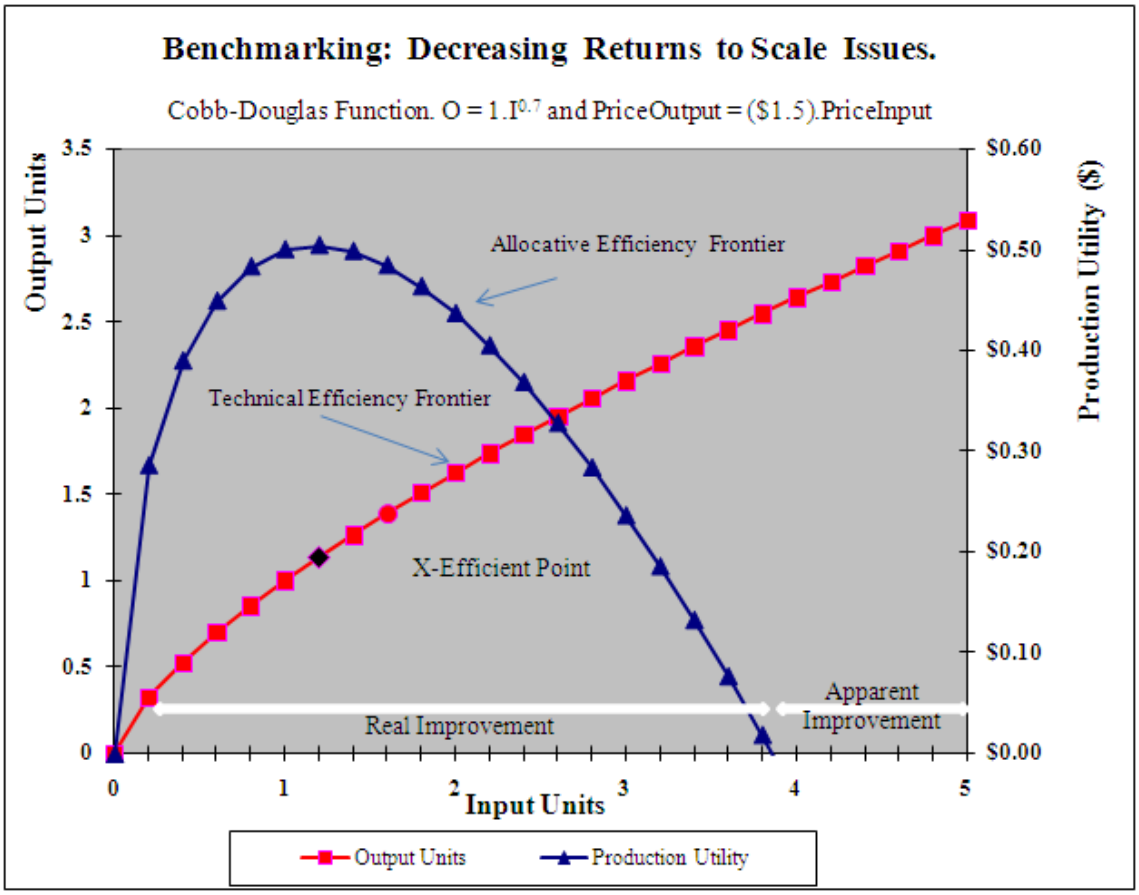

Figure 7-1. Real and Apparent Benchmarking Improvement Strategies.

If the anomalar obtains the details of the exemplar's production technology (in this case the production factor α) but production costs remain confidential, the benchmarking outcome may not be effective even if it is more technically efficient than before. If the anomalar implemented a feasible production strategy with 5 input units with the exemplary production technology, the resulting 3.1 output units lie on the efficient frontier but could also deliver negative utility. This exemplar's production environment is sensitive to production costs and offers highest utility at low input levels. This benchmarking exercise would be ineffective unless the anomalar also obtains sufficient cost information to determine the locus of exemplary utility and establishes an improved operating point for the utility of its state of affairs. In this example there is also an optimum operating point. However any operating point below 3.9 input units would deliver positive utility (and welfare) at current costs and be effective if it improved anomalar welfare.

Implementation criticisms of benchmarking (Chapter 2) cited the degree of trust and level of detailed information that anomalars receive from exemplars as key contributors

to successful benchmarking programmes. Operating costs were cited as particularly sensitive and are generally held confidential. As Figure 7-1 demonstrates, detailed knowledge of exemplary production technology is insufficient as detailed knowledge of production costs must also be known to assure welfare-enhancing strategies. Improvement claims based on apparently favourable changes to the magnitudes of anomalar environmental variables (e.g. volume related improvements) should stand suspended until supported by favourable changes in overall utility and welfare.

In summary, successful benchmarking strategies improve, rather than optimise. Successful practitioner efforts include any alliances between the exemplar and anomalar that overcome barriers to determining the causal relata establishing the technical and allocative efficiency of the exemplar's state of affairs, or more practically, the production technology and its operating costs. Anomalars must also be able to establish the effect of these relata on their overall welfare, not just the welfare of a particular state of affairs. Unless these relata can be obtained reliably and their effects examined in relation to overall anomalar welfare there are good reasons to doubt that practitioner efforts will be effective.

7.1.3 BT3. Strong Supervenience and Logical Entailment.

Perfectly effective benchmarking under this logical condition reduces to perfect imitation of the exemplar's state of affairs by the anomalar. Although this is an unlikely situation between competitors or unrelated organisations, it is relied upon for effective internal and (some) functional benchmarking exercises where processes operate identically in different organisations (e.g. financial or material requirements planning processes).

Strong supervenience and logical entailment relates to benchmarked states of affairs not the entire organisation. Thus perfectly effective benchmarking for one state of affairs does not automatically obtain effective benchmarking at organisational level unless BT2 is obtained. This distinguishes internal from functional benchmarking.

If an internal exemplar is benchmarked, it might be reasonable to expect overall welfare to improve as long as transferred relata do not trigger unfavourable reactions elsewhere and reduce overall welfare. In such cases, the teleological trajectory has also been

altered and requires restoration before the benchmarking exercise can be deemed effective.

If an external exemplar is benchmarked, it is unreasonable to expect that exemplar and anomalar states of affairs will share a common teleological trajectory. Hence it is unreasonable to assume improved anomalar welfare results solely from improvements to one of its state of affairs. Another example of this would be improvements to the wrong state of affairs.

If the exemplar's teleological trajectory resulted in rapid abandonment of its benchmarked state of affairs due to some proprietary advancement, the anomalar might still enjoy short term benefit, but will face longer term isolation from exemplary practice and an overall loss of welfare.

In Chapter 2, many citations placed good benchmarking practices beyond that of imitation. The theoretical implication of these citations is the conjunction of BT3, BT1 and BT2 that establishes relationships between *efficient* and *final* causal processes and improves overall anomalar welfare.

In summary, BT3 specifies how practitioners can implement perfectly effective benchmarking but also directs their attention to broader considerations prior to implementing any exemplary relata Thus perfectly effective benchmarking exercise entails change at tactical levels (environmental variables), improvement at process (states of affairs) and strategic levels (the overall organisation) within its teleological trajectory.

Finally, the requirement for a strong supervenience relationship excludes benchmarking partners where there is no such relationship.

7.1.4 BT4. Supervenience and Proper Subsets of Environmental Variables.

Potentially effective benchmarking addresses circumstances where there are some *formal* similarities between the environmental variables and states of affairs of the exemplar and anomalar, i.e. at least weak supervenience. An example of this could be a shared production technology which has been modified by the exemplar to improve overall quality.

If the anomalar's state of affairs can be improved by drawing *any* lessons from the nature of some exemplary common causal relata, the condition is fulfilled. For example, the exemplar might invest in more appropriate training, or deploy some resources in subtly superior ways to confer a benchmarking advantage on the anomalar. Similarly, a small organisation might draw upon selective improvements by observing a sophisticated exemplary process pursuing objectives broader than those required to meet their needs, thus fulfilling the supervenience and entailment conditions. There is no requirement for optimality, only the necessary requirement to also satisfy BT2 and obtain overall improvement in anomalar welfare.

As with BT3, BT4 also states that benchmarking should not be expected to improve performance where there is no supervenience relationship. For example there is no basis for benchmarking improvement arising from a proposal where benchmarking partners have common environmental variables such as water, electricity or human resources (entailment) but the rules governing the manner of their use (supervenience) differ. Thus a practitioner's selection of benchmarking partners should be prioritised by the strength of their supervenience rather than by the commonality (entailment) of some of their respective environmental variables.

7.1.5 BT5. Supervenience and Dispositional Entailment.

Whereas BT3 and BT4 stressed logical (or probabilistic) entailment of properties between benchmarking partners' states of affairs, this condition recognises that effective benchmarking may arise if their teleological trajectories supervene. For example, if benchmarking partners rely on a common framework of generic organisational dispositions (e.g. lead, grow, manage, measure, value, learn, etc) to influence their teleological trajectories they also establish a supervenient relationship. Dispositional relationships provide the basis for improvements arising from what Zairi's taxonomy calls generic benchmarking (Chapter 2).

BT5 predicts that the scope of generic benchmarking has limitations. For example, dispositional characteristics may not readily identify the key environmental variables associated with an exemplary teleological trajectory or even its states of affairs. Manifestations of dispositions - such as leadership, process management, knowledge management and measurement, also contribute to the basis of the respected Baldrige

business excellence model (Chapter 2). Whilst the application of any model imposes supervenience and entailment on its subjects, the Baldrige model is non-proscriptive and generic: i.e. its structure entails a set of dispositional states of affairs behaving in accordance with selected empirical relationships that are deemed exemplary. The Baldrige model is also a teleological model: it asserts that survival is greatly enhanced if interlinked states of affairs evince particular dispositions (e.g. lead, grow, plan, learn, etc) that are empirically held to be exemplary. These dispositions may be further qualified by other exemplary factors (e.g. measure, analyse, etc) that are also primarily dispositional as they apply equally to completely different environmental variables.

Because of its generic nature, the Baldrige model can form a bridge between existing internal, competitive, functional, and generic benchmarking scores. Comparisons between an organisation's scores obtained from prior competitive, functional or generic benchmarking programmes and its Baldrige score combines BT5 with BT3 or BT4. Comparing Baldrige dispositional scores (e.g. leadership) with other benchmarking scores based on environmental variables or states of affairs contributing to these dispositions may help to isolate or infer potential improvements at these more detailed levels.

McAdam and Kelly (2002) examined the generic benchmarking aspect of this bridge. They examined eight SMEs operating in quite different markets and compared their individual Baldrige scores with those of a local quality agency's own independently derived benchmark scores based on generic success factors (e.g. services quality, time to delivery, etc). This research concluded that the business excellence model could be combined with a generic benchmarking framework and suggested it might extend to 'strategic, operational or tactical levels' – reflecting the combination of BT5 with BT3 or BT4 where other data exists. They also echoed earlier referred-to benchmarking criticisms where 'over-systematisation' of this model may 'distract SMEs from their main goals and consume scarce resources' (Chapter 2).

A similar observation arises for quality management standards such as ISO 9000, another example of process standardisation via a dispositional model. ISO 9000 (1994) seeks to standardise production processes through assurances that practice conforms to specification and ISO 9000 (2000) broadened this to include enabling (management) and supporting (measurement) processes. ISO 9000 is represented as a holistic system

of standards for production process excellence whereas Baldrige is represented as a holistic model for organisational excellence. There is a degree of overlap between them (Mann & Voss, 2000) where benchmarking is an implicit component of their assessment criteria (Porter, 2004).

BT5 provides explanation for generic benchmarking and generic business excellence models. Such models are also de-facto exemplars. If practitioners associate dispositional behaviour with tactical behaviour through successful implementation of a business excellence model or otherwise, they combine BT5 with BT1 and BT2 to form a sufficient set of conditions for an effective benchmarking process.

7.2 An Effective Benchmarking Process based on Theoretical Conditions.

The five logical conditions represent necessary conditions for effective benchmarking. The logical expression developed in Equation 6-1 is now asserted to establish sufficient conditions for effective benchmarking:

$$\mathbf{BT1 \wedge BT2 \wedge (BT3 \vee BT4 \vee BT5).}$$

Equation 7-1. Sufficient Conditions for Effective Benchmarking

Compiling the implications arising from each proposition provides the basis for a practitioner-orientated process that mirrors Equation 7-1. This process may be summarised by noting that an anomalar must have the capacity to conduct effective internal benchmarking before it can pursue effective external benchmarking with an exemplar.

This logically sufficient, effective benchmarking process (EBP) rests on seven internal and two selection steps (which may be either internal or external). Each step in the process shown in Table 7-1, overleaf, complies with the Theory of Benchmarking and collectively they meet the requirements of Equation 7-1.

An Effective Benchmarking Process based on Theoretical Conditions			
Step No	**Process Steps**	**Anomalar Practitioner Actions**	**Theoretical Basis**
1	Internal 1	**Establish a Welfare Framework and Relationships.** Formally identify principal states of affairs having a significant influence on overall economic welfare (e.g. free cash flows, EVA®, etc, relate to the economic welfare of principal states of affairs and also identify relevant key environmental variables). Quantitative techniques (e.g. AHP, PCA, CFA or root cause analysis, etc) may also contribute to this step.	BT2
2	Internal 2	**Establish the Current Organisational Trajectory**. Formally establish and confirm the mutual dependencies of future plans and proposed organisational (teleological) trajectory on principal states of affairs.	BT1
3	Internal 3	**Internal Exemplar Decision.** Re-examine current states of affairs and environmental variables for evident internal benchmarking improvements arising from Internal Steps 1 and 2. This avoids the criticism that distractions and inefficiencies might arise from unnecessary external benchmarking. Note that improved internal states of affairs *should* implicitly supervene upon and entail unimproved internal states of affairs.	BT2
4	Internal 4	**External Exemplar Decision**. Formally embark on a search for an external exemplar if internal knowledge cannot articulate the improvements necessary to increase overall welfare by an amount sufficient to satisfy those upon whom reliance is made for the continued supply and maintenance of resources.	BT2

Table 7-1. An Effective Benchmarking Process based on Theoretical Conditions.

An Effective Benchmarking Process based on Theoretical Conditions, Continued			
Step No	**Process Steps**	**Anomalar Practitioner Actions**	**Theoretical Basis**
5	Selection 1	**Benchmarking Partner Selection: Exemplar Style.** Style is the conjunction of two elements: the primary element is the alignment of respective teleological trajectories and the secondary element is the type of entailment between states of affairs. Unless the exemplar supervenes (broadly evinces the same approach to survival) on the anomalar there is no point in considering entailment.	BT1
		A 'logical style' is based on the view that an exemplar can be found that either perfectly or closely matches the anomalar's states of affairs and environmental variables. For example, organisations producing goods and services in a similar manner are candidates so long as their organisational (teleological) trajectories are consistent with the anomalar (i.e. they operate within the same nomological framework).	BT3, BT4
		A 'probabilistic style' is also a 'logical style' except that states of affairs and their environmental variables obey statistical laws. Examples of this style might be found in organisations such as those associated with process-flow manufacturing, mining, or speculative services.	BT3, BT4
		A 'dispositional style' is based on the view that any exemplar manifesting selected, superior generic behaviours (e.g. a Baldrige winner, a successful exporter, a zero-waste leader, an ISO certificated organisation, etc) is also sufficiently nomologically similar to potentially transfer some knowledge applicable to current environmental variables or states of affairs (i.e. to tactical or strategic levels) that may increase organisational welfare.	BT5
6	Selection 2	**Benchmarking Partner Selection: Exemplar Alliance.** A suitable benchmarking partner is an exemplar willing to share sufficient access and detail to establish correspondence between exemplary states of affairs (and their environmental variables) and corresponding anomalar states of affairs formalised in Internal Step 1. For example, an unsuitable benchmarking partner withholds critical elements of their production technologies or resource allocations making it difficult or impossible to transfer information between corresponding states of affairs.	BT3, BT4, BT5

Table 7-1. An Effective Benchmarking Process continued.

An Effective Benchmarking Process based on Theoretical Conditions, Continued			
Step No	**Process Steps**	**Anomalar Practitioner Actions**	**Theoretical Basis**
7	Internal 5	**Benchmarking Knowledge Transfer.** Transfer feasible exemplary information to the anomalar. Analyse exemplary information within the context of Internal Step 1. Determine hypothetical improvements. Improvements must relate to overall organisational welfare within a new trajectory, not just those environmental variables or states of affairs under consideration.	BT2
8	Internal 6	**Benchmarking Durability.** Durable improvements also favourably align the teleological trajectories of the anomalar and the exemplar. If organisational (teleological) trajectories are on the cusp of divergence, tactical gains may dissipate if the exemplar's strategic direction changes rapidly. A benchmarking exercise resulting in transient gains at the expense of strategic direction risks premature dissatisfaction by those upon whom the anomalar is reliant for the supply and maintenance of resources. This step revisits assumptions made during the Alliance step to ensure that the anomalar is satisfied with the overall nature of the improvements prior to implementation.	BT1, BT2
9	Internal 7	**Implement Improvements.** Embed feasible exemplary information within anomalar states of affairs. Update Internal Steps 1 and 2 to reflect the new organisational (teleological) trajectory. Continue at Internal Step 1.	BT1, BT2

Table 7-1. An Effective Benchmarking Process continued.

