Proliferation and propagation of breakthrough performance management theories and praxes
Guest Editors: Vinh Sum Chau and Jonathan Liu

669 Editorial advisory board
670 Guest editorial
675 Quality management as a determinant factor of productivity: a systematic literature review
Bertha Viviana Ruales Guzmán, Alessandro Brun and Oscar Fernando Castellanos Dominguez
699 Performance dialogue: a framework to enhance the effectiveness of performance measurement systems
Jerome Couturier and Nikolaos Sklavounos
721 Revisiting strategy mapping for performance management: a realist synthesis
Ryan Armstrong
753 Competency based superior performance and organizational effectiveness
Sateesh V. Shet, S.V. Patil and Meena R. Chandawarkar
774 Objectivity in performance ranking of higher education institutions using dynamic data envelopment analysis
Anup Kumar and Rajiv R. Thakur
797 High performance work practices and organizational performance-mediation analysis of explanatory theories
Naval Garg
817 Cognitive leader–member exchange differences between supervisors and subordinates
Yan Shen
838 Managing the survivor syndrome as scenario planning methodology ... and it matters!
Hong T.M. Bui, Vinh Sum Chau and Jacqueline Cox

ISBN 978-1-83867-887-6

www.emeraldinsight.com/loi/ijppm
<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dr Rabiul Ahasan</td>
<td>University of Oulu, Finland</td>
</tr>
<tr>
<td>Professor Bjorn Andersen</td>
<td>Norwegian University of Science and Technology, Norway</td>
</tr>
<tr>
<td>Professor Jiju Antony</td>
<td>Heriot-Watt University, UK</td>
</tr>
<tr>
<td>Umit Sezer Bititci</td>
<td>Heriot Watt University, UK</td>
</tr>
<tr>
<td>Professor Mike Bourne</td>
<td>Cranfield School of Management, Centre for Business Performance, UK</td>
</tr>
<tr>
<td>Dr Thomas F. Burgess</td>
<td>Leeds University Business School, UK</td>
</tr>
<tr>
<td>Dr Yu-wang Chen</td>
<td>The University of Manchester, UK</td>
</tr>
<tr>
<td>Dr Andre A. de Waal</td>
<td>Maastricht School of Management, The Netherlands</td>
</tr>
<tr>
<td>Dr Ibrahim Duyar</td>
<td>University of Arkansas Little Rock, USA</td>
</tr>
<tr>
<td>Dr Patrizia Garengo</td>
<td>University of Padova, Italy</td>
</tr>
<tr>
<td>Dr Richard Greatbanks</td>
<td>University of Otago, New Zealand</td>
</tr>
<tr>
<td>Professor Claire Hannibal</td>
<td>Liverpool John Moores University, UK</td>
</tr>
<tr>
<td>Dr Jacky Holloway</td>
<td>Northampton Business School, UK</td>
</tr>
<tr>
<td>Dr Rajashekharaih Jagadeesh</td>
<td>SDM Institute for Management Development, India</td>
</tr>
<tr>
<td>Professor Wagner Cezar Lucato</td>
<td>Nove de Julho University, Brazil</td>
</tr>
<tr>
<td>Professor Jill MacBryde</td>
<td>University of Strathclyde, UK</td>
</tr>
<tr>
<td>Dr Jane Maley</td>
<td>Charles Sturt University, Australia</td>
</tr>
<tr>
<td>Dr Keja Mendibil</td>
<td>University of Stirling, UK</td>
</tr>
<tr>
<td>Dr Matteo Mura</td>
<td>University of Bologna, Italy</td>
</tr>
<tr>
<td>Professor Sai Nudurupati</td>
<td>Gandhi Institute of Technology and Management, India</td>
</tr>
<tr>
<td>Professor Low Sui Pheng</td>
<td>National University of Singapore, Singapore</td>
</tr>
<tr>
<td>Professor Zoe Radnor</td>
<td>University of London, UK</td>
</tr>
<tr>
<td>Dr Nick Rich</td>
<td>Swansea University, UK</td>
</tr>
<tr>
<td>Professor Arash Shahin</td>
<td>University of Isfahan, Iran</td>
</tr>
<tr>
<td>Associate Professor Sanjay Kumar Singh</td>
<td>Abu Dhabi University, United Arab Emirates</td>
</tr>
<tr>
<td>Dr Alison Smart</td>
<td>University of Glasgow, UK</td>
</tr>
<tr>
<td>Professor Amrik Sohal</td>
<td>Monash University, Australia</td>
</tr>
<tr>
<td>Professor Karen Spens</td>
<td>Hanken School of Economics, Finland</td>
</tr>
<tr>
<td>Dr Jitesh J. Thakkar</td>
<td>Indian Institute of Technology Kharagpur, India</td>
</tr>
<tr>
<td>Dr Mike Tse</td>
<td>University of York, UK</td>
</tr>
<tr>
<td>Professor Eileen van Aken</td>
<td>Virginia Tech, USA</td>
</tr>
<tr>
<td>Dr Lihong Zhang</td>
<td>University of Manchester, UK</td>
</tr>
<tr>
<td>Professor Li Zheng</td>
<td>Tsinghua University, People's Republic of China</td>
</tr>
</tbody>
</table>
Guest editorial

Proliferation and propagation of breakthrough performance management theories and praxes

Introduction

Major theories on, and empirical research into, performance management related issues have proliferated in the past recent decades. These have covered a broad range of disciplinary areas – some of which are at their point of inception and some by borrowing and transferring to the neighboring discipline. A special issue of the *International Journal of Productivity and Performance Management* (*IJPPM*), published about a decade ago, took stock of the extent of this multidisciplinarity, but recognized “a number of barriers to effective knowledge transfer […] together with some gaps in the scope of research methods and theory-building” (Holloway, 2009, p. 398). For example, “although formal opportunities can arise to sell the contribution of performance management research, based in business schools, to potential clients in public and commercial organisations, many more informal opportunities arise through contacts with students, their employers, our friends and acquaintances […] [but] we all need a genuine shared language to surface assumptions and raise awareness of research in other disciplines that are interdependent of our own” (Holloway, 2009, p. 399).

A decade on, it is time for a new special issue of the *IJPPM* to revisit these concluding lines and go further to examine how/if these theories/research have impacted practices for the range of stakeholders, and also in neighboring disciplines. The articles included in this special issue look broadly at the proliferation (widespread) and propagation (deliberate attempt to implant in other disciplines/contexts) of breakthrough (significant, high impact, renowned) performance management theories and praxes (cf. practices, for its acceptability and custom use). In other words, performance management research has now surpassed multidisciplinarity and interdisciplinarity, but not necessarily for the better, and its usefulness may be overstated, contrived or deemed quackery, in the discipline or context of application. Now, some decades on since “performance management” as a common language phrase has become prominent, readers may decide after considering these articles if this view still stands or can be falsified.

Led by the performance management special interest group of the British Academy of Management (a leading authority on the general academic field of management and vital community for scholars and international peers to meet and advance theory and discern practice of over 30 years), the special issue has tendered its interest openly and in particular across a number of its sister academies whose research focus has also been in productivity and performance management. For example, the performance management association, also incepted two decades ago, which has an established bi-annual meeting place that networks cross-disciplinary academics and practitioners to advance knowledge and insight and exchange ideas for rapid exchange and the World Confederation of Productivity Science which has existed since 1968 with a global network which aims to explore productivity and performance management as a mechanism for creating wealth and quality of life. Incidentally, contributions have been received from scholars of these and other academies from all over the globe across which interest and research into the subject area has proliferated. These include the UK, wider Europe, India and China, suggesting broad extendibility of the field of performance management, and not just from the USA where most of their theories and frameworks originated or were incepted.
Overview of the contributions

The eight contributions chosen for inclusion in this special issue individually examine an aspect of a performance management breakthrough theory or praxis, and together sequentially present how productivity and performance management relate. They cover the forms of systematic reviews, conceptual propositions, empirical investigations and reflective practice.