7.3 Summary of Results.

The Theory of Benchmarking's logical conditions identify three benchmarking rules for successful practitioner efforts. These rules may be summarised as:

- Change must be viewed as holistic (BT1),
- Change must improve (BT2),
- The basis for change necessarily involves a purposeful alignment of relationships (supervenience) and properties (entailment: BT3 $\vee$ BT4 $\vee$ BT5).

Change requires knowledge of current and final states. Improvement requires assessment of these states. Successful benchmarking necessarily requires an anomalar to identify and associate current practices with its espoused strategic direction before it considers the task of selecting an external exemplar.

Although benchmarking always requires an exemplar, it does not follow that it needs to be from a different organisation as different 'forms' of states of affairs may suffice. Indeed, there may only be a notional exemplar, reflecting a more efficacious *material*, *formal* or *efficient* cause that is revealed when an organisation undertakes internal benchmarking steps one and two of the EBP.

The EBP emphasises welfare improvement through the application of supervenience and entailment. Theory also illustrates how well-known business excellence models can assist practitioners by providing an exemplary supervenient and dispositionally entailed framework for implementing the internal benchmarking steps of the EBP. Even though these models address the holistic and nomological requirements for successful practitioner efforts, an overall improvement in welfare is also necessary and no less challenging.

Zairi's taxonomy (Chapter 2) encompasses four forms of benchmarking practice (internal, competitive, functional and generic), but this new Theory identifies no essential need for different forms as Benchmarking is Benchmarking. This theoretical conclusion supports an earlier empirical observation (Chapter 2.3.1) that current forms of benchmarking only appeared to identify the locus of the exemplar. There is likely to be pragmatic distinction between the selection of internal and external exemplars. This distinction may arise if the practitioner undertakes pragmatic steps to obtain consensual access to exemplary relata. It may be pragmatic (but logically unsafe) to assume that both an internal exemplar and anomalar share the same teleological trajectory, especially in large, complex and operationally diverse organisations. It would be patently unsafe to assume that any external exemplar shared the same teleological trajectory as the anomalar. If an external anomalar were selected, its teleological suitability must also be established using the principles of supervenience and entailment. Different forms of supervenience (weak,

strong or global) and entailment (logical, probabilistic or dispositional) serve to exhaust the question of an exemplar's appropriateness.

Practitioners will find that successful benchmarking opportunities most likely arise from situations where exemplars and anomalars share simple states of affairs that have direct, observable relationships with welfare-orientated financial measures such as free cash flow or EVA®. Situations such as these obtain the necessary supervenience and entailment conditions for effective benchmarking with much greater success than is likely from situations where there are complex states of affairs or volumetric imperatives (market share, revenue, production units) lacking clear linkages to organisational welfare.

The EBP addresses the two key tasks, what to benchmark and with whom, by way of generic procedures. Practitioners following these procedures may distinguish between effective and ineffective efforts at any stage on the basis of theoretical rather than empirical compliance.

The next chapter will discuss the Theory of Benchmarking and its EBP with respect to the Theory Filter or validator developed in Chapter 4.

Chapter 8. Validation of the Theory of Benchmarking.

8 Approach to Theoretical Validation.

According to Popper (1987; 1996) the 'correct method of critical discussion starts from the question: what are the consequences of our thesis or our theory? Are they all acceptable to us?' Weick (1989) addressed the question of 'acceptability' claiming good theory to be 'plausible, and more plausible if it is interesting (versus obvious, irrelevant or absurd), obvious in novel ways, a source of unexpected connections, high in narrative rationality, aesthetically pleasing or correspondent with presumed realities'. Chapter 4 identified three filters or validation criteria that improve the task of determining acceptability and determining correspondence with presumed realities.

Introductory Section 1.2 identified the particular challenge of theory validation solely in the presence of praxis. The Theory of Benchmarking developed in this thesis (the Theory) and its Effective Benchmarking Process (EBP, Table 7-1) have been developed solely in reliance of the objectives of such praxis: exemplar-driven organisational improvement. This Chapter examines the degree to which the Theory and the EBP can explain efficacies and inadequacies within current praxis, satisfy its critics and extend knowledge of benchmarking without sole reliance on empirical methods.

There is a significant body of praxis, 'the dataset', to draw upon as practitioners have produced numerous benchmarking definitions and have created scores of benchmarking frameworks or procedures describing implementation in either generic or particular situations.

Kozak *et al* (2001) cited the existence of over forty different frameworks originating from organisations, consultancies and individuals. Andersen *et al* (1999) analysed sixty different (although unspecified) frameworks in their reformulation of the familiar PDCA framework. Anand *et al* (2008) compared thirty five frameworks in their search for an exemplary benchmarking framework citing Watson's (1993) earlier survey of sixty nine frameworks. All of these frameworks are factor-based sequences of processes identifying empirical factors (and sub-factors) having some degree of implementation provenance, but not necessarily reproducibility by others. These frameworks have already been criticised (Francis *et al.*, 2007) for their causal (factors specifying 'what', rather than 'how') and

logical (processes emphasising increased necessity rather than reduction to sufficiency) inadequacy. In summary, Spendolini (1992) and Anand *et al* (2008) have, between them, distilled at least fifty nine extant frameworks into their respective examples of an exemplary framework and thus provide an efficient and reliable way to validate the Theory against the overwhelming majority of current praxis.

A key result of the Theory was presented in Chapter 7. Its causal and logical constructs were presented within a *sufficient* practitioner process to obtain *effective* benchmarking.

The provisional definition of benchmarking presented in Chapter 2 recognised its potential for exemplary contributions to organisational teleology via a causal mechanism founded on an axiom of organisational ontology. This provisional definition was extended to its final form in Chapter 6 by incorporating necessary conditions for delivering effective benchmarking. In particular, the final definition included a theoretical requirement for feasibility, underpinned by *formal* and *material* causal relata and obtained by the application of supervenience and entailment.

This chapter will analyse the consequences of theoretical and practical outcomes of this research against the theory filter or validator outlined in Chapter 4 and critically discuss the degree to which the Theory and the EBP are validated with respect to epistemological, associative and logical elements.

Two principal epistemological elements must be examined. Benchmarking's current definitions as well as its numerous implementation frameworks must be analysed with respect to the Theory's conditions and its final definition.

8.1 Validation: Benchmarking Definitions.

A provisional definition presented in Section 2.5 described the objective of benchmarking but was silent on how it functioned Establishment of the Theory improved this definition (Chapter 6.3) by appending its functionality. This definition will now be validated against the 'dataset' of current epistemological attributes presented in Table 2-1.

The review of current benchmarking definitional elements shown in Table 2-1 cited the use of 'feedback mechanisms' to trigger organisational improvement. These current

definitional elements also contain implicit associations. The feedback mechanism is implicitly associated with best practice or superiority since its recognition involves consideration of current and superior states of affairs and outcome results in a decision (to act or not). The definition of benchmarking developed in this research admits purposeful change from an inferior to a superior state of affairs and embodies all of these epistemological attributes, as well as establishing their feasibility and manner of use.

Again, the current definitional attributes of continuous improvement and adaptation/modification are firmly embedded within the primal axiom, BT0, and Peircean Causation, BT1, of this Theory. The pursuit of survival is continuous (Pfeffer, 1997). Adaptation/modification is an attribute of BT1, where an organisation's teleological trajectory is developed to survive.

The principal differences between current benchmarking definitions and the Theoretical definition developed in this thesis are *purpose* and *process*. The Theoretical definition refers to the teleological and transformational nature of benchmarking to recognise Theoretical conditions BT1 and BT2. Current literature, where it does refer to teleology and transformation, places them within benchmarking frameworks or processes rather than its definition. The implications of this omission have also been reported in literature: manifested as practitioner confusion, task underestimation and bad process.

It might be argued that intentional improvement or pursuit of best practice or superiority is implicitly teleological (survival) within a competitive environment. Even so, its explicit omission weakens current definitions and triggers criticisms of benchmarking. For example, if improvement is treated purely as an *efficient* causal activity with respect to its immediate rather than holistic effects, it falls prey to Peirce's criticism that doing so invites 'chaos' (Chapter 3.8.2). In the absence of an accompanying theory, concepts such as best practice, superiority and continuous improvement are decoupled from their *final* cause (transformation) and increase the likelihood of deficiencies such as task underestimation and ineffective practitioner effort.

Deficiencies are most evident within organisations where apparent internal benchmarking successes could disharmonise broader, important resources and lower overall welfare. The essence of improved welfare is its holistic rather than local nature, but this is not evident in current benchmarking definitions although it appears within literature. For

instance, Maire *et al* (2005) observe that 'customers' voices' must contribute to the identification of best practice but they direct their observation to a benchmarking implementation framework rather than to its definition.

Transformation is an essential requirement of any effective benchmarking exercise, yet it is remiss to ignore its feasibility if practitioners wish to avoid ineffective effort. A succinct warning was provided in Chapter 2: 'all improvement is change, but not all change is improvement' (Harrington, 1995). Clearly exemplary relata are implicitly beneficial to the exemplar, but it remains unproven that they can beneficially transform the anomalar until the theoretical constraints of supervenience and entailment are satisfied, i.e. (BT3$\vee$BT4$\vee$BT5). The degree of implicit optimism embedded in benchmarking's current definitional elements lures practitioners into a complex and often confusing process (Alstete, 2008) that, as the EBP shows, is challenging enough with the benefit of theoretical support let alone without it.

Internal benchmarking should satisfy the theoretical requirement for the exemplar to be supervenient upon and entail the anomalar's environmental variables (e.g. BT3) as both are located within the same organisational system. It is unsurprising to see internal benchmarking cited as preferred over external forms (Bhutta *et al.*, 1999; Southard & Parente, 2007), or cited as a major contributor to failure where ignored (Huq *et al.*, 2008) or underutilised as a basis for external benchmarking (Elmuti & Kathawala, 1997). These criticisms are addressed by the first three steps in the EBP (Chapter 7) as 'necessary' steps within a 'sufficient' process.

Confusion arising from current definitions of benchmarking have been reported by Alstete (2008). The terms 'benchmarks and benchmarking' were both perceived and used differently in different industries and in differently sized organisations. The observation that 'a more precise terminological use of "real" benchmarking should be promoted by management leaders, educators and writers' (Alstete 2008) is consistent with the objectives of this research.

8.1.1 Summary

Current benchmarking definitions are generally incomplete as they omit two necessary Theoretical elements (BT1 and BT2) that serve to establish the true nature and magnitude of the task. These necessary elements may be implicit within current

benchmarking frameworks but by the time that stage is reached practitioners are implementing rather than planning and crystallise problems that might have been avoided. Practitioner expectations are also reinforced by business excellence frameworks that insist upon the inclusion of benchmarking as an instrument of best practice. Literature demonstrates it unsatisfactory to assume that a lack of theoretical provenance in the very definition of benchmarking can be compensated for by inclusion in an implementation framework. These criticisms are not eliminated by reliance on close adherence either to business excellence frameworks, which have an implicit Theoretical underpinning (Chapter 7), or to current benchmarking frameworks.

The Theory-based definition of benchmarking should advance practitioner understanding by emphasising necessary theoretical criteria addressing overarching purpose, implementation and constraints. The Theory-based definition should temper rather than re-establish current perceptions of the nature of benchmarking and improve practitioner efforts.

8.2 Epistemological Validation: Benchmarking Implementation Frameworks.

In the following sections, the earlier referred to implementation frameworks of Spendolini and Anand and Kodali represent the principal datasets against which the Theory is validated. In addition, other datasets related to the PDCA approach are also briefly examined to extend this validation.

These chosen frameworks represent a reliable dataset derived from distillation of over sixty nine examples of current benchmarking praxis. This approach is valid because whilst there are numerous implementation frameworks they have already been identified as variations on the basic elements of the PDCA framework.

8.2.1 Validation against Spendolini's (1992) Generic Benchmarking Framework.

If Spendolini's (1992) nomenclature is adopted for comparative purposes, 'elements' refer to a framework's generic structure and 'steps' or 'phases' refer to their subdivision.

This framework is arguably the first thoroughly documented approach to effective benchmarking and was prompted by the spectre of 'model wars'. By 1992 there were

already significant numbers of corporate models (Xerox, AT&T, Alcoa, etc) comprised of varying numbers of elements and steps. Spendolini observed that all models reflected a common theme, but tended to confuse principle with practice as guidelines for using them varied. Out of a pool of thirty five, twenty four exemplary organisations were surveyed in some depth to conclude that a successful benchmarking implementation programme relied upon four key principles (P), governing a five stage generic benchmarking process (B) under the umbrella of four empirical practitioner caveats (C).

Each of the sections within Table 8-2 assess Spendolini's generic framework against the Theory of Benchmarking on the basis of three scores: to indicate whether the Theory leads (T+), concurs with (T) or lags (T-) the framework.

Spendolini's Generic Benchmarking Framework (1992)	**Theory of Benchmarking's Response**	
Section 1: Benchmarking Principles		**Score**
P1. Follow a simple, logical sequence of activity	If a process is predicated on (implicit) necessary conditions they are also causal (for success). Spendolini does not claim process 'sufficiency' whereas EBP does.	T+
P2. Place heavy emphasis on planning and organisation	This is interpreted to mean that all key *efficient* causal processes must be thorough.	T
P3. Use customer-focussed benchmarking	P3 is historically limited as the broader concept of stakeholder-focussed satisfaction emerged much later. The Theory's primal axiom of survival evinced by a teleological trajectory is a more holistic approach to P3.	T+
P4. Make it a generic process.	The Theory is generic to organisational improvement using benchmarking. P4 emphasised that organisations should only have ONE standard benchmarking process to avoid internal inefficiencies in implementing their programmes. The basis for P4 was avoidance of 'model warfare'; it was not a theoretical condition.	T+

Table 8-1. Theoretical Assessment of Spendolini's Generic Framework.

Section 2: The Generic Benchmarking Process		Score
B1. Determine what to Benchmark	The primal axiom and {BT1, BT2} provide a theoretical basis as to what should be benchmarked not how it should be performed. Spendolini's advice has been extended in subsequent research by tools such as AHP and CFA to assist practitioner determination of what to benchmark. These tools remain pertinent within the Theory.	T+
B2. Form a Benchmarking Team	The Theory does not explicitly dimension resources required for effective benchmarking. *Prior* indication of resource requirements may be established because the Theory identifies the 'sufficient' steps for effective benchmarking.	T+
B3. Identify Benchmarking Partners	The Theory identifies a partner as one prepared to meet all requirements that obtain any of {BT3, BT4, BT5} for the anomalar. This is one of the principle weaknesses of Spendolini's framework if it applies to other than market giants (Xerox, AT&T, etc) who are resourced to possess intimate knowledge of their competitors and global best practices via their corporate suppliers.	T+
B4. Collect and Analyse Benchmarking Information	The Theory identifies 'feasibility' as a further constraint on exemplary completeness. The Theory also recognises Spendolini's B4 as involving both *efficient* and *final* anomalar causal relata (BT1) in transforming feasible exemplary environmental variables.	T+
B5. Take Action and Continue	The final step of the EBP concurs, although it also recognises the new teleological trajectory of the anomalar. B5 offers explicit reference to a teleological trajectory beyond that which is broadly implicit in the concept of 'continuous improvement'. There is only an implication of overall welfare improvement.	T

Table 8-2. Theoretical Assessment of Spendolini's Generic Framework, Continued.

Section 3: Benchmarking Caveats		Score
C1. Know yourself before you seek to know others	{BT1, BT2} define the degree of knowledge required to relate organisational teleology and organisational value to satisfy the primal axiom (survival).	T
C2. Focus on improvement of practices rather than objects or things.	Practices are interpreted to mean *efficient* causes at process level but do not explicitly extend causality elsewhere in the organisation. For example teleological effects and infeasibilities arising from a lack of entailment are not covered.	T+
C3. Follow the programme diligently	If an operational process is predicated on (implicit) necessary conditions they are also Humean/Mill/Peircean *efficient* causal (for success).	T
C4. Resource the process	The Theory does not explicitly dimension resources. The Theory provides implicit *prior* dimensioning of resources to address supervenience and entailment relationships (internal or external) with the exemplar.	T+

Table 8-3. Theoretical Assessment of Spendolini's Generic Framework, Continued.

In all respects the Theory and its associated EBP 'leads' in nine out of the thirteen of Spendolini's categories and concurs with the remainder. Nonetheless, Spendolini's framework remains relatively vague although practitioners within Xerox clearly enjoyed greater success in its deployment than their counterparts elsewhere – most likely because of their standardised approach and internal support structures.

As with analysis of current benchmarking definitions in Table 2-1, it is prudent to extend the benefit of doubt to criteria that may be implicitly teleological or transformational. For example, in comparing Xerox's performance with Japanese exemplars, Kearns, observed that 'in category after category, the difference wasn't 50 percent better or anything like that; it was almost always over 100 percent!' (Kearns & Nadler, 1992). Performance gaps of this magnitude in key organisational processes implicitly encompass anomalar welfare but in other cases the nuance might be less obvious. As with the definition of benchmarking, practitioner efforts would be enhanced if frameworks avoided doubt and affirmed the necessity for overall welfare improvement.

Other contemporary benchmarking implementation frameworks (Camp, 1989; Drew, 1997; Bhutta *et al.*, 1999; Carpinetti *et al.*, 2002; Zairi *et al.*, 2003) reflect the same level of concordance with the Theory as they all share a common approach although each may differ in the degree to which steps or phases within each PDCA element isolate or prioritise particular actions their authors believe effective.

8.2.2 Validation against Anand and Kodali's Universal Conceptual Benchmarking Framework.

A recently published framework has been included in this epistemological validation because it has been distilled from the attributes of a great many other frameworks using Spendolini's framework to form an 'exemplary' benchmarking system. The following framework has been presented by its authors as empirical, 'highly conceptual and has not been validated by implementation in industries to determine its effectiveness'.

Anand & Kodali (2008) benchmarked thirty five current benchmarking frameworks using Spendolini's PDCA-based methodology to establish the best set of exemplary characteristics amongst them. Their complete framework and validation analysis based on the Theory is shown in Appendix 3, Table A3.1 (as a result of its size). The authors raised two research questions and repeated an earlier one:

1. How can it be proven that the practices realised are actually the best?

2. Is there need for separate classifications schemes within benchmarking?

3. Can different, seemingly unique frameworks for each type of benchmarking be replaced with a better universal framework, embodying the best practices they share?

Anand and Kodali's principal conclusions were:

1. There are really two kinds of benchmarking: internal and external,

2. Greater detail within the benchmarking framework addresses the pitfalls of benchmarking,

3. Academic-based benchmarking frameworks fare less well in contribution to best practice and would benefit from experiential insight of industrial/organisational frameworks,

4. Popular benchmarking frameworks could be represented as a universal framework comprised of twelve major steps or phases - altogether comprising fifty four sub-steps.

The first principal conclusion attracts practical empathy, but as observed in Chapter 2 and explained in Chapter 7.8, there are no theoretical constraints on the locus of an exemplar. Any distinction between internal and external benchmarking is only to be reflected in the extent of practitioner administration (legal, procedural, financial) necessary to obtain consents and access to exemplary states of affairs. The following anecdote experienced by the author of this thesis serves to illustrate that internal and external exemplars should be treated similarly. In a New Zealand example during the late 1980's, an overseas Branch of a particular Trading Bank was both independently strident and exemplary. Management re-established its states of affairs within the teleological trajectory of the parent organisation before transferring its exemplary relata to other Branches. In this case, 'strident independence' created an unacceptable teleological trajectory within the overall organisation and, until this was remedied, precluded internal adoption of otherwise exemplary operational (*efficient* causal) processes.

The second principal conclusion echoes earlier observations that increasingly refined steps within benchmarking frameworks merely simulate the theoretical requirements of supervenience and entailment. Applying increasingly specific rules or 'more rules' may be an effective *a posteriori* technique for improving empirical frameworks, but can also lead to unintended consequences, such as confusion (Alstete, 2008) or complexity (Deros *et al.*, 2006). This Theory explains the efficacy of the second conclusion in some (mainly simple) situations, and also explains its inefficiency in complex situations.

The third principal conclusion is in similar vein to the second: simulation of theoretical requirements is performed better by experienced practitioners.

The final principal conclusion was that an exemplary benchmarking framework could be constructed from 'best practices' within extant practitioner frameworks. This 'exemplary' framework appears in Appendix 3 and has been assessed in a manner similar to that of the Spendolini generic framework. Applying the Theory to this framework leads to the following observations and conclusions.

a) The majority of the steps within Anand and Kodali's proposed 'exemplary' framework are administrative. 'Identify a leader of the team', 'provide training', 'prepare a proposal', 'obtain management support', etc are all wise administrative matters. However, literature also associates many of these administrative steps with efforts to mitigate the risks associated with current benchmarking frameworks, such as the risks of project overrun and over-estimation of success (Chapter 2.4.2). These risks are explained by the Theory as practitioner efforts that seek to reduce failure by enforcing greater control over benchmarking framework processes without exercising greater control over their essence! The Anand and Kodali framework does not improve understanding of the essence of benchmarking as it offers no basis for improvement other than existing praxis: this 'exemplary' benchmarking framework retains its *a-theoretical* status.

b) Anand and Kodali's framework is a strident example of uncompetitive homogeneity – an earlier criticism noted in Chapter 2.4.1 (Kozak *et al.*, 2001). Belief that 'best benchmarking practice' must reside within current practices rather than elsewhere is dangerous and exposes practitioners to levels of risk that would be untenable if applied to most common organisational practices. Innovation easily dispels the assumption that current practices circumscribe the boundaries of knowledge. It scarcely needs justification that survival depends not only upon the known, but also upon research and the adoption of novel techniques that seek to improve the teleological trajectory – i.e. competitive advantage.

c) A further and final conclusion arising from the validation of the Theory against Anand and Kodali's 'universal' framework is the unlikelihood that continued empirical refinement of PDCA-based frameworks will ever explain effective benchmarking. Improvements to current frameworks have yielded increasingly

unwieldy and complex processes that, in the absence of a theoretical foundation for effective benchmarking, only continue the trend of approximating their undiscovered underlying theoretical criteria.

None of this denies that existing PDCA-based frameworks may be relied upon by practitioners benchmarking *simple and well-linked* states of affairs. However these same frameworks are patently deficient in complex situations and incapable of generalisation because of their vague approximation of underlying theoretical necessities and perpetuate the risks of exposure to the well-documented criticisms and implementation risks of benchmarking.

It is acknowledged that the EBP is silent on the many practical and perhaps essential administrative processes necessary to obtain consent for effective external-exemplar benchmarking. In the light of the analysis and discussion in Table A3-1, it may be a fruitful area for future organisational methods research to consider the inclusion of some administrative guidance within the EBP to assist practitioners.

8.2.3 Validation against other Benchmarking Frameworks.

Spendolini's (1992) conclusion that benchmarking frameworks were essentially generic has endured but not without numerous grafts and branches on its family tree.

Other frameworks have evolved from the basic PDCA philosophy to address particular combinations of organisational function (e.g. strategy, processes, etc), industry type (e.g. manufacturing, health) or scale (e.g. SMEs). These branches of the basic PDCA philosophy remain empirical.

Some benchmarking frameworks have also been grafted onto others. Kaplan and Norton's (1992) Balanced Scorecard (BSC) concept has been examined as a framework for benchmarking strategic intent using AHP methodologies (Ragavan & Punniyamoorthy, 2003; Punniyamoorthy & Murali, 2008) to implement what this Theory identifies as necessary conditions BT1 and BT2. The BSC is implicitly founded on BT0 as it explicitly maps (most of) those upon whom reliance is made for the supply and maintenance of resources: customers, staff and shareholders. However the BSC is less cognizant of other reliance as its recognition of suppliers and the community in which the organisation operates is at best implicit.

In another recent example Deros *et al* (2006) examined PDCA frameworks with the objective of improving SME (automotive sector) benchmarking. They examined nine benchmarking implementation frameworks covering multiple aspects of organisational performance (strategy, organisation, process, design and technology). Five of the nine frameworks were applicable to particular industries (electronics, automotive/aerospace, petrochemical, and public sector) and the remainder were generic.

Two principal conclusions arose from their research:

1. Current frameworks (whether academic or practitioner based) are overly prescriptive, most suited to larger organisations and unsuited to SMEs.

   This conclusion can be explained by the Theory developed in this research. Current prescriptive frameworks have been deployed to empirically simulate the Theory's logical conditions. Empirical simulation of welfare dependencies, organisational teleology, supervenience and entailment has generated sets of rules to guide practitioners by prescription rather than logic. Even so, the Theory probably cannot simplify its causal foundations to such a level of inconsequentiality that all practitioner difficulties vanish, but it will replace empirical prescription with logical explanation.

2. A modified, PDCA-based framework was proposed to suit (manufacturing) SMEs, based on simple, tangible and SME-friendly measures. Examples such as 'reject or rework %', 'Work In Progress (WIP) levels', or 'lead times' can be benchmarked either internally or externally and implemented gradually in pursuit of better rather than best practice.

   This conclusion identifies two further issues, both consistent with the Theory but one contrary to the tenor of current benchmarking definitions.

   a. SMEs generally have simpler states of affairs hence these tangible efficiency measures directly relate to expenditures on labour, materials or working capital. Such metrics also have direct, inverse relationships with free cash flow and organisational welfare. Theory identifies these conditions as satisfaction of BT1,

BT2 and BT3 (for internal benchmarking). It is because SMEs are *simple* rather than *small* that they more easily obtain theoretical sufficiency for effective benchmarking. In this Theory, organisations of *any* size having simple states of affairs should find it much easier to achieve effective benchmarking compared to their counterparts with complex states of affairs.

b. The conclusion that gradual implementation leading to superior rather than best practice is consistent with this Theory, but contrary to current benchmarking definitions. If 'superior' is combined with 'continuous', cyclic benchmarking processes should gradually converge to the 'best feasible' practice (provided cycle times exceed the rate of exemplary progress, etc). This Theory eschews 'best' in favour of a 'superior state of affairs' because there is no optimum teleological trajectory associated with the primal axiom (Chapter 5.4). For SMEs and even for large organisations, the concept of "the best" practice is likely to be a comparatively local phenomenon as it is impractical to determine the best exemplar on any other basis. The locus of 'local' is effectively 'global' for very large organisations.

Collins *et al* (2006) also addressed the issue of how to prove that best practices realised are actually the best and advanced the use of multi-attribute utility theory to handle diverse factors contributing to organisational improvement. Their technique established priorities and compared the effects of trade-offs. Methods such as this assist with the implementation of theoretical proposition BT2 and the first two steps of the EBP. This establishes welfare relationships and the current teleological trajectory. Collins *et al* also concur with the theoretical position that 'best' only implies relative superiority.

8.3 Associative Validation: Extant Paradigms.

Benchmarking's associative provenance has been discussed in Chapter 4.2. Improvement inspired by better purpose has already been summarised as a general rule pertaining to any organisational paradigm: BT0, survival is a *sine qua non* of organisational ontology. With this axiom, the Theory of Benchmarking's five conditions apply within any organisational paradigm that might also, for other reasons, be deemed 'incommensurate' with its competitors.

The Theory possesses two intrinsic components that categorically establish benchmarking's logical locus within any paradigm or between any so-called 'incommensurate' organisational paradigms – namely supervenience and entailment. If there is no supervenience, there can be no effective benchmarking. If there is no entailment of any kind, there can be no successful benchmarking. Supervenience is implicit where there is solely dispositional entailment (Appendix 2). If there are any instances of 'cross-paradigm' supervenience, an 'apparently incommensurate' exemplar may indeed become a participant in an effective benchmarking process. This was remarked upon by Popper (1996), who held that even if there are apparently different axiomatic frameworks, critical rational discussion tests whether their constructions are mutually illuminating. In organisational benchmarking, not all exemplary dispositional characteristics will be so unacceptable or paradigmatically constrained that they preclude supervenience.

The locus of benchmarking was proposed to be 'equidistant' from the meta-centres of Burrell and Morgan's organisational perspectives (Figure 4-3) reflecting generic support of exemplar-driven improvement within any paradigmatic framework. Benchmarking contributes to what Popper terms 'critical rational discussion' because it has an objective role in the pursuit of a better teleological trajectory. The *final* cause of benchmarking is congruent with that of its adherent: to improve the satisfaction of those upon whom reliance is made for the supply and maintenance of resources.

8.4 Further Logical Validation: *a Priori* or *a Posteriori*?

The causal or logical nature of the Theory is implicit in its axiomatic basis (BT0) and corresponding postulates (BT1 – BT5) which have been defended and warranted in Chapter 6. There is also implicit logical validation in the preceding sections of this Chapter through the application of the Theory Filter or Validator (Chapter 4) to the Theory. However, there is a further logical step that appeals in the light of a key criticism of benchmarking: the ability of its practitioners to distinguish between effective and ineffective efforts.

The Theory has been demonstrated to explain historical criticisms and establish a 'logically sufficient framework' that provides practitioners with knowledge of their potential for successful effort. But to what extent is this knowledge experiential or innate? For instance, is the Theory based on *a priori* knowledge, as Galen Stawson colloquially

observed, such that 'you can see that it is true just lying on your couch' (Sommers, 2003)? This raises two final validation questions:

1. Is the Theory *a priori* justified?

2. Does the Theory provide practitioners with an *a priori* approach to benchmarking?

Knowledge formed from combinations of the experiential and innate is conventionally deemed *a posteriori* – even if heavily based on the latter. To be justifiably *a priori*, Kant (1787/1965) argued that knowledge must be absolutely independent of all experience – i.e. innate. But Kant (1783/2004) also deemed logical progression from perceived 'basic principles' to be *a priori* knowledge even though such principles might not necessarily be wholly obvious at a particular time. Event causation was, for Kant, a 'basic principle'; for Hume, an experiential phenomenon; for Peirce a complex of probabilistic and teleological phenomena (Chapter 3); and for physicists, entirely uncaused (e.g. describing the spontaneous, random appearance of particles). This progression suggests that the concept of justified *a priori* knowledge depends on the depth of human perception (true belief) at some point in time and risks subsequent defeat by progressive insight!

Thus rationalists such as Bonjour (1998) and Russell (2008) follow Kant by claiming that a proposition is justified *a priori* so long as there is no appeal to experience – other than to the extent that experience might be necessary purely to understand the proposition itself.

An *a priori* justification does not necessarily guarantee truth. If a proposition is based on a perceptively true (but actually false) proposition, *a priori* justification leads to false knowledge based on true belief. Thus a 'safer' path to *a priori* knowledge proceeds from 'truth assured' propositions – such as 'A = A', or 'All brothers are male'.

This Theory is founded on a primal axiom – a truth assured proposition of the form (A = ¬ B); explicitly (Alive = ¬ Dead). It is a truth assured proposition to say that an organisation must exist to survive and survive to exist. It is further reasoned that this must also be its *final* cause – since there can be no other. The Theory then proceeds to hold that survival requires continuous nourishments (permissions and other resources,

etc) that are only provided in exchange for the satisfaction of their suppliers (i.e. the continuous creation of welfare).

Moving onto the other conditions raises the next question: is BT1 justified *a priori*? Peircean causation is a more difficult issue to contain within an analytical proposition. This causal framework asserts that *efficient* cause is subject to chance and there can be 'no *efficient* cause without *final* cause', thus appearing to combine both rational and experiential components. Peirce's example of the relationship between the Sheriff and the Law (Chapter 3.8.2) is illuminating. The Sheriff exemplifies e*fficient* cause but is directionless unless actions are based on a *final* cause - embodied in the Law. The actions of the ideal Sheriff are rational with respect to the Law (analytical propositions). Yet the Law is a developmental teleology, or a current teleological trajectory within human ontology. Its purpose is determined by legislators applying their experience and judgement to obtain states of affairs (a trajectory) they believe best suit their society's advancement and survival. Thus Peirce's Sheriff and Law example is a composition of justified *a priori* and *a posteriori* knowledge (see also Chapter 3.8.2). If generalised within an organisational framework it implies that recognition of the need to survive is *a priori* knowledge but its pursuit - any corresponding developmental teleologies or teleological trajectories - requires *a posteriori* knowledge since all possible paths to survival are subject to the experience and judgement of their creators. As discussed earlier (Chapter 5.5), there is no *a priori* optimum or even feasible path that obtains survival.

Implementation of the Theory in any organisational situation is a two-stage process that involves practitioners:

1. understanding the meaning of the theoretical conditions and
2. applying the theoretical conditions.

The first stage is a purely logical process that does not require experience beyond that required to understand the meaning of any nomenclature describing the Theory's conditions. The sufficient conditions for an effective benchmarking process then follow from the logical expression $BT1 \wedge BT2 \wedge (BT3 \vee BT4 \vee BT5)$ (Equation 7-1) which describes what must be done on an *a priori* basis.

The second stage describes how Practitioners apply the Theory. Application requires knowledge that can only be obtained from an intimate understanding of current organisational states of affairs – e.g. knowledge of internal systems and processes, their relationships with each other together with their contribution to organisational purpose – i.e. *a posteriori* knowledge.

Current empirical benchmarking frameworks have attracted criticism for being predominantly experiential, to such an extent that practitioner experience, rather than reliance on sets of rules, is cited as essential for success (Chapter 2.4.2).

On the basis of this brief analysis, the Theory is founded on a justified *a priori* proposition, one of its conditions (e.g. Peircean causation) appears to contain both justified *a priori* and *a posteriori* constructions; application of the Theory appears justified *a priori* in its explanation of an effective benchmarking process but *a posteriori* in its implementation. This contrasts with current benchmarking practices that are represented by frameworks that are entirely experiential - *a posteriori*.

8.5 Conclusion.

The original research question was provoked by the *a-theoretical* nature of benchmarking, evinced by inability of current benchmarking practices to distinguish successful from unsuccessful efforts. This Theory establishes a logical, causal relationship between an effective benchmarking process and the efforts of practitioners. Although it is also important for any proposed theory to rationalise and extend current paradigmatic frameworks as well as explain current epistemology, it is principally through causal structures that explanation and extension of current knowledge progresses.

This Chapter concludes that the Theory of Benchmarking developed in this thesis has been validated against current exemplary praxis (epistemology) – explaining their effectiveness and ineffectiveness. Validation of logical consistency was established in Chapter 6 and further extended in this Chapter to include the parts of the Theory that are *a priori* and *a posteriori*. The final validation demonstrated the Theory's ability to constructively co-exist within current organisational paradigms.