The first article by Ruales Guzman et al. provides a systematic literature review of quality management in relation to productivity. This review is meaningful, as quality management has long been researched or presented as core to productivity due to quality being defined as a level ultimately required by the customer. The various theories relating to quality management go back to the pioneering Japanese principles of Crosby (1979) and practitioners involved in auto manufacturing (e.g. Feigenbaum, 1951) of over half a century ago. The article finds that productivity is now better represented by other variables not considered within the early quality movement, such as human resources, top management and process management. The authors argue, these go beyond manufacturing (and productivity) and relate more closely to performance management, which is broader. This identification is important as the present journal has grown to be more comprehensive in this way, against which the topics within its scope have also proliferated over the years.

The second article, by Couturier and Sklavounos, focuses on the performance management aspect to examine specifically the meaning of the now commonly used expression, performance management system (PMS). It begins with a review of the extant literature in the PMS area and then narrows down to an emerging area, identified as “dialogue.” Based on Mengis and Eppler’s (2008) framework for conversation management from the knowledge management perspective, the authors present a framework that enhances the effectiveness of PMSs as the growing importance of communication assists in coordinating the prerequisites known to be important within a PMS, both theoretically and usefully in practice. Hence, this article has demonstrated two key contributions – first, the proliferation of the performance dimension to the knowledge management domain as well as its proliferation to the communication literature; and second, the beneficial interdisciplinary integration for the performance management and communications literature.

Armstrong’s article which follows takes the specific aspect of strategy mapping, which originated in the 1990s and earlier, but popularized by Kaplan and Norton’s (2000) seminal work, and revisits it using a realist synthesis. While the idea of balancing different perspectives of an organization to augment the financial one to gain a comprehensive understanding of a company’s performance status and then to map out the cause and effect variables dates back a couple of decades and is popular in usage (Rigby and Bilodeau, 2015), the article argues that strategy maps are still underused and has greater potential than its current usage. Arguing also the under-utility of the realist synthesis in employing systematic literature reviews, practice may be suffering from a potential mismatch of the way strategy maps fit specifically within performance management. Hence, organizational effectiveness, perhaps through better alignment or other approaches, sits at the focal point for the future of performance management research.

The next article, by Shet et al. does just that by looking explicitly at organizational effectiveness in the form of superior performance for the firm. Based on the leadership competency model (LCM) which has emerged more recently in the people-based literature, the authors trace the need to appraise the workforce in a futuristic manner to the early work of McGregor (1957) with the origins dating back to the once dominating and breakthrough Taylorist perspectives of over a century ago. Following on from the need to maximize the potential of PMSs from a human, rather than machine-based humanistic, perspective, this article examines contributions in this area through competency-based techniques.
In so doing, the authors examine the in-depth meaning of competencies and contribute by developing a new scale for measuring competencies. They argue that few empirical studies have integrated LCM with other human resource processes. From the empirical results of their structural equation modeling analysis, then they argue that performance management initiatives can be accelerated by building better performance cultures within the organization.

The fifth contribution in this special issue, by Kumar and Thakur, follows on from the view of the need to develop performance measures and evaluates the use of methods to determine performance success. Thus, the well-known technique of data envelopment analysis for efficiency and ranking is critiqued and applied to the increasingly important context of higher education institutions. Tracing such an approach back to Cameron (1978), the authors utilize a more emerging technique known as dynamic data envelopment analysis to meet the needs of ranking which is now favoured and considered of greater use in the international league tables. The method factors in different societal, economic and political contexts to evaluate efficiency, providing more meaning evaluations to a broader range of stakeholders interested in the rankings.

The next article, by Garg, also considers the efficiency of organizations, from a well-known cognitive work theory, known as high-performance work practices (HPWP) from around the 1990s (e.g. Huselid, 1995) although their origins may have been earlier under different guises. Recognizing a number of empirical researches have taken place that considered the impact of HPWPs on organizational performance, missing is the mediating role of explanatory theories. The article therefore tests job satisfaction and organization citizenship behavior, which come from psychological impact theory, as mediators of HPWP and firm performance and finds partial mediation effects on the relationship. In this way, the performance dimension of a firm not only has moved on to consider human practices, but even further and deeper, to the cognitive and psychological aspects that support that underlying performance.

In a similar vein, the next article by Shen examines the cognitive underpinning of human performance through the leader-member exchange relationship between supervisor and subordinates. In this study, Shen reveals the influence of organizational identification on organizational citizenship behavior and the moderating effects of turnover intention. Based on social exchange theory, famous in the 1960s (Blau, 1964), Shen finds the quality of relationships within a firm is important, due to the positive moderating effects found present between the superior and subordinate, and can improve overall firm performance and individual satisfaction. Overall, she argues that organizations should strive to improve a subordinate’s organizational identification to encourage extra-role behaviors. This research was limited in that the quantitative study does not define or provide a richer context of the circumstances in which the “quality” of the relationship of employees in the firm is understood.

Hence, the last article by Bui et al. fills this gap by providing a reflective practice perspective on the relationship between employees and employer, to understand the cognitive emotions and feelings of employees who escape redundancy during major organizational downsizing. The article augments scenario planning methodology, a well-known technique first used in the 1960s (Wack, 1985), with those specific emotions known in the human resource management literature as the survivor syndrome (Brockner et al., 1986) to show they benefit one-another. Taking the 2008 global financial crisis (see Chau et al., 2012) as an extreme context of major economic turbulence in which redundancies were most likely, the article drew on the feelings of survivors who were interviewed to understand their feelings better. The authors find the feelings could be grouped within seven themes, and argue they should be used within scenario planning methodology to create more accurate scenarios as no other kind of research participant can
assist to craft hypothetical situations (scenarios) more accurately than those who almost befallen them. This is a praxis of performance management that can be improved by first improving the theory on which it is based.

Implications and future directions
The articles in this special issue have presented a sequential exploration of theories that relate to one-another and are combined in novel and valuable ways that have extended our existing knowledge about productivity and performance management. From these, and following up on the early queries raised by Holloway (2009), we are now able to throw some light on three fundamental questions:

(1) Is performance management still engaged in the multidisciplinarity/interdisciplinarity usefulness debate? Yes, but perhaps not intentionally. The contributions have shown that new constructs that have not previously been considered, but which they have had their place in neighboring disciplines, play a role in the performance management relationship researched. They are not questioned as to whether or not research should engage with other specific chosen disciplines, but rather that the research has involved cut across the realms of other disciplines which have subsequently been identified.

(2) Is there a shared language that has emerged from across the disciplines to which performance management relates that is understandable to a wider audience of stakeholders? Perhaps yes. As a broader group of scholars from other disciplines, practitioners from a range of industries and sectors who find use in performance management, and policy makers and beneficiaries of research have become more engaged in performance management research as research subjects or make use of its findings, a common language of terms relating to performance management improvement has emerged. Common terms, like high performance, organizational effectiveness, efficiency, superior performance, strategy alignment and communication, leadership and cognition, and methodological phrases like moderating/mediating effects, dynamic analysis and systematic review, have come into frequent use.

(3) Has research in performance management proliferated and/or propagated appropriately within their associated praxes? Definitely so. In the way performance management research has involved broad and new, intervening and related variables, there is evidence of pushing out the frontiers in the field by allowing neighboring disciplines to learn from its mainstream and established conceptions. The consideration of a theory being breakthrough or not is one that has made significant impact at the time and then stood a further test of time in terms of its validity and resilience against new and emerging ideas. However, breakthrough ideas vis-à-vis steady ones, or step and momentum changes, to borrow suitable terms from the established field (Lorange et al., 1986), are but part and parcel of the same thing. Old and dated ideas become augmented with new (breakthrough) ideas, which in turn become established/dated and get augmented with even newer ideas. These are then tested/applied to nascent contexts in the form of their associated praxes.

So where does this leave us for the future, and for an International Journal of Productivity and Performance Management? The contributions in this special issue have indeed offered broad coverage and extensive scope of the theme of breakthrough theories, as dictated by a changing society and business, but this is just the start of a life-long journey of evolvement in performance management research and praxis. Immediate future research might like to follow up on how truly different new ideas in the field are from their established
breakthrough ones that are claimed to be augmenting. But for the longer term, we expect new breakthrough theories and associated praxes to proliferate, and we look forward to the pleasure of your company in joining us in that journey.

Vinh Sum Chau
University of Kent, Canterbury, UK, and
Jonathan Liu
Regents University London, London, UK

References
Quality management as a determinant factor of productivity

A systematic literature review

Bertha Viviana Ruales Guzmán
Department of Systems and Industrial Engineering,
Universidad Nacional de Colombia, Bogotá, Colombia

Alessandro Brun
Department of Management, Economics and Industrial Engineering,
Politecnico di Milano, Milan, Italy, and

Oscar Fernando Castellanos Dominguez
Department of Systems and Industrial Engineering,
Universidad Nacional de Colombia, Bogotá, Colombia

Abstract

Purpose – The purpose of this paper is threefold: first, to analyse the current state of the literature on the relationship between quality management (QM) and productivity as a performance indicator; second, to identify the key constructs of QM practices related to productivity; and, finally, to reveal whether QM can actually be regarded as a determinant of productivity.

Design/methodology/approach – This research was carried out through a systematic literature review, considering 150 papers that studied this relationship between 1997 and 2017 and another 37 papers on the internal determinants of productivity.

Findings – The findings revealed that human resource management, top management and process management were the more relevant constructs of QM practices related to productivity. In addition, 89 per cent of the internal determinants of productivity were related to the proposed constructs of QM practices, which suggest that QM is a determinant factor of productivity.

Originality/value – This review analysed the literature on the relationship between QM and productivity, as few studies have done before, generating original, interesting and useful findings that can guide future research and that also represent a useful tool for researchers, practitioners, managers and policy makers.

Keywords Quality management, Productivity, Literature review, Determinant factor

Paper type Literature review

1. Introduction

One of the notable topics in the growing proliferation and propagation of breakthrough performance management theories and praxes is the study of the relationship between performance and quality management (QM). QM has been widely studied as a mechanism for strengthening performance and considered a distinctive organisational capability, a competitive strategy and a competitive advantage within the theory of the resource-based view of a firm (Elshaer and Augustyn, 2016; Del Río et al., 2017). The study of the relationship between QM and performance is a current topic and still of interest, with a growing trend for researchers and practitioners, as evidenced by the large number of publications in different contexts and disciplines. However, the large number of publications has focused mainly on studying the relationship with performance using general approaches such as firm; very few studies have looked at the relationship with more specific indicators of interest to organisations such as productivity.

QM has been recognised as a management philosophy that is related to productivity, profitability, firm performance (FP) and competitiveness and consolidated by Quality Gurus such as Feigenbaum (1951, 1961), Crosby (1979), Ishikawa (1976, 1985), Deming
According to Sedani and Lakhe (2011), the most popular approaches of QM are total quality management (TQM) and the ISO 9000 International Standards series; for this reason, this paper approached QM with these two approaches. TQM is defined as the integration of all functions, processes and levels of an organisation in order to achieve continuous improvement of the quality of products and customer satisfaction (Ross, 1993; Dubey and Gunasekaran, 2014), and the ISO 9000 International Standards series approaches QM through certification. The first version of the ISO 9000 standard was disseminated in 1987, later revised and updated in 1994, 2000, 2008 and 2015. According to Fonseca (2015), ISO 9000 cannot be considered a TQM, but he suggested, as Lizarzaburu (2016) did, that the ISO 9001:2015 version is closer to TQM because it is based on the seven principles of QM that are more consolidated, in contrast to the eight principles of the previous version (2008). The current principles of ISO 9001 are: customer focus, leadership, engagement with people, process approach, improvement, evidence-based decision making and, finally, relationship management.

Regarding productivity, it has been defined as the efficiency in the conversion of inputs to outputs (Syverson, 2011) and as an operational concept in terms of a saleable, quality product output per unit of input (Shahin, 2008); in summary, it is typically expressed as an output-input ratio (Solow, 1957; Chew, 1988; Tangen, 2005; Shahin, 2008; Syverson, 2011). On the other hand, productivity has been identified as the most important driver of long-term economic growth (Harris and Moffat, 2015) and one of the most vital factors affecting manufacturing company competitiveness (Tangen, 2005), and used as an indicator of the current and real situation of the economy of a firm, industry or country (Miranda and Toirac, 2010).

Studying the relationship between QM and productivity is important for both researchers and practitioners and was proposed several decades ago, with contributions from Deming (1982), Saraph et al. (1989), Flynn et al. (1994, 1995), Hendricks and Singhal (1997) and Samson and Terziiovski (1999). Despite the importance of this relationship, most of the studies have focused on the relationship between QM and general performance approaches, as evidenced in literature reviews, such as the study by Sousa and Voss (2002), where the impact of QM on FP was discussed; the research of Nair (2006), who identified which QM practices (QMp) are positively related to improved performance through a meta-analysis of correlation; and the study by Ebrahimi and Sadeghi (2013), which analysed QM and performance relationships. The valuable contributions of these reviews to the field of QM are evident; however, they have focused on studying the relationship between QMp (as an operationalization of QM) and general performance approaches, and in addition, they have approached literature review with a mainly narrative approach and have not been often totally encompassing.

Therefore, our research sought to provide a systematic literature review (SLR) of the current state of the literature on the relationship between QM and productivity as a performance indicator. In addition, it also aimed to identify the key constructs of QMp related to productivity and to reveal whether QM can actually be regarded as a determinant of productivity. To develop this research, a rigorous, well-defined and unbiased process was adopted using protocols that include comprehensive searches for all of the potentially significant studies (Tranfield et al., 2003; Tavares et al., 2016). The steps suggested in the literature for an SLR (namely: planning the review, conducting the review, and findings and discussion) were used as a research methodology. Two separate units of analysis were necessary to achieve the proposed goals, and the samples resulted in two totally different sets of papers, with no overlap, that encompass 150 papers for the first unit of analysis and 37 for the second one.

The main findings were: only 49 of the 150 papers considered productivity as a performance indicator in their analysis; the study of the relationships between QM and
performance and between QM and productivity is still in force, with an increasing trend over time; most of the studies have been conducted in Asia and Europe and few in countries of Latin America and the Caribbean; the majority of the studies focus on manufacturing firms, while few studies have analysed specific sectors; questionnaires were the most used resource, and few studies used more than one resource for data collection; QMp was the most used operationalization for QM; firm and financial were the most studied types of performance; productivity was mostly measured with the Likert scale; the most relevant constructs of QMp related to productivity were human resource management, top management and process management; and QM can be considered a determinant of productivity, since 89 per cent of the 36 internal determinants of productivity were related to the identified constructs of QMp.

The remainder of the paper is organised as follows: the next section describes the research methodology used in this study, the findings and discussion are seen in Section 3, and Section 4 shows the conclusions of this paper and provides directions for future research.

2. Research methodology
The SLR research method was used to accomplish the stated goals. A systematic review is a typology of reviews in which a search, analysis and evaluation of the research evidence is carried out using a protocol that results in a transparent report of the methods to facilitate its replication (Grant and Booth, 2009). The SLR method has been used in the research of different disciplines and fields of study, such as the research of Seuring and Müller (2008), in which a systematic review of sustainable supply chain management was carried out; the study by Colicchia and Strozzi (2012), which investigated the process of knowledge creation, transfer and development from a dynamic perspective within the context of supply chain risk management, and the study by Lopes et al. (2016), which analysed the links between lean manufacturing practices and organisation performance.