The results of these validations determine current benchmarking praxis- its definitions and multiplicity of implementation frameworks – to be simulations of this Theory's causal relata: supervenience, entailment, welfare improvement and teleology. Continuous improvement of these simulations have increased the complexity of current benchmarking frameworks, reduced their efficiency and unnecessarily multiplied their forms. This Theory only recognises a single form of benchmarking but the EBP (pragmatically) recognises that an external exemplar might incur administrative rather than theoretical effort to determine the feasibility of a particular benchmarking proposition.

This Theory of Benchmarking views organisations as assemblies of states of affairs, each having appropriate resources or environmental variables, whose behaviours represent trajectories continuously developed (whether consciously or not) in support of a *final* cause through satisfaction of the suppliers and maintainers of these resources. Rules that determine the anomalar's current trajectory in pursuit of the *final* cause may be compared with those of exemplars. Provided there is a nomological congruence (supervenience) between these rules there is an opportunity for effective benchmarking if it is also both feasible (entailment) and efficient (welfare improvement) to transfer exemplary characteristics to the anomalar's environmental variables.

Many other phenomena have been explained through validation of this Theory of Benchmarking:

- Excellence frameworks. The Malcolm Baldrige National Quality Award exemplifies welfare improvement by benchmarking organisational forms against exemplary dispositions.

- Simplicity. This is a more significant factor than size in obtaining effective benchmarking because supervenience and entailment relationships are easier to establish. This explains the effectiveness of internal benchmarking within SMEs, the use of consistent approaches within larger organisations and the wisdom of self-knowledge before knowledge of others.

- Strategic Intent. The necessity of 'strategic intent' within some current frameworks has been identified as a teleological relatum.

- Stakeholder Consultation. Inclusion of other stakeholder voices within current frameworks has been identified as adherence to the primal axiom, that of survival.

The Theory explains current criticisms of benchmarking and addresses the research question through the establishment of a sufficient condition that obtains effective benchmarking. While it is difficult to deny that experience is an important component of any organisational change programme, the conditions for effective benchmarking are logical, do not depend on practitioner experience and can be determined prior to the commencement of a proposed benchmarking programme.

Chapter 9. Conclusions.

9 What has been learned?

Pursuit of the original research question, '*What is the theoretical framework for benchmarking'*, has established a series of principal findings and a number of consequential contributions. This concluding chapter examines the significance of resolving the research question and consequential contributions to new definitions of benchmarking and states of affairs. Additionally, several practitioner tools have been produced together with a generic process establishing necessary conditions for validating *potentially theoretical* constructions.

There are also a number of ongoing research avenues arising from this Theory. The practical aspects of implementing the Theory within different organisations may give rise to comparative case studies. There appears to be potential to extend the methodologies used to establish a theory of benchmarking into a generic approach to organisational theorising and the causal nature of states of affairs.

9.1 Principal Findings.

Critics of benchmarking have identified many of its incongruities. These served to form the research question and motivate its progress. The Theory of Benchmarking developed in this thesis validates the hypothesis that the practice of benchmarking embodies causal elements, which if identified, might permit distinction between effective and ineffective effort. These findings conclude that the proliferation of benchmarking's numerous types, forms and frameworks have been attempts to simulate its causal engine.

This research concludes that enduring expectations of discovering the underlying nature of benchmarking through continued incremental refinement of its current empirical frameworks will be unfruitful.

In benchmarking terms, the appropriate exemplar for benchmarking theory is causation. Effective benchmarking is a transfer of welfare between two causal engines. These causal engines are formed from conjunctions of Aristotelian, Humean-Mill and Peircean causation and an ontological axiom that necessitates the continuous creation of this welfare. The locus of these engines is immaterial: what is material is the ability to resolve

their respective causal relata. The rules for supervenience and entailment provide this resolution.

Causation remains controversial. Is it a suitable basis for a theoretical platform? All theories are intrinsically controversial hence the pivotal test of any nascent theory is the acceptability of its outcomes to those dependent upon them. It is unlikely that thousands of years of debate into the nature of causation will suddenly precipitate in unanimity. Some hold causation to be an unnecessary construction; some accept its *efficient* nature and others concur but further recognise that people and organisations are also teleological.

Explanation of effective versus ineffective benchmarking is a principal outcome of this research. The Theory of Benchmarking developed in this thesis concludes that several essential theoretical conditions for effective benchmarking have reposed within its implementation frameworks. Their unexpressed theoretical consequences reduce to simulations of the causal engine: supervenience, entailment, welfare improvement and teleological conjunction. This explains why effective efforts are most evident in situations where there is organisational simplicity (e.g. SMEs). Strong supervenience and logical entailment, even between diverse operations, are hallmarks of simple, straight-forward organisational structures.

Another introductory observation was increasing epistemological fusion between benchmarking and its supporting methodologies. The appealing concept of benchmarking had, over the years, also become associated with an increasingly sophisticated array of methods and systems (e.g. DEA, AHP, PDCA frameworks, etc). This research has addressed the essence of benchmarking whilst demonstrating such methods and systems to be proper. They are tools to be used in the satisfaction of, rather than substitution for, its theoretical principles.

A significant conclusion also arises from Chapter 8. Distillation of ‘scores’ of extant benchmarking frameworks into a single exemplar has resulted in homogeneity of praxis rather than theoretical advancement. This distillation is also evidence that there is little reason to expect significant improvements to benchmarking under its current praxis-driven paradigm. The Theory of Benchmarking developed in this research bespeaks a new paradigm. It has been validated to be *potentially theoretic* in its ability to encompass

the necessary epistemological, causal and paradigmatic attributes of currently accepted organisational benchmarking practices. There are, however, no *a priori* sufficient conditions to progress *potentially theoretic* constructions and the future acceptability of the Theory of Benchmarking developed in this thesis lies in its contribution to a new epistemology that challenges and, hopefully displaces the extant, praxis-driven paradigm.

The degree to which benchmarking frameworks have multiplied is truly remarkable but it would be remiss to omit adding yet another: the first founded entirely on theoretical principles. The Effective Benchmarking Process (EBP) is a practitioner-orientated benchmarking tool based on conditions sufficient for its effectiveness. The EBP demonstrates that there is but a single form of benchmarking; that the locus of the exemplar is of administrative rather than theoretical interest and that the overall success of any benchmarking effort relies on self-knowledge. Organisations that cannot explain existing relationships between their states of affairs or those that cannot establish them on their teleological trajectory are unlikely to improve by looking elsewhere. Analogously, it makes little sense to seek directions from an external source without first having adequate knowledge of one's own vehicle, a reason for its use and a clear destination in mind.

9.2 Additional Findings.

A number of contributions were also developed to support and exemplify aspects of the Theory of Benchmarking. These included new definitions of benchmarking and states of affairs, a theory provenance process and practitioner tools.

9.2.1 Definitions of Benchmarking and States of Affairs.

A new high-level definition of benchmarking has been developed:

'Benchmarking is an exemplar-driven teleological process operating within an organisation with the objective of intentionally changing an anomalar's existing state of affairs into a superior state of affairs via the transformation of feasible exemplary relata'.

This definition combines theoretical and empirical provenance which advances current definitions because it explains what it is and how it achieves its purpose. Both are important practitioner concepts.

Organisational states of affairs have also been defined in terms of their contribution to this causal engine of benchmarking:

'States of affairs represent the status of organisational competencies at some point in time that can be gauged according to some consistent metric and establish the teleological trajectory or sustainability of an organisation.'

The importance of the teleological trajectory cannot be understated as it is also the basis of theoretical explanation for respected business excellence frameworks such as NZBEF, MBNQA, EFQM, etc.

9.2.2 A Theory Provenance or Validation Process.

Theory building is an important aspect of knowledge advancement but the actual process of theorising is understated. It is often re-stated less authoritatively as, for example, model or framework building. One objective of this research was to develop a theoretical framework. The Theory of Benchmarking is presented as far as it can be: as *potentially theoretical*, on the basis of avoidance of particular and known *a-theoretical* constructions. Advancement beyond this stage relies upon whether new epistemologies develop – or not, depending on the acceptability of this Theory's outcomes to scholars and practitioners in the benchmarking community. Whilst the 'theory filter' was primarily devised to validate this Theory, it appears to have broader potential application. The filter or validator presented in this research advances the epistemology of theorising by identifying *a-theoretical* constructions and, counterfactually, *potentially theoretical* constructions. It also raises the question as to whether theorising, as was found with benchmarking, might also benefit from the application of Occam's razor and be reduced to a single form based on the concepts of supervenience and entailment.

9.2.3 Practitioner Tools.

Appendix 1 illustrates the use of DEA as a tool to identify the *efficient* components of exemplary trajectories and potentially remedial strategies for other anomalars. There are

also numerous other tools that may assist with this task, but a precondition for their effective application is the existence of one of three possible supervenience and entailment relationships whose disjunction obtains theoretical sufficiency for effective benchmarking: i.e. $(\mathbf{BT3} \vee \mathbf{BT4} \vee \mathbf{BT5})$.

Already described, the EBP is also a practitioner tool, but its theoretical bias may need to be tempered with practitioner-orientated, administrative support.

9.3 Further learning.

Opportunities for extension and further learning are presented at two levels: theoretical and operational.

9.3.1 Improving the Causal Engine: Teleological Trajectories.

There is scope for further consideration of the nature of benchmarking's causal engine. For example, singular and probabilistic causation have been referred to as potential variations on Peircean causation. Furthermore, applications of Peircean causation to broader aspects of organisational behaviour may also advance the understanding of sustainability and in doing so further simplify linkages between exemplars and anomalars. Scenario analysis is an example of a teleological tool that might contribute to this task. There are both theoretical and operational aspects of such research.

9.3.2 Theorising: Paradigmatic Conjunction.

The apparent lack of a reliable method of assessing the projection or triangulation of a nascent theory onto a set of paradigms in satisfaction of the theory filter or validator prompts the question as to whether there might not be a better approach. The Theory of Benchmarking in this thesis has applied the concepts of supervenience and entailment to establish the feasibility of including particular relata in a benchmarking process. This raises the possibility of broadening the application of these concepts to re-visit the nature of paradigmatic conjunction. A better understanding of the theory of paradigmatic conjunction would strengthen this component of the theory filter and improve its reliability as a validator of a potentially theoretical construction.

9.3.3 Theory-based Case Studies of Effective Benchmarking.

Practitioner-orientated tools are valuable methods for advancing learning. The Effective Benchmarking Process emphasises theoretical necessities and currently lacks the abundant levels of administrative support found in current frameworks. The application

of the concepts of supervenience, entailment and welfare improvement in ‘marketplace’ situations may not always be straightforward and any tools or templates that simplify this process improve the effectiveness of the theory-based process. A case-study approach might achieve this objective and contribute to a new theory-based epistemology of benchmarking. This future research theme is predominantly operational.

9.4 Finale.

The *final* cause of this thesis is the satisfaction of its readers and its author. It is source of satisfaction and challenge to divine the extent organisational study draws upon such diverse disciplines of scholarship; both ancient and modern.

References.

Adorno, T. W. (2001). *Kant's 'Critique of Pure Reason'*. Palo Alto, CA: Stanford University Press.

Ahmed, P. K., & Rafiq, M. (1998). Integrated benchmarking: a holistic examination of select techniques for benchmarking analysis. *Benchmarking for Quality Management & Technology, 5*(3), 225-242.

Alstete, J. W. (2008). Measurement benchmarks or "real" benchmarking?: An examination of current perspectives. *Benchmarking: An International Journal, 15*(2), 178-186.

Anand, G., & Kodali, R. (2008). Benchmarking the benchmarking models. *Benchmarking: An International Journal, 15*(3), 257-291.

Andersen, B., & Moen, R. M. (1999). Integrating benchmarking and poor quality cost measurement for assisting the quality management work. *Benchmarking: An International Journal, 6*(4), 291-301.

Anderson, A., & McAdam, R. (2004). A critique of benchmarking and performance measurement. *Benchmarking: An International Journal, 11*(5), 465-483.

Andrews, K. R. (1980). *Concept of Corporate Strategy*. Homeland, IL: Irwin.

Aquinas, T. S. (1258/1957). Summa Contra Gentiles. In O. P. Joseph Kenny (Ed.), *Online English Translation http://www.diafrica.org/kenny/CDtexts/ContraGentiles.htm*. New York: Hanover House.

Argyris, C. (1977). Organizational learning and management information systems. *Accounting, Organizations and Society, 2*(2), 113-123.

Argyris, C. (1999). *On Organizational Learning*: Blackwell Publishers.

Aristotle. (1984). The Complete Works of Aristotle - the revised Oxford Translation. In J. Barnes (Ed.), (Vol. 1). New Jersey: Princeton University Press.

Armstrong, D. M. (1969). Dispositions Are Causes. *Analysis, 30*(1), 23-26.

Arrow, K., Dasgupta, P., Goulder, L., Daily, G., Ehrlich, P., Heal, G., *et al.* (2004). Are We Consuming Too Much? *Journal of Economic Perspectives, 18*(3), 147-172.

Bacon, F. (1620/2000). The New Organon (pp. 105). West Nyack, NY, USA: Cambridge University Press.

Banker , R., Cooper, W., Sieford, L., Thrall, R. M., & Zhu, J. (2004). Returns to Scale in Different DEA Models. *EJOR, 154*, 345-362.

Banker, R. D., Charnes, A. C., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science, 30*(9), 1078-1092.

Banker, R. D., & Thrall, R. M. (1992). Estimation of returns to scale using data envelopment analysis. *European Journal of Operational Research, 62*(1), 74-84.

Baumol, W. J. (2002). *The Free-Market Innovation Machine: Analyzing the Growth Miracle of Capitalism*. Princeton, NJ: Princeton University Press.

Belton, V., & Vickers, S. P. (1993). Demystifying DEA - A Visual Interactive Approach based on Multiple Criteria Analysis. *Journal Operational Research Society, 44*(9), 885.

Bhutta, K. S., & Huq, F. (1999). Benchmarking best practices: an integrated approach. *Benchmarking: An International Journal, 6*(3), 254-268.

Bjurek, H. (1996). The Malmquist Total Factor Productivity Index. *Scandinavian Journal of Economics, 98*(2), 303-313.

Blackburn, S. (2007). Supervenience Revisited. In R. Shafer-Landau & T. Cuneo (Eds.), *Foundations of Ethics: An Anthology* (pp. 439ff). Malden: Blackwell Publishing.

BNQP. (2008). Criteria for Performance Excellence (pp. iv): Department of Commerce: National Institute of Standards and Technology.

BonJour, L. (1998). *In Defense of Pure Reason: A Rationalist Account of A Priori Justification*: Cambridge University Press.

Bourgeois, L. J. (1979). Toward a method of middle-range theorizing. *Academy of Management Review, 4*(3), 443-447.
Bourne, M., Mills, J., Wilcox, M., Neely, A., & Platts, K. (2000). Designing, implementing and updating performance measurement systems. *International Journal of Operations & Production Management, 20*(7), 754-771.
Boyer, E. L. (1990). *Scholarship reconsidered: Priorities of the Professoriate*. Princeton, NJ: Carnegie Foundation for the Advancement of Teaching
Bunge, M. (1959). *Causality - The place of the causal principle in modern science*. Cambridge, MA: Harvard University Press.
Burrell, G., & Morgan, G. (1979). *Sociological paradigms and organisational analysis: elements of the sociology of corporate life*. London: Heinemann
Büyüközkan, G., & Maire, J. L. (1998). Benchmarking process formalization and a case study. *Benchmarking for Quality Management & Technology, 5*(2), 101-125.
Camp, R. C. (1989). *Benchmarking: The Search for Industry Best Practices that Lead to Superior Performance*. Milwaukee,WI: ASQC Quality Press.
Campbell, A. (1999). Tailored, not benchmarked. A fresh look at corporate planning. *Harvard Business Review, 77*(2), 41-50.
Carpinetti, L. C. R., & de Melo, A. M. (2002). What to benchmark? *Benchmarking: An International Journal, 9*(3), 244-255.
Cartwright, N. (2007). *Hunting Causes and Using Them*. Cambridge: Cambridge University Press.
Cassell, C., Nadin, S., & Gray, M. O. (2001). The use and effectiveness of benchmarking in SMEs. *Benchmarking: An International Journal, 8*(3), 212-222.
Cefis, E., & Marsili, O. (2005). A matter of life and death: innovation and firm survival. *Industrial and Corporate Change, 14*(6), 1167-1192.
Chappell, V. (Ed.). (1999). *Hobbes & Bramhall on Liberty & Necessity*. Port Chester, NY: Cambridge University Press.
Charnes, A. C., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*(6), 429-444.
Clarke, A., & Manton, S. (1997). A benchmarking tool for change management. *Benchmarking: An International Journal, 3*(3), 248-255.
Collins, T. R., Rossetti, M. D., Nachtmann, H. L., & Oldham, J. R. (2006). The use of multi-attribute utility theory to determine the overall best-in-class performer in a benchmarking study. *Benchmarking: An International Journal, 13*(4), 431-446.
Cooper, W. W., Seiford, L. M., & Zhu, J. (2004). *Handbook on Data Envelopment Analysis*. London: Kluwer Academic Publishers.
Cox, S. E. (2003). *A Theory of Benchmarking with Applications to Software Reverse Engineering.* University of Toronto, Toronto.
Dattakumar, R., & Jagadeesh, R. (2003). A review of literature on benchmarking. *Benchmarking: An International Journal, 10*(3), 176-209.
Davidson, D. (1970/2001). Mental Events. In D. Davidson (Ed.), *Essays on Actions and Events* (Vol. Philosophical Essays 1, pp. 209ff). Oxford: Oxford University Press.
Davidson, D. (1980). *Essays on Actions and Events*. Oxford: Clarendon Press.
Davidson, D. (1993). Causal Relations. In E. Sosa & M. Tooley (Eds.), *Causation* (pp. 81). Oxford: Oxford University Press.
Davis, B. (1993). *The thought of Thomas Aquinas*. Oxford: Oxford University Press.
Delahunty, R. J. (1985/1999). *Spinoza, The Arguments of the Philosophers*. London: Routledge.
Deming, W. E. (1982). *Quality, productivity, and competitive position*: Massachusetts Institute of Technology, Center for Advanced Engineering Study, Cambridge, MA.
Deming, W. E. (1986). Out ofthe Crisis. *MIT Center for Advanced Engineering Study, Cambridge, MA.*