In this study, the steps suggested by the above-mentioned papers and by the specific paper on the SLR method (e.g. Carnwell and Daly, 2001; Tranfield et al., 2003, 2004; Cronin et al., 2008; Nightingale, 2009; Randolph, 2009; Tavares et al., 2016) were used as a guide to carry out this research. These steps included: planning the review: the research questions (RQ) and the protocol to delimit the unit of the analysis (inclusion/exclusion criteria) were defined; conducting the review: involved the identification of keywords, application of protocol to delimit the unit of the analysis (inclusion/exclusion criteria), review of abstracts and review of full-text of selected papers; and findings: included analysis and reporting.

2.1 Planning the review
Our review protocol encompassed two units of analysis to achieve the purposes proposed in this study. The first unit of analysis was used to achieve the first purpose. This unit of analysis included the literature on the relationship between QM and performance and between QM and productivity, since productivity is an indicator of performance and frequently analysed within general performance. In order to develop this purpose, the proposed RQ were:

*RQ1.* What is the current state of the literature on the relationship between QM and performance?

*RQ2.* What is the current state of the literature on the relationship between QM and productivity?

The identification of key constructs of QMp related to productivity was carried out only with the literature that analysed the relationship between QM and productivity as a
performance indicator and that reported the QMp related to productivity in the results. The RQ proposed for this purpose was:

**RQ3.** What are the key constructs of QMp related to productivity?

A second unit of analysis, totally different from the first, was used to identify the internal determinants of productivity at the plant or firm level. Subsequently, a theoretical relationship was proposed between the findings of these two units of analysis in order to reveal whether QM can actually be a determinant of productivity. The RQ at this stage was:

**RQ4.** Can QM be a determinant of productivity?

The following framework (Figure 1) was proposed for this research considering the objectives and previous RQ.

The proposed criteria for the inclusion and exclusion of documents in the two units of analysis are shown in Table I.

The inclusion criteria were taken into account in the selection of filters in each database, and the exclusion criteria were applied in the review of the abstracts of each of the selected documents or when the paper was read in its entirety. The databases used in the identification of the literature of the two units of analysis were Scopus and Web of Science since, according to Mongeon and Paul-Hus (2016) and Aghaei et al. (2013), these databases are still the main sources for citation data and are the two most extensive databases. On the other hand, the Source types selected in this study were Journals since, as Cronin et al. (2008) stated, journals are regarded as being more up-to-date than books as sources of information.

---

**Table I.** Inclusion/exclusion criteria

<table>
<thead>
<tr>
<th>Criteria</th>
<th>First unit of analysis</th>
<th>Second unit of analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Inclusion</strong></td>
<td>All subject areas</td>
<td>Engineering; economics, econometrics and finance; business, management and accounting; decision sciences and multidisciplinary</td>
</tr>
<tr>
<td>Subject area</td>
<td>All subject areas</td>
<td>All kinds of papers (empirical and theoretical developments and reviews)</td>
</tr>
<tr>
<td>Document type</td>
<td>Empirical articles</td>
<td>All papers until October 2017</td>
</tr>
<tr>
<td>Time frame</td>
<td>January 1997 to October 2017</td>
<td>Publications focusing on public policies and determinants of productivity in macroeconomic level</td>
</tr>
<tr>
<td>Exclusion</td>
<td>Publications related to lean without a distinction of results between TQM, JIT and TPM</td>
<td>Theoretical or anecdotal studies</td>
</tr>
</tbody>
</table>

---

**Figure 1.**
Research framework
2.2 Conducting the review
To conduct the review, “QM” and “Productivity” were selected as the main keywords for the first unit of analysis, and “Determinants” and “Productivity” were selected for the second one. Related keywords or synonyms (Table II) were identified through the review of the papers that included the main keywords and their respective bibliographic citations.

Both units of analysis were identified in the databases, combining the keywords with Boolean operators and using the filter “article title”. In addition, an additional filter (article title, abstract and keywords) with the words “plant”, “firm”, “industry”, “manufacture” and “manufacturing” was used. With this information, the search equations were created for use in the different databases, as follows.

First unit of analysis:

\[
\text{TITLE (“quality management” OR tqm OR iso OR lean)} \\
\text{AND TITLE (productivity OR efficiency OR “technical progress”} \\
\text{OR performance OR profitability) AND TITLE – ABS – KEY} \\
\text{(plant OR firm OR industry OR manufacture OR manufacturing). (1)}
\]

Second unit of analysis:

\[
\text{TITLE (determinants OR “determining factors” OR “decisive factors”} \\
\text{OR “Factors influencing” OR “Factors affecting” OR “influence factors” OR “Affecting factors” OR “important factors” OR “Key factors”} \\
\text{OR “factors that influence”)AND TITLE(eficiency OR productivity OR “operational performance” OR “Technical progress” OR “Production performance”} \\
\text{OR “manufacturing performance” OR “plant performance”) AND TITLE – ABS – KEY} \\
\text{(plant OR firm OR industry OR manufacture OR manufacturing). (2)}
\]

Subsequently, the papers were selected. Figure 2 shows the selection of the papers used in the first unit of analysis. With Equation (1), we found 672 papers in the Scopus database and 318 in Web of Science. However, after applying the inclusion/exclusion criteria, the final

<table>
<thead>
<tr>
<th>First unit of analysis</th>
<th>Second unit of analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main keywords</strong></td>
<td><strong>Related keywords</strong></td>
</tr>
<tr>
<td>Quality management</td>
<td>ISO</td>
</tr>
<tr>
<td></td>
<td>TQM</td>
</tr>
<tr>
<td></td>
<td>Lean</td>
</tr>
<tr>
<td>Productivity</td>
<td>Performance</td>
</tr>
<tr>
<td></td>
<td>Profitability</td>
</tr>
<tr>
<td></td>
<td>Efficiency</td>
</tr>
<tr>
<td></td>
<td>Technical progress</td>
</tr>
</tbody>
</table>

Table II. Main keywords and related keywords
sample was 483 documents. It is important to note that, in this study, lean articles with a global result and without a separation between TQM, JIT and TPM were not taken into account because these studies would not clearly identify the results for TQM, which is the topic of interest for this study. Then, all of the abstracts were reviewed in their entirety, and 150 documents were selected and used to analyse the current state of the literature on the relationship between QM and performance (RQ1). Afterwards, only 49 of the 150 papers were selected to analyse the current state of the literature on the relationship between QM and productivity (RQ2) because only these papers considered productivity as a performance indicator. Finally, 28 out of the 49 articles were used to answer RQ3 and to identify the key constructs of the QMp related to productivity. Both databases detected the same articles, which evidence the strength and rigorousness of the search. Additionally, articles cited in references (snowball search) were used as additional sources, but not many additional articles were found.

Similarly, the selection of the papers used to identify the second unit of analysis is shown in Figure 3, resulting in a sample of 37 papers. It is important to emphasise that the 37 papers found in this section were completely different from the 150 of the first unit of analysis, with which the internal determinants of productivity were identified and the possibility of QM being one was considered (RQ4).

3. Findings and discussion
The findings are divided into three sections. In Section 3.1, the first unit of analysis was used to analyse the current state of the literature on the relationship between QM and performance (RQ1) with 150 papers. In addition, in this section, the current state of the literature on the relationship between QM and productivity (RQ2) was analysed with the 49 papers that studied this relationship (see in greater depth the explanation of the selection of the 49 articles in Section 3.1.1). In Section 3.2, the key constructs of the QMp related to productivity were identified using 28 papers (RQ3). Finally, in Section 3.3, the second unit of analysis, with 37 papers, was used to identify the internal determinants of productivity and to reveal whether QM is one (RQ4).

3.1 Current state of the literature on the relationship between QM and performance and between QM and productivity (RQ1 and RQ2)
In Sections 3.1.1–3.1.6, the first unit of analysis was used, in which two samples were analysed. A sample of 150 papers was used for the analysis of the current state of the literature on the relationship between QM and performance (represented in the figures with the black colour and the legend performance), and another sample of 49 papers was used for the analysis of the current state of the literature on the relationship between QM and productivity (represented in the figures with the grey colour and the legend productivity).

![Figure 2. Selection of papers in the first unit of analysis](image)
![Figure 3. Selection of papers in the second unit of analysis](image)
The studies of Seuring and Müller (2008), Colicchia and Strozzi (2012) and Lopes et al. (2016) were used as a model in order to carry out this analysis, evaluating aspects such as the type of performance studied, the distribution of papers published over time, the classification by region, sector, type of data collection, operationalization of QM and other relevant factors of the groups of papers.

3.1.1 Type of performance studied. The range of performance approaches in the empirical studies was broad, which were classified into the categories: FP, financial performance (FnP), operational performance (OP), productivity or TFP (Pty), quality performance (QP), manufacturing/plant performance (MP), innovation performance (IP) and combination of performances (CP). The FP category included the denominations organisational, business, firm and non-financial and financial performance. The CP category encompassed papers that mentioned the analysis of various performances, within which 12 different combinations were found.

Figure 4 shows that FP was the most cited type of performance, which encompasses a global vision of companies, while more specific approaches, such as OP, QP and productivity or TFP, were less studied. On the one hand, our results are in line with those by Nair (2006), who, like us, found that some studies considered a multidimensional operationalization of performance, while others considered a single performance construct. On the other hand, the results of Ebrahimi and Sadeghi (2013), who stated that OP is a primary performance measurement (since it follows directly from the actions taken during QM implementation), differed from our results since we found that the percentage of studies that directly evaluated the OP was still low. In addition, according to our findings, it was evident that the number of papers that studied the direct relationship between QM and productivity is few in number (6 per cent, corresponding to nine papers); however, 40 additional studies involved productivity as an indicator of other performances, which means that 49 papers were used to analyse the current state of the literature on the relationship between QM and productivity.

These findings suggest that the analysis of the relationship between QM and more specific indicators of great utility for companies such as productivity should be included in future research, which could generate results of interest not only for academics, but also for practitioners, managers and policy makers since, as Harris and Moffat (2015) stated, productivity has been identified as the most important driver of long-term economic growth, one of the most vital factors affecting manufacturing company competitiveness (Tangen, 2005), and used as an indicator of the current and real situation of the economy of a firm, industry or country (Miranda and Toirac, 2010).
3.1.2 Distribution of papers published over time. The papers that studied the relationship between QM and performance and between QM and productivity were classified into three periods of time, seven years each (Figure 5).

The current state of the literature on both relationships has shown a growing trend, which demonstrates that this topic is still of interest despite the fact that it was addressed for the first time several decades ago by scholars such as Deming (1982, 1986), Saraph et al. (1989), Flynn et al. (1994, 1995), Hendricks and Singhal (1997) and Samson and Terziovski (1999). This trend can be attributed on the one hand to the importance of strengthening performance and productivity for company competitiveness and, on the other hand, as Sedani and Lakhe (2011) affirmed, to the fact that QM is currently seen as a pervasive management practice in modern business management, which generates great interest in studying the possible relationships and/or effects between these two variables.

3.1.3 Distribution of papers by region. The classification of the sample was done in eight categories (Figure 6). The region “Asia” included the countries India, Jordan, Taiwan, Singapore, China, Malaysia, Thailand, Iran, Turkey, Palestine, Qatar, Pakistan, Vietnam, Japan, Indonesia and the Philippines. “Europe” included Spain, Italy, Portugal, Greece, Serbia and one study in the European continent generally without specifying any countries. “Oceania” covered Australia and New Zealand. “North America” (North A) included the USA and Canada. “Africa” covered Tunisia, Ghana, Libya, Mauritius and Egypt. The category “Various Regions” corresponded to studies that analysed several countries from different regions. “Latin American and Caribbean countries” (Latin A) involved one study conducted in Brazil and one that included 31 Latin American and Caribbean countries. Finally, the category “Not Reported” (NR) corresponded to papers that did not report or did not clearly define a country or region of study.
Our findings show that the region Asia had the greatest amount of papers for both analysed relationships, followed by Europe and Oceania, while the regions Africa and Latin America and Caribbean countries had a lower concentration of studies (Figure 6). The more studied countries were India (13 per cent of papers), Australia (11 per cent) and Greece (8 per cent), representing a greater number of research than those found in regions such as Africa (6 per cent) and Latin America (1 per cent). The interest in conducting research in Asia and Europe can be attributed to the large number of ISO 9001 certified firms in these regions. According to the ISO (2017) Survey, 52 per cent of certifications are in Asia and the Pacific, and 37 per cent are in Europe, while in Latin America and Africa, the percentage is lower (4 and 1 per cent, respectively). Regarding to the high number of studies conducted in India can be attributed to the fact that, as Sedani and Lakhe (2011) affirmed, India has emerged as one of the top 10 countries in recent years, becoming the fourth largest manufacturing economy in the world.

Our findings can be explained from different perspectives and even supported by research, such as Ebrahimi and Sadeghi (2013), who asserted that most of the research studies on QM–performance relationships have been conducted in developed countries, with few studies in developing countries; however, these findings also show that it is necessary to guide future research towards conducting studies in regions of developing countries, which have specific economic and social contexts, and in which different findings could be obtained that would be of interest to the field of knowledge.

3.1.4 Distribution of papers by sector. This classification was made with seven categories (Figure 7), namely, manufacturing (manuf), manufacturing and service (manuf and serv), service (serv), automotive industry (automotive), NR, specific sectors (specific) and cross sectors (cross). The specific sectors category involved papers from 19 different sectors, such as hotel industry, textile industry, pharmaceutical industry, stock exchange, food manufacturing and cement manufacturing, among others. The cross sectors category included manufacturing and non-manufacturing; manufacturing, service and construction industries; manufacturers, service providers and wholesale traders; manufacturing, construction, retail and services sectors; and manufacturing, service, and computer and construction industries.

Manufacturing was identified as the sector in which more studies have been carried out (Figure 6). This result is in line with the study by Nair (2006), who found that most of the studies have focused on manufacturing. In addition, as Ebrahimi and Sadeghi (2013) argued, manufacturing firms have adopted QM principles to a larger extent than service ones (3 per cent), which explains the substantial difference in the number of studies between these two macro-sectors.

Figure 7. Distribution of papers by sector
In the paper by Sousa and Voss (2002), the need to increase research on the studied topic in specific sectors was manifested; a need that, according to our results, has not yet been satisfied because, although specific sectors were studied by a large number of papers, it is important to note that this category involved 19 sectors, which is equivalent to an average of only 1.1 per cent for each one. As a guide for future research, carrying out studies on the relationship between QM and productivity in specific sectors is recommended using as a model the studies carried out in the manufacturing sector since this is the most approached sector. This paper suggests that studies be carried out in relevant sectors for each of the regions with developing countries with key sectors for sustainability, such as the agro-industry or agro-food sector.

3.1.5 Type of data collection resource. Six data collection resources were singled out in this study (Figure 8): questionnaire (Quest), secondary sources (Ss), interview and questionnaire (Int+Quest), secondary source and questionnaire (Ss+Quest), interview and secondary source (Int+Ss) and interview, secondary source and questionnaire (Int+Ss+Quest). The category secondary sources encompassed the studies that collected data from sources such as other studies, company documents, government institutions, and data from institutions outside the company. Questionnaire was the most used data collection resource (70 per cent), followed by secondary sources. The popularity of questionnaire among researchers can be attributed to the fact that it allows easy and quick data collection, without the need for the presence of the researcher in the investigated units or direct contact with the interviewee, using phone calls or e-mails, which facilitates the evaluation of a large sample. In addition, sometimes questionnaire surveys are administered by market research companies, so that the research team can focus on the data analysis.

The percentage of articles that used more than one data collection resource was very low (11 per cent) when compared to those that used only one (89 per cent). The use of more than one resource of data collection in research implies the use of more economic resources, time and people, among others, but offers the possibility of triangulating data; a practice that, according to Voss et al. (2002), ensures validity in the results.