Deros, B. M., Yusof, S. M., & Salleh, A. M. (2006). A benchmarking implementation framework for automotive manufacturing SMEs. *Benchmarking: An International Journal, 13*(4), 396-430.
Diogenes_Laertius. (1853). In C. D. Yonge (Ed.), *The Lives and Opinions of Eminent Philosophers* (pp., Paragraph XII). London: Henry G. Bohn.
Dowling, J., & Pfeffer, J. (1975). Organizational legitimacy: Social values and organizational behavior. *Pacific Sociological Review, 18*(1), 122-136.
Drew, S. A. W. (1997). From knowledge to action: the impact of benchmarking on organizational performance. *Long Range Planning, 30*(3), 427-441.
Dubin, R. (1976). Theory building in applied areas. *Handbook of industrial and organizational psychology*, 17-39.
Ducasse, C. J. (1926). On the Nature and Observability of the Causal Relation. *Journal of Philosophy, 23*(3), 125–136.
Ducasse, C. J. (1969). *Causation and the types of necessity*. New York: Dover Publications.
Easterby-Smith, M., Burgoyne, J., & Araujo, L. (1999). *Organizational learning and the learning organization* Sage: London.
Elmuti, D., & Kathawala, Y. (1997). An overview of benchmarking process: a tool for continuous improvement and competitive advantage. *Benchmarking for Quality Management & Technology, 4*(4), 229-243.
Elnathan, D., Lin, T. W., & Young, S. M. (1996). Benchmarking and management accounting: a framework for research. *Journal of Management Accounting Research, 8*(2), 37-54.
Elster, J. (1983). *Explaining technical change*. Cambridge: Cambridge University Press.
Fara, M. (2006). Dispositions. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2006 ed.).
Färe, R., Grosskopf, S., Kirkley, J. E., & D, S. (2000). *Data Envelopment Analysis (DEA): A Framework for Assessing Capacity in Fisheries When Data are Limited.* Paper presented at the Proceedings of the Tenth Biennial Conference of the International Institute of Fisheries Economics and Trade, Corvallis, OR.
Farrell, M. J. (1957). The Measurement of Productive Efficiency. *Royal Statistical Society, A*(120), 253-281.
Francis, G., & Holloway, J. (2007). What have we learned? Themes from the literature on best-practice benchmarking. *International Journal of Management Reviews, 9*(3), 171-189.
Frankl, V. E. (1992). *Man's Search for Meaning* (4 ed.). Boston, MA: Beacon Press.
Friedman, M. (2002). *Capitalism and Freedom* (40th Anniversary Edition ed.). Chicago: University of Chicago Press.
Gabbey, A. (2002). The Cambridge Companion to Newton. In B. Cohen (Ed.), (pp. 336). West Nyack, NY: Cambridge University Press.
Garvin, D. A. (1993). Building a learning organization. *Harvard Business Review, 71*(4), 78-91.
Gaukroger, S. (2002a). *Descartes' Natural Philosophy*. London: Routledge.
Gaukroger, S. (2002b). *Descartes' System of Natural Philosophy*. West Nyack: Cambridge University Press.
Ghoshal, S. (2005). Bad Management Theories Are Destroying Good Management Practices. *The Academy of Management Learning and Education, 4*(1), 75-91.
Gioia, D. A., & Pitre, E. (1990). Multiparadigm perspectives on theory building. *Academy of Management Review, 15*(4), 584-602.
Han, J. K., Kim, N., & Srivastava, R. K. (1998). Market Orientation and Organizational Performance: Is Innovation a Missing Link? *Journal of Marketing, 62*(4), 30-45.
Harrington, H. J. (1995). The new model for improvement: total improvement management. *Management Decision, 33*(3), 17-24.
Hawthorne, J. (2004). Inductive Logic. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2008 ed.).

Hobbes, T., Lessay, F. (1688/2000). Historica Ecclesiastica. In G. A. Rogers & T. R. Sorell (Eds.), *Hobbes and History* (Vol. ebrary, Inc, pp. 151, 155, 156). London: Routledge.
Hodgson, P. E. (2003). Galileo the Scientist. *Logos: A Journal of Catholic Thought and Culture, 6*(3), 13-40.
Huber, G. P. (1991). Organizational learning: The contributing processes and the literatures. *Organization Science, 2*(1), 88-115.
Hulswit, M. (2000). Peirce on Causality and Causation. In J. Queiroz & R. Gudwin (Eds.), *Digital Encyclopaedia of C.S. Peirce* (pp. III.I): http://www.digitalpeirce.fee.unicamp.br.
Hulswit, M. (2002). *From Cause to Causation: A Peircean Perspective* (Vol. 90). Dordrech: Kluwer Academic Publishers.
Hume, D. (1739/1999). *A TREATISE OF Human Nature: BEING An ATTEMPT to introduce the experimental Method of Reasoning INTO MORAL SUBJECTS*. Kitchener, Ontario: Batoche Books.
Hume, D. (1748/2000). *An Enquiry Concerning Human Understanding*. Kitchener, Ontario, Canada: Batoche Books.
Huq, F., Abbo, M.-H., & Huq, Z. (2008). Perceptions about benchmarking best practices among French managers: an exploratory survey. *Benchmarking: An International Journal, 15*(4), 382-401.
Jackson, A. E., Safford, R. R., & Swart, W. W. (1994). Roadmap to Current Benchmarking Literature. *Journal of Management in Engineering, 10*(6), 60-67.
Juran, J. M. (1950). The role of statistics as a tool of management*. *Statistica Neerlandica, 4*(1-2), 69-79.
Kant, I. (1783/2004). Prolegomena to Any Future Metaphysics that Will be Able to Come Forward as Science: With Selections from the Critique of Pure Reason. In G. Hatfield (Ed.), (pp. 40). West Nyack, NY: Cambridge University Press.
Kant, I. (1787/1965). *Critique of Pure Reason, trans. Norman Kemp Smith*. New York: St Martin's Press.
Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard--measures that drive performance. *Harvard Business Review, 70*(1), 71-79.
Kearns, D. T., & Nadler, D. A. (1992). *Prophets in the Dark: How Xerox Reinvented Itself and Beat Back the Japanese*. New York: Harper Collins.
Kenny, A. (2002). *Aquinas on Being*. Oxford: Oxford University Press.
Kim, J. (1973). Causes and Counterfactuals. *Journal of Philosophy, 70*(17), 570-572.
Kim, J. (1984). Concepts of Supervenience. *Philosophy and Phenomenological Research, 45*(2), 153–176.
Kozak, M., & Nield, K. (2001). An Overview of Benchmarking Literature: Its Strengths and Weaknesses. *Journal of Quality Assurance in Hospitality & Tourism, 2*(3/4), 7-32.
Kuhn, T. S. (1996). The Structure of Scientific Revolutions (ed.). *Chicago: The University of Chicago Press, 152*, 378.
Kumar, S., & Chandra, C. (2001). Enhancing the effectiveness of benchmarking in manufacturing organizations. *Industrial Management and Data Systems, 100*(1), 80-89.
Kyrö, P. (2003). Revising the concept and forms of benchmarking. *Benchmarking: An International Journal, 10*(3), 210-225.
Lancaster, K. J. (1966). A New Approach to Consumer Theory. *The Journal of Political Economy, 74*(2), 132.
Leibfried, K. H. J., & McNair, C. J. (1992). *Benchmarking: A tool for continuous improvement*. Essex Junction, VT: Oliver Wight Publications Inc, with arrangement from Harper Collins Publishers Inc.
Levinthal, D. (1992). Surviving Schumpeterian Environments: An Evolutionary Perspective. *Industrial and Corporate Change, 1*(3), 427-443.
Lewis, D. (1973). Causation. *Journal of Philosophy, 70*(17), 556-567.

Lewis, M. W., & Grimes, A. J. (1999). Metatriangulation: Building theory from multiple paradigms. *Academy of Management Review, 24*(4), 672-690.
Liang, Y. (2004). *Benchmarking Theory Applied to State Parks: an exploratory study.* Oklahoma State University.
Llewelyn, S. (2003). What counts as "theory" in qualitative management and accounting research? Introducing five levels of theorizing. *Accounting, Auditing & Accountability Journal, 16*(4), 662.
Locke, J. (1690/2001). *An Essay Concerning Human Understanding*. Kitchener, Ontario, Canada: Batoche Books.
Mackie, J. L. (1973). *Truth, Probability and Paradox: Studies in Philosophical Logic*: Oxford University Press.
Mackie, J. L. (1975). Causes and Conditions. In E. Sosa (Ed.), *Causation and Conditionals* (pp., p62). London: Oxford University Press.
Maire, J.-L., Bronet, V., & Pillet, M. (2005). A typology of "best practices" for a benchmarking process. *Benchmarking: An International Journal, 12*(1), 45-60.
Mandik, P. (2004). Supervenience. In C. Eliasmith (Ed.), *Dictionary of Philosophy of Mind. URL=<http://philosophy.uwaterloo.ca/MindDict/supervenience.html>*.
Mann, R., & Voss, M. (2000). An innovative process improvement approach that integrates ISO 9000 with the Baldrige framework. *Benchmarking: An International Journal, 7*(2), 128-145.
McAdam, R., & Kelly, M. (2002). A business excellence approach to generic benchmarking in SMEs. *Benchmarking: An International Journal, 9*(1), 7-27.
McAdam, R., & McCreedy, S. (1999). A critical review of knowledge management models. *The Learning Organization: An International Journal, 6*(3), 91-101.
McLaughlin, B., & Bennett, K. (2005). Supervenience. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (*Fall 2006,* http://plato.stanford.edu/archives/fall2006/entries/supervenience. ed.).
Merton, R. K. (1968). *Social Theory and Social Structure, 1968 Enlarged Edition*. New York/London: Free Press.
Merz, J. T. (1884/2001). *Leibniz*. Kitchener, Ontario, Canada: Batoche Books.
Mill, J. S. (1872/1973). A System of Logic Ratiocinative and Inductive. In J. Robson (Ed.), *Collected Works of John Stuart Mill, Vol. VII* (8th ed., pp. Book III, Chapter v, §2, p326). Toronto: University of Toronto Press.
Montagnes, B. (2001). *Doctrine of the Analogy of Being According to Saint Thomas Aquinas*. Milwaukee, WI, USA: Marquette University Press.
Morgan, G. (1980). Paradigms, metaphors and puzzle solving in organization theory. *Administrative Science Quarterly, 25*(4), 605-622.
Morgan, G. (1983). More on Metaphor: Why we Cannot Control Tropes in Administrative Science. *Administrative Science Quarterly, 28*(4), 601-607.
Moriarty, J. P. (2007). *Division Benchmarks for New Zealand Tourism and Characteristic and Related Industries 1999-2003. Yield Report* (No. 6a). Canterbury, NZ: TRREC, Lincoln University.
Moriarty, J. P. (2008). Challenges for the Hospitality Industry: Beyond Price. In J. Fountain (Ed.), *New Zealand Tourism & Hospitality Research Conference 2008*. Hanmer, Canterbury, New Zealand [in print].
Moriarty, J. P., & Smallman, C. (2009). En-Route to a Theory of Benchmarking *Benchmarking: An International Journal, 16*(3).
Nadler, S. (2002). *Spinoza's 'Ethics'.: An Introduction*. West Nyack, NY, USA: Cambridge University Press.
NIST, N. I. o. S. (2007). *Malcolm Baldrige National Quality Awards, Scoring Guidelines, Consensus* Retrieved 15 December, 2006, from

http://www.quality.nist.gov/Word_files/2007%20Consensus%20Materials/Consensus_Toolkit_Final.doc, P9
NZBEF. (2007). *New Zealand Business Excellence Framework*. Retrieved 10 March, 2008, from http://www.nzbef.org.nz/about_framework.htm
Partovi, F. Y. (1994). Determining what to benchmark: an analytic hierarchy process approach. *International Journal of Operations & Production Management, 14*(6), 25-39.
Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*: Cambridge University Press.
Peirce, C. S. (1898/1992). Reasoning and Logic of Things. In K. L. Ketner (Ed.) (pp. p197). Cambridge, MA: Harvard University Press.
Peirce, C. S. (1935). Collected papers of Charles Sanders Peirce. v VI Exact Logic. In C. Hartshorne & P. Weiss (Eds.). Cambridge, MA: Harvard University Press.
Peirce, C. S. (1998). The Essential Peirce: Selected Philosophical Writings. In N. Houser & C. J. W. Kloesel (Eds.). Bloomington: Indiana University Press.
Pfeffer, J. (1997). *New Directions for Organization Theory: Problems and Prospects*: Oxford University Press US.
Pfeffer, J., & Salancik, G. R. (2003). The External Control of Organizations: A Resource Dependence Perspective In M. J. Handel (Ed.), *The Sociology of Organizations: Classic, Contemporary, and Critical Readings* (pp. 234). Thousand Oaks: Sage Publications Inc.
Pinder, C. C., & Bourgeois, V. W. (1982). Controlling Tropes in Administrative Science. *Administrative Science Quarterly, 27*(4), 641-652.
Plato. (1793). Plato: The Cratylus, Phædo, Parmenides and Timæus of Plato. With notes on the Cratylus, and an explanatory introduction to each dialogue. In T. Taylor (Ed.), *Eighteenth Century Collections Online* (pp. 455). London: Gale Group.
Poole, M. S., & Van de Ven, A. H. (1989). Using paradox to build management and organization theories. *Academy of Management Review, 14*(4), 562-578.
Popper, K. (1987). The Myth of the Framework. In J. C. Pitt & M. Pera (Eds.), *Rational Changes in Science: Essays on Scientific Reasoning* (pp. 34 ff): Springer.
Popper, K. R. (1996). *The Myth of the Framework: In Defence of Science and Rationality*. London: Routledge.
Porter, L. J. (2004). *Assessing Business Excellence: A Guide to Business Excellence and Self-assessment*: Butterworth-Heinemann.
Psillos, S. (2002). *Causation and Explanation*. Montreal & Kingston: McGill-Queens University Press.
Punniyamoorthy, M., & Murali, R. (2008). Balanced score for the balanced scorecard: a benchmarking tool. *Benchmarking: An International Journal, 15*(4), 420-443.
Radnor, Z., & Robinson, J. (2000). Benchmarking innovation. *Creativity and Innovation Management, 9*(2), 3–13.
Ragavan, P., & Punniyamoorthy, M. (2003). A Strategic Decision Model for the Justification of Technology Selection. *The International Journal of Advanced Manufacturing Technology, 21*(1), 72-78.
Roxbee_Cox, J. W. (1975). Mackie On Dispositional Properties. *The British Journal for the Philosophy of Science, 26*(3), 232-234.
Rozeboom, W. W. (1973). Dispositions Revisited. *Philosophy of Science, 40*(1), 59-74.
Russell, B. (1979). *History of Western Philosophy* (2 ed.). London: Unwin Paperbacks.
Russell, B. (2008). *A Priori* Justification and Knowledge. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall ed.).
Shewhart, W. A. (1980). *Economic Control of Quality of Manufactured Product*. Milwaukee, WI: American Society for Quality Control.
Singh, J. V., Tucker, D. J., & House, R. J. (1986). Organizational legitimacy and the liability of newness. *Administrative Science Quarterly, 31*(2), 171-193.
Skorupski, J. (1998). *The Cambridge Companion to Mill*: Cambridge University Press.