In future research, the use of more than one data collection resource is suggested in order to obtain more reliable results, which would be more useful for companies. The use of more than one data collection resource can be applied within different methodological approaches, one of which is case study; a method that, according to Ebrahimi and Sadeghi (2013), involves several resources at the same time.

In addition, Table III shows that the use of more than one data collection resource began in 2003 in the selected samples, but the highest concentration was found in 2010. This finding shows that, in the last decade, the interest of researchers in including more than one data collection resource has been growing, which reinforces the direction of
future research towards the use of multiple data collection resources to increase the validity of results.

3.1.6 Operationalization of QM. The operationalization of QM was classified into five groups: QMp; certified vs non-certified firms; NR; implemented TQM; and quality award. The operationalization through QMp was the most used (63 per cent) for both relationships, followed by certified vs non-certified companies (14 per cent). Similarly, the findings of Nair (2006) showed that the operationalization of the QM was carried out mainly through multidimensional construct (QMp) or single construct; unfortunately, in this study, the distributions were not presented for each one of the operationalization. Furthermore, in our findings, the QMp was named with different terms, such as TQM factors (e.g. Fotopoulos and Psomas, 2010), QM criteria (e.g. Sadiq Sohail and Hoong, 2003), TQM elements (e.g. Meftah Abusa and Gibson, 2013), QM dimensions (e.g. Sharma, 2006), TQM measures (e.g. Akgun et al., 2014), TQM variables (e.g. Terziovski, 2006) and critical success factors of TQM (e.g. Mehralian et al., 2017).

These findings reveal that there is a lack of standardisation in the terms used for QMp, which implies confusion in academics and industries. Researchers should make an effort to standardise the vocabulary used for the operationalization of QM in order to facilitate the compression of studies, facilitate the search for information and contribute to the solid construction of knowledge. Therefore, this paper used the term “QMp” since it was the most frequent operationalization found in the scientific literature.

3.1.7 Productivity as a performance indicator. This section and Section 3.1.8 used the 49 papers that studied the relationship between QM and productivity as the sample. The productivity was addressed with different approaches (Table IV), which was viewed through FP in 29 papers while only 9 papers directly studied productivity.

These findings reveal that productivity does not have a clear and standardized conceptualised across studies, and, in many cases, it has been confused with performance since, as Tangen (2005) affirmed, the concepts of productivity and performance are often mixed up and considered interchangeable. These findings can be attributed to weak knowledge on the concept of productivity, its implications and its importance, which has led to this indicator being relegated to a second rank and neglected or ignored by those who influence production processes and by researchers, managers and policy makers. A clear
conceptualization of the variables is recommended for future research in order to obtain reliable results and adequate interpretations. Because of the lack of a defined performance umbrella that covers productivity, and in concordance with the definition of productivity (relationship between ratios of output to the inputs), in this research we suggest that this indicator should be covered by OP, and OP should be covered by FP.

3.1.8 Productivity measurement techniques. Measurement techniques were classified into four groups (Figure 9). The category “other measures” included specific indicators of productivity. The results showed that this variable has been mostly measured with the Likert scale (71 per cent), followed by Cobb–Douglas, stochastic frontier or DEA (14 per cent) and less frequency with the use of specific indicators of productivity “other measures” (8 per cent).

The Likert scale generally measures variables through an evaluation based on perceptions or on the concepts that the respondent has about the variable, in this case productivity. In addition, evaluations are often done with a single respondent (often a top manager)
who evaluates all the variables. These conditions imply a risk for the objectivity and impartiality of results since, as Tangen (2005) affirmed, “many people who make decisions (top managers) to improve plant efficiency do not know what Productivity is and often confuse it with performance”.

According to these findings, future research should increase the use of specific indicators of productivity as a measurement technique in order to obtain reliable and valid data that subsequently allow for the generation of proposals for improvement. In addition, the visions of respondents from other levels and departments should be included in order to increase the reliability and validity of studies.

3.1.9 Other relevant factors of the papers. In this analysis, the total of the sample of the first unit of analysis was taken into account (150 papers). It was found that 131 of the 150 papers used a cross-sectional study. On the one hand, 50 per cent of these 131 papers evaluated the “relationship”, “association” or “correlation” between variables, which demonstrated the consistency between the hypotheses or RQ and the type of study (cross-sectional study). On the other hand, 40 per cent evaluated the “impact”, “effect” or “influence” between the variables, thus showing the inconsistency between the hypotheses or RQ of these papers and the type of study addressed (cross sectional study). Finally, 10 per cent of the papers did not propose hypotheses or RQ.

Our findings reflect the inconsistency that exists in the analysed literature for the type of study proposed (cross-sectional or longitudinal) and the hypotheses, objectives or RQ. Future research should employ a rigorous use of terms for relationships and impacts, with an adequate selection of the type of study, preferably using longitudinal studies to evaluate the impact or effect of QM on productivity or performance over time. The data collection resources of a longitudinal study could include secondary sources such as files and databases, among others, providing greater availability of information over a long period of time.

3.2 Constructs of QMp related to productivity (RQ3)
In this section, the sample of 49 papers that studied productivity was used. Only 29 of the 49 papers operationalized the QM through QMp; therefore, these 29 papers were selected to identify the QMp constructs related to productivity (RQ3). One of the 29 papers did not find any relationship between QMp and productivity; therefore, this study was not considered, and the final sample was 28 papers (Figure 10). The QMp identified in the sample were 38, classified in the eight proposed constructs described in Table V.

The results of this section were organised in Figure 10 taking into account the classification of the proposed QMp constructs. In this figure, the 28 papers that identified the QMp related to productivity were coded with a number, which were classified within the performance from which it addressed productivity (column papers that report QMp for the performance that involves productivity) or were directly classified in productivity when their results clearly showed the QMp related to productivity (column papers that report QMp for productivity). The frequency represents the amount of QMp related to productivity and reported by each author. This frequency can be evaluated by each construct, author and performance or productivity.

Our findings reveal that, on the one hand, 17 of the 28 papers (61 per cent) studied productivity in their data collection, but, in their research results, they did not report the QMp related to productivity since they only reported the QMp related to the general performance from which it was addressed. On the other hand, only 11 papers (39 per cent) identified the QMp related to productivity in their research results. The QMp constructs with the highest citation frequency in the papers that report QMp for the performance that involves productivity and in the papers that report QMp for productivity were human
Construct QMp

Table V. Constructs of QMp

<table>
<thead>
<tr>
<th>Construct QMp</th>
<th>QMp</th>
<th>Construct QMp</th>
<th>QMp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top management</td>
<td>Leadership</td>
<td>Process management</td>
<td>Process management</td>
</tr>
<tr>
<td></td>
<td>Top management support</td>
<td></td>
<td>Business and service process management</td>
</tr>
<tr>
<td></td>
<td>Top management commitment</td>
<td></td>
<td>Quality of process, product/service</td>
</tr>
<tr>
<td>Human resource</td>
<td>Factual approach to decision making</td>
<td>Process management</td>
<td>Process management</td>
</tr>
<tr>
<td>management</td>
<td>Human resource management</td>
<td></td>
<td>Business and service process management</td>
</tr>
<tr>
<td></td>
<td>Training and education</td>
<td>Supplier management</td>
<td>Supplier management</td>
</tr>
<tr>
<td></td>
<td>Employee relations</td>
<td></td>
<td>Supplier management</td>
</tr>
<tr>
<td></td>
<td>Employee involvement/participation</td>
<td>Supplier management</td>
<td>Supplier management</td>
</tr>
<tr>
<td></td>
<td>Employee empowerment</td>
<td>Supplier management</td>
<td>Supplier management</td>
</tr>
<tr>
<td></td>
<td>Reward and recognition</td>
<td>Customer focus</td>
<td>Customer focus</td>
</tr>
<tr>
<td></td>
<td>Favourable work environment and culture</td>
<td>Customer focus</td>
<td>Customer focus</td>
</tr>
<tr>
<td></td>
<td>Information and analysis</td>
<td>Continuous improvements</td>
<td>Continuous improvements</td>
</tr>
<tr>
<td>Process control</td>
<td>Systemic approach and documentary evidence for quality system</td>
<td>Continuous improvements</td>
<td>Continuous improvements</td>
</tr>
<tr>
<td></td>
<td>Quality data and reporting</td>
<td>Continuous improvements</td>
<td>Continuous improvements</td>
</tr>
<tr>
<td></td>
<td>Measuring results and performance</td>
<td>Continuous improvements</td>
<td>Continuous improvements</td>
</tr>
<tr>
<td></td>
<td>Process monitoring and control</td>
<td>Feedback</td>
<td>Feedback</td>
</tr>
<tr>
<td></td>
<td>Process control</td>
<td>Prevention of non-</td>
<td>Prevention of non-conformance</td>
</tr>
<tr>
<td></td>
<td>Selective application of tools and techniques</td>
<td>conformance</td>
<td>Prevention of non-conformance</td>
</tr>
<tr>
<td></td>
<td>Statistical tools</td>
<td>Product/service design</td>
<td>Product/service design</td>
</tr>
<tr>
<td></td>
<td>New product quality</td>
<td>Interfunctional design</td>
<td>Interfunctional design</td>
</tr>
</tbody>
</table>