Smallman, C. (2007, July 5-7). *Bourdieu's Cat: On Data Indeterminancy in the Study of Organisational Practice.* Paper presented at the 23rd European Group for Organisational Studies Colloquium, Wirtschaftsuniversität Wien, Vienna, Austria, p6.
Smith, P. (1997). Model misspecification in Data Envelopment Analysis. *Annals of Operations Research, 73*, 233-252.
Sommers, T. (2003). *The buck stops - Where? Interview: Galen Stawson (The Believer).* Retrieved 13th August, 2008, from http://www.believermag.com/issues/200303/?read=interview_strawson
Sosa, E., & Tooley, M. (1993). *Causation*. Oxford; New York: Oxford University Press.
Southard, P., B , & Parente, D., H (2007). A model for internal benchmarking: when and how? *Benchmarking: An International Journal, 14*(2), 161-171.
Spendolini, M. J. (1992). The Benchmarking Process. *Compensation & Benefits Review, 24*(5), 21.
Stalnaker, R. (1996). Varieties of Supervenience. *Philosophical perspectives, 10*, 221-241.
Stavins, R. N., Wagner, A. F., & Wagner, G. (2003). Interpreting sustainability in economic terms: dynamic efficiency plus intergenerational equity. *Economics Letters, 79*, 339-343.
Stewart, G. B. (1991). *The Quest for Value: the EVA Management Guide*. NY: Harper Collins.
Sutton, R. I., & Staw, B. M. (1995). What Theory Is Not. *Administrative Science Quarterly, 40*(3), 371-384.
Tooley, M. (1987). *Causation: A Realist Approach*. Oxford: Clarendon Press.
Tooley, M. (1993). Causation. In E. Sosa & M. Tooley (Eds.), *Causation* (pp. 172). Oxford: Oxford University Press.
Van de Ven, A. H., & Poole, M. S. (1995). Explaining development and change in organizations. *Academy of Management Review, 20*(3), 510-540.
Van de Ven, A. H., & Poole, M. S. (2005). Alternative Approaches for Studying Organizational Change. *Organization Studies, 26*(9), 1377.
van Helden, G. J., & Tillema, S. (2005). Search of a Benchmarking Theory for the Public Sector. *Financial Accountability & Management, 21*(3), 337–361.
Vaziri, H. K. (1992). Using competitive benchmarking to set goals. *Quality Progress, 25*(10), 81-85.
Wallace, W. (2000). *Kant*. Kitchener, Ontario, Canada: Batoche Books.
Watson, G. H. (1993). *Strategic Benchmarking: How to Rate Your Company's Performance Against the World's Best*. New York, NY: Wiley.
Watson, G. H. (1994). A Perspective on Benchmarking: Gregory H. Watson in conversation with the Editor. *Benchmarking: An International Journal, 1*(1), 5-10.
Weick, K. E. (1989). Theory construction as disciplined imagination. *Academy of Management Review, 14*(4), 516-531.
Weick, K. E. (1995). What theory is not, theorizing is. *Administrative Science Quarterly, 40*(3), 385-390.
Weissman, D. (1965). *Dispositional Properties*. Carbondale and Edwardsville: Southern Illinois University Press.
Welch, S., & Mann, R. (2001). The development of a benchmarking and performance improvement resource. *Benchmarking: An International Journal, 8*(5), 431-452.
Willmott, H. (1993). Breaking the Paradigm Mentality. *Organization Studies, 14*(5), 681.
Wöber, K. W. (2002). *Benchmarking in Tourism and Hospitality Industries: The Selection of Benchmarking Partners*. Wallingford: CABI Publishing.
Wolfram_Cox, J. R., Mann, L., & Samson, D. (1979). Benchmarking as a mixed metaphor: Disentangling Assumptions of Competition and Collaboration. *Journal of Management Studies, 34*(2), 285-314.
Woolhouse, R. (1993). Descartes, Spinoza, Leibniz: The Concept of Substance in Seventeenth-Century Metaphysics, *The Review of Metaphysics*: Routledge.

Xerox. (1979). Benchmarking, *Internal Publication*. London: Xerox Corporation.
Yasin, M. M. (2002). The theory and practice of benchmarking: then and now. *Benchmarking: An International Journal, 9*(3), 217-243.
Zairi, M. (1992). The art of benchmarking: using customer feedback to establish a performance gap. *Total Quality Management & Business Excellence, 3*(2), 177-188.
Zairi, M. (1994). *Measuring Performance for Business Results*: Springer.
Zairi, M. (1996). *Effective Benchmarking: Learning from the Best*: Chapman & Hall.
Zairi, M. (1997). Benchmarking: Towards being an accepted management tool or is it on its way out? *Total Quality Management, 8*(2-3), 337-338.
Zairi, M., & Ahmed, P. K. (1999). Benchmarking maturity as we approach the millennium? *Total Quality Management, 10*(4-5), 810-816.
Zairi, M., & Baidoun, S. (2003). Understanding The Essentials of Total Quality Management: A Best Practice Approach – Part 1, *Working Paper Series: No 03/05*. Bradford: University of Bradford.
Zairi, M., & Leonard, P. (1994). *Practical Benchmarking: The Complete Guide*: Chapman & Hall.

Appendix 1. An Algorithm for Establishing the Efficient Production Frontier from a set of Technically Efficient Exemplars.

1. Technical Efficient Exemplars.

If a benchmarking exercise establishes a set of exemplars, it is unlikely that their operating parameters will align precisely with those of the anomalar. The most pragmatic improvement available to the anomalar, if it is not already included in the set of data that established the exemplars, is to position itself as best it can on the efficient frontier that describes the dominating technology employed by these enterprises. If the anomalar is included in the benchmarking dataset, methodologies such as DEA will have established the optimal parameter set that position the anomalar on the efficient frontier. DEA, for example, achieves this by assigning linear combinations of exemplary inputs to re-establish each of the anomalar's inputs. Alternatively, analysis of slack variables identifies shifts in input levels that need to occur to achieve the efficient frontier.

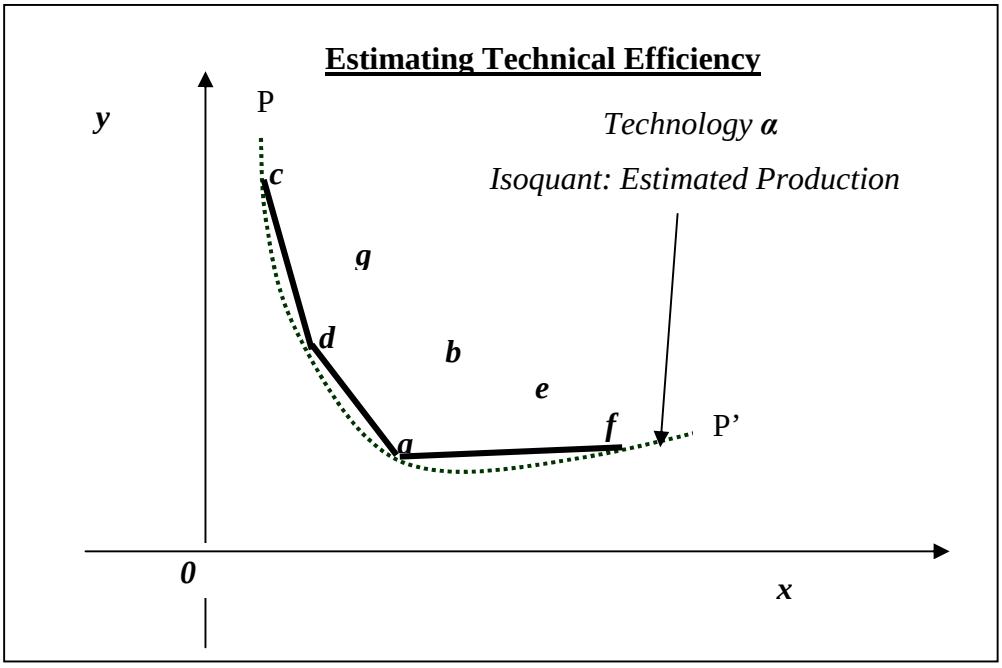

Figure A1- 1. Estimating Technical Efficiency.

In the Figure A1-1, the operating parameters (resources *y* & *x*) for exemplars {a, c, d, f} are already known. Suppose an anomalar learns of these exemplars and wishes to establish its own best operating practice purely on the basis of this knowledge.

For a simple two dimensional example such as above, a production frontier piece or segment is a line of the form $y = \alpha.x + \beta$, where α is the slope and β the intercept on the y (output) axis. In the case of multiple inputs and outputs, parameters are matrices representing a hyper-plane in as many dimensions as are formed by the sum of inputs and outputs.

2. Determining the Efficient Frontier.

If we assume linear combinations of exemplar variables (as assumed by DEA), we may say that the exemplar set of r members, each of which has n outputs (y) and m inputs (x) may be described as a vector equation, E of the form

$$E = Outputs - Inputs - \mu = 0 \text{ , where}$$

$$Outputs = \sum_{i=1}^{n} u_i.y_{ir} \ ,$$

$$Inputs \ = \sum_{j=1}^{m} v_j.x_{jr}$$

Equation A1- 1. Frontier System of Equations

Where u and v are the parametric coefficients of respective outputs and inputs and parameter μ is a scalar representing returns to scale (RTS) behaviour for exemplars on the efficient frontier. For example: μ equals zero indicates constant RTS, μ negative indicates increasing RTS and μ positive indicates decreasing RTS (Banker & Thrall, 1992, p 79).

In the multiplier model of Data Envelopment Analysis, the task is to maximise the output ($u.y$) of the r^{th} exemplar for a given set of inputs ($v.x$) as follows:

$$\text{Max } \sum_{i=1}^{n} u_i.y_{ir} - \mu_r \text{ (Outputs)}$$

Equation A1- 2. Envelopment Form of Solution

Subject to the solution lying on the efficient frontier and the proportion of positive input quantities needed to achieve this not exceeding 100% – i.e.

Subject to

$$\sum_{i=1}^{n} u_i.y_{ir} - \sum_{j=1}^{m} v_j.x_{jr} - \mu_r \leq 0$$

$$\sum_{j=1}^{m} v_j.x_{jr} = 1$$

$u_i \ , \ v_j \geq \boldsymbol{\varepsilon}$**,** a very small positive real number

Equation A1- 3. Linear Program for Efficient Frontier

The solution for which is a hyper-plane in the form E, above describing the efficient frontier at the r^{th} exemplar's location. These equations are known as the BCC multiplier form of

DEA extended to address variable returns to scale by the addition of μ_r (Banker *et al.*, 2004). A linear programming solver (e.g. MathCAD™ or GAMS™) or specialised DEA solver (e.g. DEA Solver) may be used to process this problem.

Solutions to this formulation are representations of the relative technical efficiency of exemplars (generally 100%), equations of the hyper-planes forming the efficient frontier at exemplar nodes and an indication of the scale behaviour of each exemplar. Sorting the solutions in terms of ascending μ provides the piece-wise efficient frontier from IRS through CRS to DRS behaviours (Respectively, Increasing Returns to Scale, Constant Returns to Scale and Decreasing Returns to Scale).

If anomalars are included in the dataset, their relative technical efficiencies will be less than 100% and each will have a theoretical improvement policy calculated to re-position it onto the efficient frontier.

Careful attention needs to be paid to the nature of any improvement policy as its feasibility in practice is a function of the modelling assumptions and veracity of the data. Note that DEA treats all error or uncertainty as a contributor to technical efficiency. An example of this would be an assumption that all inputs can be re-sized without restraint in sympathy with an improvement policy. In practice the capacities of some inputs, such as capital assets, may be inconsistent with such an assumption and the model requires re-configuration to recognise this. Moriarty (2008) provides an example of benchmarking hospitality organisations under capacity constraints within the New Zealand Tourism Sector using Färe's (2000) capacity constrained DEA approach.

If anomalars are not included in the dataset, efficient frontier equations form a basis for improvement. Key anomalar attributes most closely aligned with a comparable exemplar serve to indicate which equation applies. An example might be the use of "revenues" or "customer volumes" as a comparability factor to assist with selection of the appropriate equation.

3. Determining the Piecewise Frontier: A Representative Example.

The single input, single output example shown in Figure A1-2 reflects the performance of states of affairs from 12 enterprises, some of which are potentially exemplary:

data :=

"Enterprise"	"Output"	"Input"
1	2.6	3.8
2	4	6
3	1	2
4	4	7
5	3.2	6.5
6	2.9	5
7	1.7	2.8
8	3	3
9	3.5	4
10	2	2.5
11	2	6
12	2	5

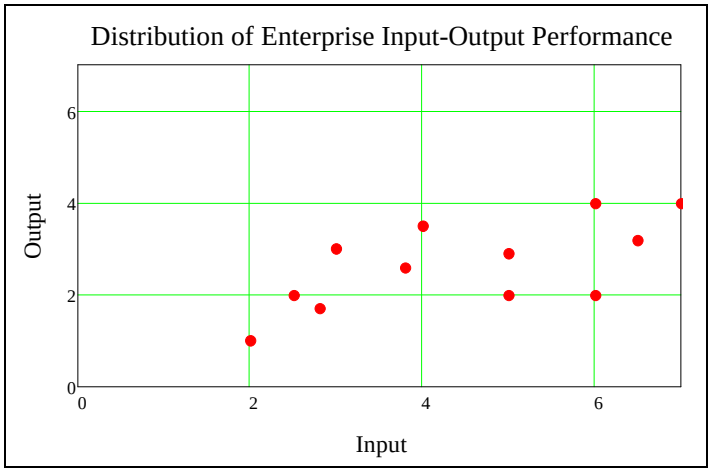

Figure A1- 2. Potentially Exemplary States of Affairs.

The optimum solution, derived from a linear programming solution of the "BCC" based on Equation A1-3 is shown in Table A1-1 below:

Enterprise	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
u	0.263	**1/6**	½	0.143	0.154	0.2	0.357	⅓	¼	0.4	0.167	0.2
v	0.132	⅔	¼	0.571	0.308	0.1	0.179	⅓	½	0.2	0.083	0.1
μ	-0.395	**1⅔**	**-¾**	1.429	0.462	-0.3	-0.536	**0**	¾	-0.6	-0.25	-0.3
Efficiency	**0.74**	**1**	**1**	**0.86**	**0.52**	**0.59**	**0.84**	**1**	**1**	**1**	**0.42**	**0.5**

Table A1- 1. DEA Model Optimum Solution for Potential Exemplars.

Enterprises 2, 3, 8, 9 and 10 are exemplars with relative technical efficiencies of 100%. Note that the normalised equations (with respect to u_r) describing the hyper-planes (straight-lines in this example) are given in Table A1-2 in the form $y_r - (v_r / u_r).\ x - \mu_r / u_r = 0$:

Enterprise	Output (*y*)	Input ($-\boldsymbol{v}_r / \mathbf{u}_r$)	$-\mu_r / u_r$ (μ')	Scale
2	1	0.25	2.5	$\mu' > 0$, Dec RTS
3	1	2	-3	$\mu' < 0$, Inc RTS
8	1	1	0	$\mu' = 0$, Const RTS
9	1	0.5	1.5	$\mu' < 0$, Dec RTS
10	1	2	-3	$\mu' > 0$, Inc RTS

Table A1- 2. Piecewise Equations for the Efficient Frontier.

Enterprise 2 is an exemplar, with its efficiency frontier equation of the form $y - 0.25x - 2.5 = 0$ and Enterprise 3 is described as $y - 2x + 3 = 0$. Note that Enterprises 3 and 10 are collinear but Enterprise 8 also has its own plane. The piece-wise frontier of the efficient set of exemplars is illustrated in Figure A1-3.

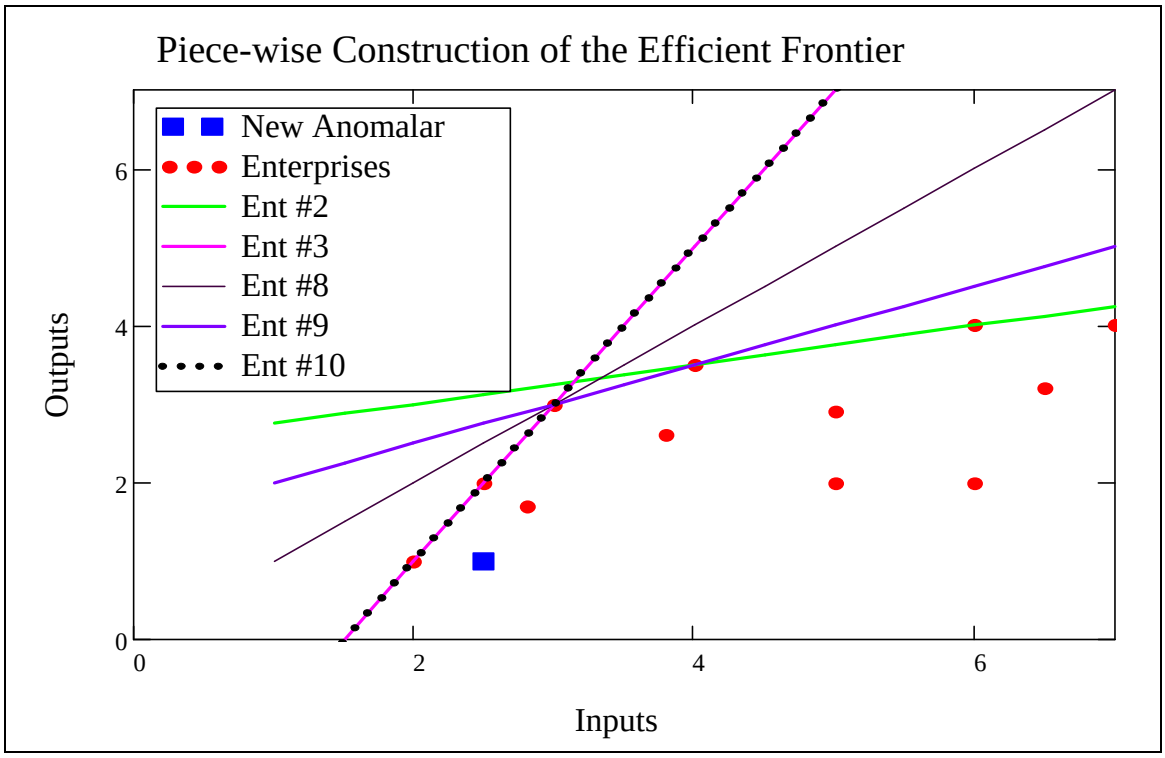

Figure A1- 3. Piece-wise Construction of the Efficient Frontier.

This Figure also reflects "returns to scale". Note that exemplars 3 and 10 have gradients greater than 45^o indicating that outputs scale greater than inputs; exemplar 8 has gradient exactly 45^o indicating that outputs scale as per inputs and the remainder have outputs increasing at lesser rates than corresponding inputs.

4. Applying the Efficient Frontier.

If a new anomalar, not in the original dataset, with, say, (Input, Output) pair (2.5, 1) wished to identify its input orientated exemplar, it would look to the performance of the nearest horizontal equation for exemplary performance. This places it on the increasing returns to scale frontier next to Enterprise 3. It would need to reduce its inputs from 2.5 to 2 whilst maintaining unity output. This improvement in the new Anomalar would be calculated

directly if included in the dataset and modelled according to an input-orientated envelopment model, e.g. a Charnes Cooper Rhodes (CCR) model (Cooper *et al.*, 2004, p 13).

This example demonstrates the use of DEA to establish technical efficiency amongst a group of enterprises and to derive improvement factors for anomalars external to the group. Technical efficiency also needs to be combined with allocative efficiency to ensure that exemplary pricing accompanies exemplary technology.

Appendix 2. Entailment and Supervenience.

1. Introduction to Entailment and Supervenience.

This appendix examines the application of two concepts that outline relationships between causal relata, environmental variables and states of affairs that might establish a benchmarking opportunity between an exemplar and anomalar. The concept of entailment is a rational encompassment of two sets of facts, models or formulae. The concept of supervenience is that of a nomological or metaphysical dependency between the whole and its parts.

In this Appendix the following hierarchies will be used: If **A** or **B** represent Organisations, $\{\mathbf{S_A}\}$ or $\mathbf{\{S_B\}}$ represent their (respective) states of affairs and $\mathbf{\{EV_A\}}$ or $\mathbf{\{EV_B\}}$ represent the respective environmental variables (resources, etc) that are associated with a corresponding state of affairs: i.e. $\mathbf{\{EV_A\}} \subseteq \mathbf{S_A} \subseteq \mathbf{\{S_A\}} \subseteq \mathbf{A}$.

2. Entailment.

The concept of entailment refers to a type of concurrence between two sets of properties or relationships. There is entailment if all of the properties or relationships within one of the sets are the same as at least one of the properties or relationships of the other set. These relationships or concurrences may assume many forms; they might be strictly logical- where two sets are described by deterministic relationships or 'fuzzier' where the relationship is described by probabilities, tendencies or dispositions. This concept is pertinent to benchmarking as many forms of exemplary relationships or properties represent potential avenues for anomalar improvement. Three forms of entailment (logical, dispositional and probabilistic) will be examined for their contribution to a theory of benchmarking.

2.1. Logical Entailment.

The unconditional truth of sets of properties is a necessary attribute of logical entailment. Suppose there are two sets of models (formulae or formal sets of facts) A and B. If every model or interpretation that makes all of the members of set A true also make at least one of the members of set B true then A entails B, written $A \models B$.

For example, if A is the model with environmental variables (*a*, *b*) and the state of A, $S_A = a + b$. Also if B is the model with environmental variables (*a*, *b*, *c*) and the state of B, $S_B = a + b + c$. Then A $\models$ B if every value of *a* and *b* applying to S_A makes at least one instance of S_B the same as S_A (i.e. true). For this example, it is only true when *c* is zero. Generally, the relationship is $S_A \subseteq S_B$, as shown in Figure A2-1 below.

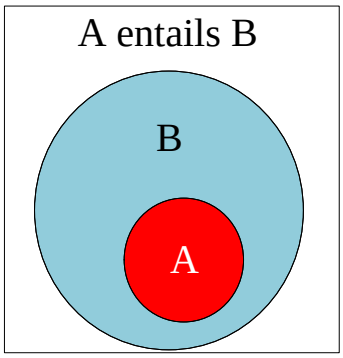

Figure A2- 1. Entailment.

Logical entailment is also a property of deterministic causal theory. Humean *efficient causation* (Chapter 3.5.6) holds that cause is the unconditional antecedent of an effect. We may say that if A is the cause of the effect B, A $\Leftarrow$ B, then every instance of A that gives rise to the effect B also satisfies A $\models$ B.

2.2. Dispositional Entailment.

This form contrasts with logical entailment in that it refers to properties that are manifestations of tendencies or propensities rather than purely categorical. There is robust philosophical debate about the ontological reality of such properties as some hold that some underlying essence or cause accounts for a disposition (Armstrong, 1969; J. L. Mackie, 1973) whereas others insist on their reality calling them qualifying properties (Weissman, 1965; Roxbee_Cox, 1975). Fara (2006) and Rozeboom (1973) provide examples of each: 'fragility, striving, responsibility, solubility, courageous and agile' are dispositional properties whereas 'massiveness and triangularity' appear to be categorical. The definition of dispositional entailment is similar to its logical precursor in that F expresses a disposition if and only if (iff) there is an associated manifestation and conditions of manifestation such that, necessarily, an object is F only if the object would produce the manifestation if it were in the conditions of manifestation (Fara, 2006, Paragraph 1). Observance of the disposition of exemplary manifestations towards some

desirable characteristic motivates the anomalar to benchmark against these conditions in order to improve. A symbol for dispositional entailment is ⫤, written A ⫤ B.

2.3. Probabilistic Entailment.

This 'hybrid' form contrasts with both dispositional and logical entailment as it relates to partial entailment based on 'a less than conclusive basis' for a proposition to be true (Hawthorne, 2004). A formal description of a probabilistic relationship, P, between a property, C and a conjunction of statements S, is P [C|S] = r; where $0 < r \leq 1$. A benchmarking example would be expressed along the lines that "the probability of failure (C) given the application of (process x, using materials, y with purity z) (S) is under 1%". Entailment is then based on a formula similar to its dispositional counterpart:

F probabilistically entails iff the probability of state of affairs with property C based on an associated set of statements, S, is r, then, necessarily, a state of affairs is F only if P [C|S] = r. The symbol denoting probabilistic entailment is ⊢, written A ⊢ B.

3. Entailment and Causation.

Both dispositional and probabilistic entailment reflect less deterministic relationships than logical entailment – as can be deduced if manifestations become categorical of if probabilities give way to certainties.

So we can say that ⊨ ⊢ ⫤, i.e. logical entailment entails probabilistic entailment and both entail dispositional entailment – however the converse is not the case.

Entailment has causal implications and it is informative to place it within a benchmarking framework.

Logical entailment is an essential characteristic of Humean-Mill *efficient c*ausation: the unconditional antecedents of the effect are its cause. E.g.

$\{EV_{cause}\} \models \{EV_{effect}\}$. E.g. {spark, gunpowder} ⊨ {spark, gunpowder, gases}.

If anomalars expect imitation to be the effect of benchmarking, then logical entailment of the exemplar's environmental variables and associated state of affairs is required. If anomalar

expectations are that organizational learning, some degree of improvement or triangulation are the effects of benchmarking, then dispositional or probabilistic entailment are required.

Dispositional entailment is a characteristic of both Humean-Mill and Peircean causation as dispositions can describe unconditional as well as teleological properties. Eggs have a fragile disposition and an unconditional antecedent of a hammer upon an egg has the effect of shattering it. Teleological dispositions also entail: food not only obtains the effect of nourishment (Humean-Mill *efficient c*ausation) but also satisfies the property (purpose) of long-run good health (Aristotelian and Peircean Final Causation). We may say that a property, 'good diet', dispositionally entails the property, 'good health'. In benchmarking, we might say that 'ethical behaviour' (for the sake of patronage and organisational survival) dispositionally entails 'success'.

Probabilistic entailment is a formal characteristic of Peircean causation as this theory holds that the nature of causation contains probabilistic elements that entail both *efficient* and *final* causation. Whilst these probabilistic elements and final cause are controversial, and have been discounted by some (Elster, 1983 p18) in respect to physical science, it is difficult to discount them in organisational science where uncertainties and the pursuit of survival are omnipresent. Indeed, a state of affairs such as a 'process' may be conditional on the properties of respective environmental variables to the extent that its status is satisfactory, even exemplary, if its outcomes lie between certain limits. For example, the efficacy of successive batches of recipe items such as soaps or cosmetics may be acceptable if their ingredients entail particular distributions of purities and proportions. Anomalars benchmarking such states of affairs should seek probabilistic entailment of the exemplar rather than logical entailment to obtain improvements.

4. Entailment and Benchmarking.

Critics of benchmarking cite the absence of theoretical mechanisms that enable distinction between effective and ineffective efforts (Chapter 2.4.2). Entailment is a theoretical mechanism relating properties of environmental variables that constitute exemplar and anomalar states of affairs. Entailment does not suggest that the laws associated with these properties are identical. For example most organisations have the environmental variable 'electricity' but each may use it differently. Thus entailment is not a sufficient condition for benchmarking, and may not even be necessary if benchmarking is simply used to stimulate

lateral thinking. Knowledge of property entailment can provide necessary conditions that distinguish between efforts seeking imitation, likely improvement and those unlikely to offer any improvement. For example:

- Imitation entails strict logical entailment of the exemplar and anomalar environmental variables. I.e. $\{EV_A\} = \{EV_B\}$ which is consistent with $S_A \subseteq S_B$. If imitation also means performance achievement within a set of limits to an accepted level of likelihood, then strict probabilistic entailment is necessary.

- Likely improvement arises from entailment in general. If $\{EV_A\} \subset \{EV_B\}$, there is still entailment (logical, dispositional or probabilistic), but as respective EV are only partial aligned, improvement rather than imitation might be obtained (if, as we shall see in the next section, a supervenience relationship also exists).

- Unlikely improvement arises from a lack of logical or probabilistic entailment. However in situations where there are no logical or probabilistic common properties shared between anomalar and exemplar EV, i.e. $\{EV_A\} \neq \{EV_B\}$, it is possible that a nomological relationship might exist and that there be some form of dispositional entailment. Such an example might be a successful 'conservation strategy' operating elsewhere on different environmental variables which might offer teleological rather than *efficient* causal improvements.

As entailment only relates to properties, a further consideration that assists benchmarking is the existence of some nomological relationship between the exemplar and anomalar.

5. Supervenience.

The notion of a nomological or metaphysical dependence of one state of affairs upon another is supervenience. A symbol denoting a supervenience relationship is ►, i.e. A ► B. Supervenience arose from the philosophy of mental characteristics where it was contended that they were dependent upon physical characteristics (Davidson, 1970/2001) such that changes in mental characteristics cannot occur without changes occurring in physical characteristics. This concept encompasses dependencies of a broader type and it is within this context that they are pertinent to a theory of benchmarking. In this work, causation theory is presented as a key explanatory driver for a theory of benchmarking. Cause

presupposes a relationship between the antecedents of the effect and the effect itself – a supervenience of the cause on the effect.

In Kim's (1984) view, a realist would probably accept that causation – the set of causal relationships – is dependent upon our knowledge of explanatory and epistemic relationships of our world – its nomology and metaphysics. Supervenience reflects the dependencies between crucial aspects of a whole and that of its parts. This holistic dependency can include its existence and nature (Kim 1984). Supervenience is not a restatement of reductionism, where one state of affairs is able to be defined in terms of another. Supervenience is a metaphysical or nomological relationship between sets of facts or properties and is not a semantic relationship between theories or languages (Stalnaker, 1996).

The distinction between entailment and supervenience lies in the distinction between properties and laws. Entailment refers to the properties of two sets of environmental variables and supervenience refers to the laws or nomological dependencies between them. Supervenience and entailment may or may not co-exist, depending on the circumstances. For example, the surface area of a sphere supervenes on its volume (and vice versa) as the law (formula) governing each of them only involves constants and the sphere's radius. A sphere's volume logically entails its surface area as each share a single property, radius.

In another example, consider two properties {being a brother}, {being a sibling}. Now {being a brother} logically entails {being a sibling}. Note that {being a sibling} does not logically entail {being a brother}, neither does {being a sibling} supervene on {being a brother}, but it would supervene on {children of the same parents}. A sister and an only child have identical properties in that they are both ⌐ {being a brother}, yet one is a sibling and the other not (McLaughlin & Bennett, 2005).

Another example offered by Mandik (2004) illustrates the application of supervenience and entailment to simple physical properties. Force is a function of mass and acceleration. The environmental variables that determine Force are {mass, acceleration}. We may also say that acceleration is a function of position – i.e. its environmental variable is {position}. Thus facts pertaining to Force relate to facts pertaining to mass and position. In terms of dependency, Force supervenes upon an object's mass and position as none of these can alter

without altering Force, yet the reverse is not true. The causal relationship between Force and its environmental variables is given by the relationship:

$$Force = mass.\frac{d}{dt}\left(\frac{d(position)}{dt}\right)$$

Equation A2- 1. Example of Supervening Properties

Supervening properties need not be the properties upon which they supervene. Although Force supervenes on position, it is not the same as position. There is also an entailment relationship in this example as position logically entails Force, but not the converse. Force is also a vector whereas position is a scalar.

Consequently, if an exemplar adopted a management law of rewarding success, it might supervene on an anomalar that also obeyed the law of "rewarding success" – even though there might not be entailment of their respective environmental variables.

In essence, all that is being said is that for X to supervene on Y, are that X's properties 'covary' with Y's properties (McLaughlin and Bennett, 2005, Sect 3.7). This is sufficient for benchmarking where the objective is not necessarily to imitate exemplary behaviour, but to derive benefit from it to the greatest extent possible.

5.1. Varieties of Supervenience.

There are three different levels of supervenience – weak, strong and global. There is also a temporal form that relates dependencies over time.

The essential difference between each level of supervenience is its scope of application. Strong supervenience describes nomological dependence over all known states of affairs. Global supervenience describes this dependence over a set of states of affairs and weak supervenience refers to a particular state of affairs.

If A and B represent two non-empty families of properties within a given set of possible states of affairs $\{S_A, S_B, \ldots, S_n\}$, levels of supervenience can be presented as follows (Stalnaker, 1996, p226): (Note, in this work, A generally denotes the anomalar and B, the exemplar, hence the definitions will be ordered to favour this convention)

- B weakly supervenes on A iff individual elements within any single state of affairs $\{S_i\}$ can differ with respect an B property only if they differ with respect to some A property.

- B strongly supervenes on A iff individual elements within the same or different states of affairs $\{S_A, S_B\}$ can differ with respect with respect to an B property if and only if they differ with respect to some A property,

- S_B globally supervenes on S_A iff two possible states of affairs $\{S_A, S_B\}$ within a domain of two worlds, $\{w_1, w_2\}$, differ with respect to the distribution of S_B properties iff they also differ with respect to the distribution of S_A properties.

That is, indiscernibility with respect to A entails indiscernibility with respect to B within the same or differing states of affairs (Kim, 1984; Blackburn, 2007). Formally:

B *weakly* supervenes on A, $B \blacktriangleright_w A$, if and only if, necessarily ($\Diamond$), for any property F in B, if anything *x* has F, then there is at least one property G in A such that *x* has G, and if anything y has G it also has F, i.e.,

$$\text{iff} \quad \Diamond \forall F \in B\,[Fx \rightarrow \exists G \in A(Gx \wedge \forall y(Gy \rightarrow Fy))]$$

Equation A2- 2. Weak Supervenience

(Kim, p 158)

B *strongly* supervenes on A, $B \blacktriangleright_s A$, if and only if necessarily ($\Diamond$), for any property F in B, if anything *x* has F, then there is at least one property G in A such that *x* has G, and necessarily, if anything y has G it also has F, i.e.,

$$\text{iff} \quad \Diamond \forall F \in B\,[Fx \rightarrow \exists G \in A(Gx \wedge \Diamond \forall y(Gy \rightarrow Fy))]$$

Equation A2- 3. Strong Supervenience

(Kim, p 158)

B *globally* supervenes on A, $B \blacktriangleright_g A$ if, for these two worlds w_1, w_2:$\{ B , A \} \subset \{ w_1, w_2\}$, and only if necessarily, for any property F in B, for everything that has F in w_1 there is at least one identical property G in A where necessarily everything that has G in w_2 also has F, i.e.,

$$\text{iff } \Diamond\forall x\in w_1 \forall F\in B\,[Fx \rightarrow \exists G\in A\,(Gx \wedge \Diamond\forall y\in w_2\,(Gy \rightarrow Fy))]$$

Equation A2- 4. Global Supervenience

The difference between these varieties lies in the extent to which properties are shared. It is possible that not all of the properties shared by *x* and *y* in A are also shared in B. If for some F in B, *x* has F but *y* does not, then weak supervenience prevails. Specifying the necessity for both *x* and *y* to share all properties of both A and B leads to strong supervenience.

Narrowing the domain from all worlds to particular worlds – as would occur if the referents occupied a particular paradigmatic framework – leads to global supervenience: i.e. globality within that framework, but not elsewhere.

This opens up the potential to use supervenience to distinguish between incommensurable organisational theories. Theories receiving support from global supervenience (in confined domains) are less acceptable than those receiving support from weak and strong supervenient relationships (in unconfined domains). Global and Strong supervenience are identical if w_1 and w_2 encompass the same domains.

Figure A2-2 represents the relationship B►A, in various domains.