Notes: FP, firm performance; QtyP, quality performance; OP, operational performance; Pty, productivity; Frequency, frequency of occurrence. *The paper reference number can be found in the bibliography, next to the reference.
resource management, followed by top management and process management. Additionally, the more cited QM constructs related to productivity are the three mentioned above; however, it is advisable to perform a test with empirical studies to validate this result.

Our results revealed that the number of investigations that study the identification of QM constructs related to productivity is still scarce, which prevents the consolidation of the theory of the relationship of these variables. The three QM constructs identified with the higher frequencies of citation named human resource management, top management and process management can be used as inputs for future research, where with empirical studies will test these findings in order to nourish and consolidate the theory and, at the same time, give practitioners, managers and policy makers a useful guide for strengthening productivity.

3.3 Internal determinants of productivity and its relationship with QM (RQ4)
The last section identifies the internal determinants of productivity and analyzes whether QM can be considered one with the 37 papers of the second unit of analysis. In the results reported by the different studies, 36 internal determinants of productivity were identified, which were subsequently classified within the QM constructs proposed in the previous section or in contingency factors (see Figure 11). The findings were organized in Figure 11, where each paper was coded with a number, and the citation frequency of each determinant and of each construct was calculated.

This study identified 10 determinants encompassed by human resource management construct, 14 by process management, 2 by supplier management, 2 by top management,

<table>
<thead>
<tr>
<th>QM Construct</th>
<th>Determinant</th>
<th>Frequency by determinant</th>
<th>Frequency by construct QM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human resource management</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
1 by product/service design, 2 by process control, 1 by focus customer and 4 by contingent factors. The findings also revealed that the two QMp constructs with the higher frequencies of citation and that encompassed the largest number of determinants were human resource management (with a frequency of 45 times) and process management (with a frequency of 42 times), results in line with the findings of Section 3.2 of this paper where these two constructs formed part of the three key constructs of QMp related to productivity.

The human resource management and process management constructs encompassed 24 of the 36 determinants (67 per cent), which suggests that these two constructs are relevant and should be considered in future research. The high citation frequency of human resource management can be attributed, among other factors, to the fact that many articles that study productivity often address it from the labour productivity perspective, where the main input is human resources. Taking into account the previous findings and realizing that 89 per cent of determinants are related to the constructs of QMp, QM can be considered a determinant of productivity. However, future research should look at these results with empirical studies with more indicators apart from labour productivity, such as raw materials, capital, energy and waste, among others. In addition, contingency factors should always be considered in studies, since they represent a very influential determinant for both QM and productivity.

4. Conclusions and further research
This SLR encompassed two totally different units of analysis, the first one was used to analyse the current state of the literature on the relationship between QM and productivity in order to identify the key constructs of QM related to productivity and, the second one was used to identify the principal internal determinants of productivity. Subsequently, a theoretical relationship was proposed between the findings of these two units of analysis in order to reveal whether QM is a determinant of productivity. The analysis was carried out in terms of the distribution of papers over time, with classification by region, sector, type of data collection, operationalization of QM, type of performance studied, type of Productivity measurement technique used, QMp related to productivity, main internal determinants of productivity and the relationship with QM.

This paper analysed the literature that studied the relationship between QM and productivity as few studies have done before. Previous revisions to this one used a mainly narrative approach, whereas in this systematic review, a more rigorous, well-defined and unbiased process was adopted, in which important indicators for companies, such as productivity, were taken into account, providing interesting and useful findings that will guide future research and that are a useful guide for researchers, practitioners, managers and policy makers. The main findings are summarised in the following paragraphs.

The current state of the literature on the relationship between QM and productivity has seen a growing trend, which shows that the issue still generates interest in researchers and practitioners. Asia and Europe were the regions with the higher concentrations of studies, whereas regions with developing countries such as Latin America and Africa had a lower concentration. Manufacturing was the general sector in which more studies were performed in comparison with contexts, such as specific industrial sectors. Questionnaire was the principal data collection resource, Likert scale was the most frequent measurement technique and perception of a single respondent was the most common source of information. In contrast, the use of more than one data collection resource is still scarce, and interest in its use has grown in the last decade. General approaches such as FP were the most analysed types of performance, whereas specific approaches such as productivity have received little attention. The key constructs of QMp related to productivity included human resource management, top management and process management. Furthermore, there was a relationship between 89 per cent of the internal
productivity determinants and the QMp constructs, which suggests that QM is a determinant of productivity. To conclude, some studies lacked methodological rigour because they did not have a conceptualization of the analysed variables that was clear and standardized or because they did not have consistency between the hypotheses or RQ and the type of study carried out (cross sectional or longitudinal).

This study has several implications for research, practice and society. The findings, on the one hand, show that QM is a determining factor for productivity (performance indicator of great importance for the economic growth of industries and countries) and, on the other hand, they reveal that the key constructs of QMp related to productivity are human resource management, top management and process management. These findings contribute to the consolidation of the theory of the relationship between QM and productivity, are a basis and guide for future research and are also a useful tool for managers and policy makers in the formulation of industrial policies that strengthen productivity.

Future research should conduct studies in contexts different from traditional ones, such as the regions of Latin American and African, and in specific industrial sectors in order to know their contingency factors and to propose alternatives for improvement and strengthening of these contexts. In addition, in order to obtain reliable and valid results, we suggest to use a standardisation of terms (e.g. QMp, performance, productivity, effect and relationship), use more than one resource of data collection involving visions of more than one respondent from different levels and departments, use objective productivity measurement techniques (e.g. productivity indicators) and perform more longitudinal studies that evaluate the effect of QM on productivity or performance. To conclude, the limitations of our paper are mainly focused on the low number of papers that studied the relationship between QM and important specific performance indicators such as productivity. For this reason, for future research we suggest performing empirical studies in different contexts in which this relationship is analysed and our findings are tested in order to feed the performance management theories and praxes, provide a guide for the decision making of practitioners, managers and policy makers and also contribute to the strengthening of productivity.

References


Ross, J. (1993), *Total Quality Management: Text, Cases and Readings*, St Lucie Press, Delray Beach, FL.


Further reading


Corresponding author
Bertha Viviana Ruales Guzmán can be contacted at: bvrualesg@unal.edu.co
Performance dialogue

A framework to enhance the effectiveness of performance measurement systems

Jerome Couturier
Department of Management, Escp Europe, Paris, France, and
Nikolaos Sklavounos
Department of Business Administration, University of Macedonia, Thessaloniki, Greece

Abstract

Purpose – The purpose of this paper is to provide guidelines for improving performance dialogue with a specific process and framework, leveraging existing literature.