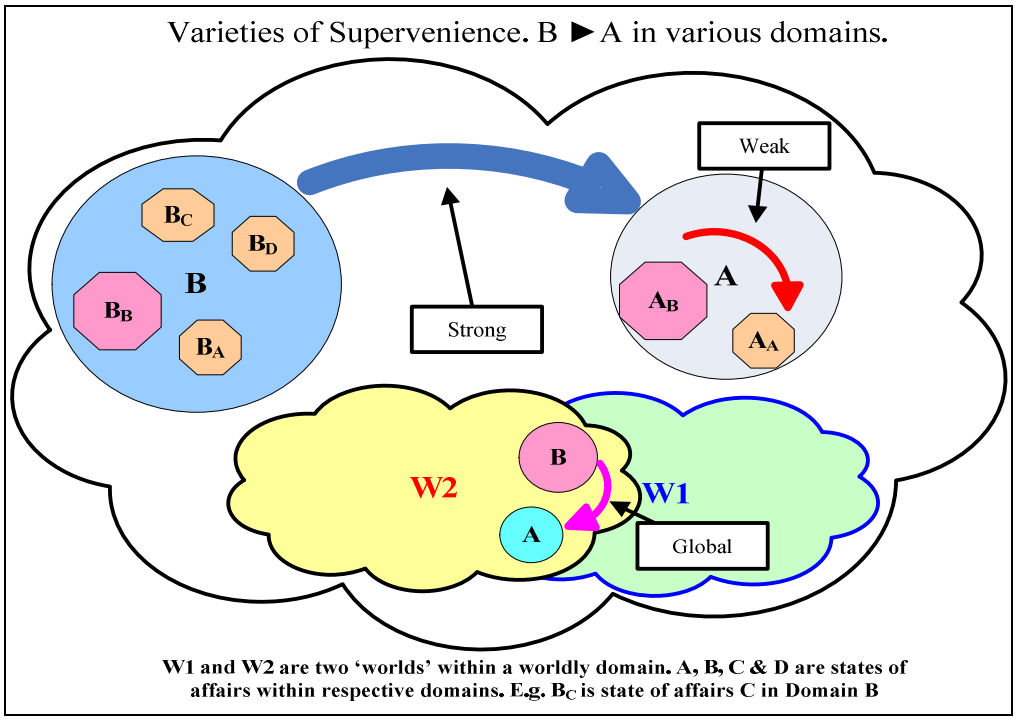

Figure A2- 2. Varieties of Supervenience.

From this it can be seen that strong supervenience entails weak supervenience and global supervenience is a localised form of strong supervenience (i.e. applying only within a fixed domain). It also follows that supervenience relationships are transitive (If A supervenes on B and B supervenes on C then A supervenes on C) and reflexive (A supervenes on A).

6. Applying Supervenience and Entailment to Benchmarking.

An anomalar A wishes to undertakes a benchmark exercise against one of its state of affairs, S_A which is determined by a corresponding set of environmental variables $\{EV_A\}$. Its objective is to obtain the greatest degree of improvement possible from a relevant exemplar B with a state of affairs S_B, determined by a corresponding set of environmental variables $\{EV_B\}$. The question is the degree to which the anomalar can be confident that the exemplar is a benchmark candidate and the extent to which it can be expected to offer improvement.

Using both the Humean-Mill and Peircean models of causation, improvements may be classified in terms of efficient and final cause. If Humean-Mill, only strict *efficient* cause applies; if Peircean, both efficient and final cause apply within the ambit of chance.

If $\{EV_A\}$ logically ($\models$) entails $\{EV_B\}$, facts that give rise to state of affairs S_A also give rise to at least one of the facts arising from S_B. Whether such facts produce a meaningful state of affairs for B is an issue, since $\{EV_B\} \supseteq \{EV_A\}$. Consider the following:

- If $\{EV_A\} = \{EV_B\}$, and B $\blacktriangleright_{s,g}$ A, the anomalar may be confident that the benchmarking exercise will be highly relevant as both the properties and laws associated with the exemplar apply to the anomalar. If B $\blacktriangleright_{w}$ A, some situations might apply where exemplar differences might not result in anomalar differences – possibly reducing the scope of improvement.

- If $\{EV_A\} \subset \{EV_B\}$, and B $\blacktriangleright_{s,g}$ A, the anomalar may be confident that some degree of relevance could result from the benchmarking exercise. If B $\blacktriangleright_{w}$ A, only some aspects of exemplar behaviour could be relevant to the exemplar.

If $\{EV_A\}$ dispositionally ($\models\!\!\models$) entails $\{EV_B\}$, manifestations of tendencies that apply to A's state of affairs also apply in part to those in B's state of affairs. If B $\blacktriangleright$ A, manifestations of these tendencies covary, thus providing a basis for pursuit of improvements.

If $\{EV_A\}$ probabilistically ($\vdash$) entails $\{EV_B\}$, benchmarking options are essentially similar to the case of logical entailment provided $B \blacktriangleright A$.

If B does not supervene on A, but there is some form of entailment, all that can be said is that the anomalar and exemplar share environmental variables having *exactly* the same properties. The absence of any property covariance precludes further improvement. E.g. the shared environmental variable 'electricity' confers no advantage to an anomalar in the absence of a law of usage.

Finally, if B does supervene on A, yet there is no logical or probabilistic entailment, the benchmarking exercise is restricted to understanding the dispositions or manifestations of exemplary behaviours that might lead to desirable changes in the anomalar's teleological trajectory.

Combining these combinations into tabular form provides a framework for the expectation of improvement that might possibly arise from a benchmarking exercise.

If A is the Anomalar, characterised by $\{EV_A\}$, and B the exemplar, characterised by $\{EV_B\}$, the following application of entailment and supervenience between these two organisations provides a basis for what might be expected from a benchmarking exercise. Note that in all cases, A or B may be replaced by a respective state of affairs, e.g. S_A or S_B and $\{EV_A\}$, $\{EV_B\}$ refer to their respective environmental variables.

Improvement Opportunity	**Humean-Mill/Peircean Model *Efficient***	**Peircean Model *Efficient***	**Peircean Model *Teleological***
Perfect Causal Improvements	$\{EV_A\} = \{EV_B\}$ $A \models B,\ B \blacktriangleright_{s,g} A$	$\{EV_A\} = \{EV_B\}$ $A \vdash B,\ B \blacktriangleright_{s,g} A$	$\{EV_A\} = \{EV_B\}$ $A \Vdash B,\ B \blacktriangleright_{s,g} A$
Potential Causal Improvements	$\{EV_A\} \subset \{EV_B\}$ $A \models B,\ B \blacktriangleright_{w} A$	$\{EV_A\} \subset \{EV_B\}$ $A \vdash B,\ B \blacktriangleright_{w} A$	$\{EV_A\} \subset \{EV_B\}$ $A \Vdash B,\ B \blacktriangleright_{w} A$
Potential Teleological Improvements	-		$\{EV_A\} \neq \{EV_B\}$ $A \Vdash B, B \blacktriangleright_{swg} A$

Table A2- 1. Feasible Benchmarking Improvement Criteria.

Conditions Precluding Improvement	$\{EV_A\} \subseteq \{EV_B\}$ And B ⌐► A	$\{EV_A\} \neq \{EV_B\}$ And B ⌐► A

Table A2- 2. Infeasible Benchmarking Improvement Criteria.

Applying the benchmarking taxonomy of Zairi (Chapter 2) to Table A2-1 relates improvement classes to classical benchmarking types.

Relationships between Zairi's Benchmarking Taxonomy and Causal Status	
Perfect Causal Improvement	Internal benchmarking, functional benchmarking (possibly), competitive benchmarking (possibly)
Potential Causal Improvement	Internal benchmarking, Competitive benchmarking, functional benchmarking
Potential Teleological Improvement	Generic benchmarking

Table A2- 3. Zairi's Benchmarking Taxonomy and Causal Status.

In summary, supervenience is a necessary condition for the expectation of improvement and entailment serves to strengthen the degree to which improvement might be expected.

Appendix 3. Theoretical Validation: the Anand & Kodali (2008) Universal Conceptual Benchmarking Framework.

In this validation analysis, Theory refers to the theoretical conditions (BT0:BT5) in Chapter 6 and the EBP Step refers to Table 7-1. This Appendix supports Chapter 8.2.2.

Theoretical Validation: Anand & Kodali (2008) Universal Conceptual Benchmarking Framework					
A&K Phase	A&K Step	Description	Theory	EBP Step	Analysis/Discussion
Team formation	**1**	Identify a leader of the team to carry benchmarking study	n.a.	Admin	These are administrative steps that should form part of organisational strategic planning. There should always be a process to establish BT1 & BT2 within every organisation. SMEs should find this much easier than VLEs
	2	Form a benchmarking team with clear-cut definition of responsibility for each team member	n.a.	Admin	
	3	Identify the capability of team and provide necessary training if required	n.a.	Admin	
	4	Identify the strategic intent/area of the business which is to be benchmarked	BT2	1	Knowledge of relationships between principal states of affairs and Welfare should be 'business-as-usual' rather triggered by a desire to perform benchmarking. Performing this process as a special requirement of benchmarking requires substantial incremental resource and likelihood of failure as it may be perceived as an ad hoc challenge to current practices or cultures.
Subject identification	**5**	Narrow down the number of subject areas (from the brainstorming stage) to a few areas in which benchmarking might have a high impact	BT2	1	
	6	Evaluate the importance of each subject area based on priorities	BT2	1	
	7	Identify benchmarking subject	BT2	1	
Customer validation	**8**	Identifying the customers for the benchmarking information	BT2	1	
	9	Identify key customer expectations	BT2	1	
	10	Validate the topic with respect to customers, company's mission, value and milestones, business needs, financial indicators, non-financial indicators, additional information that influence plans and actions	BT1, BT2.	2	This simulates the teleological nature of benchmarking. It also simulates the concept of a State of Affairs and its role in the current, developmental teleological trajectory

Table A3- 1. Theoretical Validation: the Anand & Kodali (2008) Universal Conceptual Benchmarking Framework.

A&K Phase continued	A&K Step	Description	Theory	EBP Step	Analysis/Discussion
Management validation	**11**	Prepare the mission of benchmarking and outline the purpose and scope of the benchmarking project	n.a	Admin	These Administrative processes have merit where they alert Management to the potential that may arise from improvement to a principal state of affairs whose relationship with overall welfare is known. Obtaining authority to alter a principal state of affairs is sound practice. If the EBP steps 1 & 2 were intrinsic to organisations (e.g. as within the Baldrige Framework) a key criticism of current benchmarking frameworks would be addressed.
	12	Identify different resources required for benchmarking study. It includes *all* resources.	n.a.	Admin	
	13	Prepare a proposal for benchmarking and submit it to management to get their commitment, with clear explanation on the benchmarking project, its objectives, tentative time plan of benchmarking activities with target dates, the benefits, costs involved, resources required, etc.	n.a.	Admin	
Self analysis	**14**	Understand the current situation by studying and analysing the existing information on the subject to be benchmarked	BT2	1,3	These steps simulate supervenience and entailment on the selected processes. They also seek to establish the composition (and existence) of the current teleological trajectory. There may be an implicit internal benchmarking decision in step 19. If it were explicit it would cement EBP step 3 and justify external benchmarking, thus avoiding criticism of unnecessary external benchmarking.
	15	Identify the critical success factors (CSFs) based on the subject of benchmarking, strategic intent, core competencies & capability maps	BT1, BT2.	1,2	
	16	Select the best performance measurement for CSFs	BT1, BT2.	1,2	
	17	Specify the data in terms of units and intervals to make the comparison and the analysis phase easier	BT1, BT2.	1,2	
	18	Measure the existing state of the subject to be benchmarked with respect to the CSFs	BT3	3	
	19	The subject to benchmarked is documented & characterised to determine and understand its inherent capability	BT3	3? 4? Admin	Documentation assists with the requirements of BT3

Table A3- 1. Theoretical Validation: the Anand & Kodali (2008) Benchmarking Framework, Continued.

A&K Phase continued	A & K Step	Description	Theory	EBP Step	Analysis/Discussion
Partner selection	**20**	Identify the external published information sources for collecting pre-benchmarking information by searching different technical and business journals, internal database, external databases, and public libraries	n.a	5	An administrative stage for gathering information on potential external partners based on internal specifications.
	21	Identify the potential benchmarking partners based on the above data	n.a	5	
	22	Establish the requirements for the selection of benchmarking partners or for the characterization of the degree of relevance that any particular company may have as a potential benchmarking partner	BT3, BT4, BT5.	5	Selection of a benchmarking partner solely based on the anomalous area under study does not recognise the need to satisfy a broader relationship (supervenience) with the exemplar. Partner Selection steps are principally '*efficient* causal' (note the examples in Step 24) but also need to be aligned with the exemplar's *final* cause – i.e. BT1 is missing. Also, the style of the exemplar is simulated. It is difficult to see how a dispositional or probabilistic exemplar might fit into A&K's framework.
	23	Narrow the list to few benchmarking partners by comparing the candidates			
Pre benchmarking activities	**24**	Collect lower level detail on benchmarking partner prior to contacting them (e.g. location, when did they get started, no. of employees, product line, key managers, market share, revenue profit, customer satisfaction, etc.)			
	25	Establish contact with the selected partner(s) and gain acceptance for participation in the benchmarking study	BT3, BT4, BT5, Admin	6	Administration steps (25, to 30) establish good exemplar-anomalar practices.
	26	Make an initial proposal, which includes the subject, reason for selecting the organization, what you expect from them, when to visit them, agenda for the visit, format of information that will be exchanged, etc.			

Table A3- 1. Theoretical Validation: the Anand & Kodali (2008) Benchmarking Framework, Continued.

A&K Phase continued	A&K Step	Description	Theory	EBP Step	Analysis/Discussion
Pre benchmarking activities continued	**27**	Determine the data collection method – which can be a questionnaire or site visits or interview or a combination of all methods	BT3, BT4, BT5, Admin	6 Cont.	Administration protocols (26, to 30) establish good exemplar-anomalar practices. Step 31 is an efficient practice to ensure the anomalar focuses on key environmental variables. These protocols only apply to implicit *efficient* causal relata between the benchmarking partners. In the case of other than logical entailment, exemplary causal relata may require some form of transposition or translation into feasible anomalar relata. This is not envisaged A & K's framework.
	28	Validate it after discussing with various experts including partners			
	29	Establish a protocol for performing the benchmarking study and also develop a non-disclosure agreement that tells about the information that will be shared and define the ethics of benchmarking			
	30	Prepare for reciprocal agreement, in case the benchmarking partner wishes to benchmark a different area in within the organization that wants to benchmark			
	31	Assess the information needs – write and review the questions, information required and other details to be collected with the benchmarking team members, so that there is a clear consensus and understanding about the information to be collected			
Benchmarking	**32**	Perform benchmarking study which might include collecting information through questionnaire/survey, interview, site visit, etc.	BT2	7	These steps broadly the same in both frameworks, but the EBP recognises 'feasibility' as an important relatum at this stage of the process. The concept of feasibility is not evident in the A & K framework.
	33	Collect data on methods, procedures, performance measure & practices that are considered superior			
	34	Sort the collected information & data			

Table A3- 1. Theoretical Validation: the Anand & Kodali (2008) Benchmarking Framework, Continued.

A&K Phase continued	A&K Step	Description	Theory	EBP Step	Analysis/Discussion
Gap analysis	**35**	Determine competitive gap	BT1, BT2.	8	The EBP considers 'Durability' as a hierarchical combination of improvements commencing with environmental variables, extending to states of affairs and finally, to the organisation itself. Welfare now appears implicitly in Step38. There is no distinction between local and holistic organisational improvement. Step 38 also implicitly refers to organisational teleology (management commitment)
	36	Identification of possible root causes and the superior practices that are responsible for the gap			
	37	Evaluating the nature of practices, methods, procedures (enablers) to determine their adaptability to the benchmarking company's culture by performing the feasibility study			
Action plans	**38**	Prepare the report and communicate the findings of benchmarking throughout the organization and project the benefits in terms of dollars and get the management commitment			
	39	Make results available to benchmarking partners			
	40	Establish functional goals			
	41	Project future performance			
	42	Develop the action plan with necessary recommendations & time frame for implementation			
	43	Gain acceptance from management and employees through commitment & participation, respectively, for implementing the action plans			
Implementation	**44**	Prioritize implementation of different practices			
	45	Deploy the action to the concerned product process owners with the target date for implementation & completion	Primal Axiom BT1, BT2.	9	
	46	Implement of action plans to bridge the gap			
	47	Provide training to the employees on new practices			

Table A3- 1. Theoretical Validation: the Anand & Kodali (2008) Benchmarking Framework, Continued.

A&K Phase continued	**A&K Step**	Description	Theory	EBP Step	Analysis/Discussion
Continuous improvement	**48**	Monitor results of the implemented actions	Primal Axiom BT1, BT2.	9 Cont.	Steps 48-54 reflect the tenor of EBP 9
	49	Check whether the target is reached			
	50	Recalibrate the benchmark and improve continuously			
	51	Ensure that best practices are fully integrated into process			
	52	Structure rewards system to recognize continuous improvement to the benchmarking team and the implementation team			
	53	Update the benchmarking report which provides the information on the best practices, how it was implemented in the benchmarked company and how it was adapted in the existing organization and a comparative analysis of the reported benefits, etc. which will help in learning purposes			
	54	Recycle the benchmarking process, i.e. perform new benchmarking studies for new areas/processes			

Table A3- 1. Theoretical Validation: the Anand & Kodali (2008) Benchmarking Framework, Continued.