Design/methodology/approach – Building upon Mengis and Eppler’s (2008) framework for conversation management, this study follows an action research approach, involving a process of co-creation, split into several distinct stages including two series of 20 semi-structured interviews with top executives of a major pharmaceutical company. These executives were directly involved in using the performance measurement system (PMS) in order to provide guidelines for improving performance dialogue. The data were analysed using content analysis, and the authors helped to develop a solution.

Findings – The analysis reveals a variety of recurrent communicative challenges and practices, which all appear to be characteristic for the performance dialogue process. The proposed framework consists of four separate phases, namely: data collection and identification of the main under and over-performance gaps, root cause analysis and action plans formulation, dialogue and solution implementation and dissemination of best practices.

Originality/value – This research contributes to the literature by introducing work on organisational communication into the field of PMS, by proposing a communication model for performance dialogue implementation. Furthermore, it addresses companies’ issues on how to successfully use their PMS and proposes a framework with specific prerequisites to be put into practice. Finally, this study offers a different explanation in the form of the lack of performance dialogue for the failure of PMS, compared to the current explanations found in the literature.

Keywords Innovative performance management, Performance measurement system, Performance dialogue, Sustainable performance, Organizational communication, Conversation management

Paper type Research paper

Introduction

Performance measurement is fundamental for effective organisation management. The substantial number of research studies and balanced scorecard implementations in multiple organisations demonstrate the popularity of this topic. Despite the increased interest in performance measurement in recent years, many companies fail to correctly use a performance measurement system (PMS), despite designing and introducing such a system. This paper aims to understand what happens when a company begins to use a PMS. The research focuses on the moments after the organisational leaders have designed and implemented a PMS. The main aim of this study is to demonstrate that managers can successfully use measures to manage their business and extract value from the data they collect by fostering the right conditions for internal structured meetings on performance, which is called performance dialogue.

This research is positioned vis-a-vis with some of the important work on organisational dialogue and communication, more specifically on communication models. Mengis and Eppler’s (2008) framework for conversation management is applied to analyse the use of PMS to inform managers and decisions made in a division of a major pharmaceutical company. The data are collected through two series of 20 semi-structured interviews and
analysed via content analysis. This analysis allows for the identification of some recurrent communicative challenges and practices in performance dialogue. Based on these and the existing literature, a framework for the successful implementation of performance dialogue is proposed.

After this introduction, the article reviews the relevant academic literature covering performance management and organisational dialogue and communication. It goes on to describe research questions and methodology, case findings and the proposed framework. Finally, the discussion of results and conclusions of the research are presented.

**Literature review**

**Performance measurement systems**

Performance measurement is a popular subject that has been recently investigated in numerous research and business studies (Neely *et al.*, 2005; Najmi *et al.*, 2005; Marchand and Raymond, 2008; Choong, 2013, 2014; Wieland *et al.*, 2015; Pádua and Jabbour, 2015). Businesses gradually recognised the significance of a multidimensional and balanced PMS as a tool to enable them to drive the company forward (Najmi *et al.*, 2005). Nowadays, PMSs have become multi-disciplinary and can be found in the following: the business sector (Bourne *et al.*, 2005; Franco-Santos *et al.*, 2007; Hubbard, 2009), governments and non-business sectors (Australian Government Process Interoperability Framework, 2012) and international agencies (Serrat, 2010).

High-performing organisations actively and repeatedly review their performance and assess their progress against established target values using elements, such as efficiency, effectiveness, profitability, quality of innovation and productivity (Huarng, 2011). Enterprises that achieve such high performance not only plan to maintain a predefined level of performance but also continuously do their best to optimise organisational performance by improving their performance elements (Oyemomi *et al.*, 2016).

During the years, PMSs have evolved from being based only on financial metrics that lacked the ability to adjust to the competitive changes facing the companies (Kennerley and Neely, 2003), to balanced systems that focus on both financial and non-financial indicators (Neely *et al.*, 2003). Famous examples include Kaplan and Norton’s Balanced Scorecard (1996), Neely *et al.’s* (2002) performance prism and Skandia’s Navigator, which was created by Edvinsson and Malone (1997). The aim of these systems is to help managers to “manage through measures”, in other words, to support decision making, set goals, allocate resources and inform management (Busi and Bititci, 2006).

Despite the popularity of PMS, an increasing number of authors have pointed out the difficulties of its successful implementation (McCunn, 1998). Looking at success and failure factors, many authors (Parker, 2000; Ittner and Larcker, 2003; Enoma and Allen, 2007) identify the selection of inappropriate or excessive measures as a failure factor and propose best practices for selecting these metrics. Bourne *et al.* (2000) and Neely *et al.* (2005) suggest processes for designing and implementing PMSs, while others (Moullin, 2004; Bourne *et al.*, 2002) are looking at the factors affecting these phases. The actual use of performance measures has also been researched, focusing mainly on a few tools or techniques that could be applied when assessing the implementation of strategy and challenging such strategic assumptions. (Neely and Bourne, 2002; Busi and Bititci, 2006; Neely *et al.*, 2003; Kennerley and Neely, 2003; Neely *et al.*, 2000; Goold and Quinn, 1990 in Neely *et al.*, 2005).

Micheli and Mari’s (2014) research has laid the foundations of a pragmatic epistemology of measurement in both the physical and social sciences. The researchers hold an epistemological stance, where performance measurement is regarded as a form of insight, rather than “true knowledge”. Thus, instead of being mainly aimed at providing “true descriptions” of the external world, theories can be considered “a form of insight, i.e., a way of looking at the world, and not a form of knowledge of how the world is” (Bohm, 1980, p. 4).
Bohm (1980, p. 23) also states that “measure is an insight created by man. A reality that is beyond man and prior to him cannot depend on such insight”.

In this context, it is significant that communication is regularly cited as a crucial factor within successful organisational performance (Merchant, 1989; Daft and Lewin, 1993; Tucker et al., 1996). Effective dialogue enables the sharing of employees’ experiences and encourages them to gain lessons based on these shared experiences (Malina and Selto, 2001; Schindler and Eppler, 2003). Regarding the use of PMS, Malina and Selto (2001) claim that a well-designed and effectively communicated PMS motivates and inspires both managers and employees to align their actions with the company’s strategy. On the other hand, when there are flaws in the PMS design or misunderstandings in strategic communication, problems will arise and the performance measurement initiative will not reach its performance enhancement target. Kaplan and Norton (2007), Bourne et al. (2000), De Bruijn (2002) and Bruijn and van Helden, (2006) acknowledge the importance of corporate dialogue by arguing that regular meetings among managers (or managers and employees) should be part of the PMS. An effective dialogue system, composed of regular meetings, would enable the exchange of knowledge and shared experiences among individuals and would contribute to reinforcing the organisational culture and strategy required. Creating the right structures for performance meetings to encourage creative dialogue will allow the use of information as a powerful resource for competitive advantage. Such a dialogue could allow the communication of meaningful information, instead of the raw data (Neely et al., 2003) and discussing the links between non-financial and financial measures (Malina and Selto, 2001).

In an effort to connect performance and dialogue, Moynihan (2005) conceptualises an integrative dialogue taking place based on performance information as “performance dialogue”. Performance data underlines success or failure, and the purpose of the dialogue is to identify and disseminate the reasons why success happens (Moynihan, 2005). Performance dialogue is a participative method, which assists employee engagement, motivation and performance.

Moynihan (2006) examines the Program Assessment Rating Tool (PART) in the US federal budgeting process and uses dialogue theory to improve performance in public administration. Dialogue theory emphasises the ambiguity inherent in interpreting performance information and related resource allocation choices. An exploratory test of dialogue theory is conducted through an experiment which involves graduate students assessing PART evaluations. Dialogue theory proposes that a variation in interpretation of performance information can be expected as a result of ambiguity of performance information and the influence of roles in the political process. The experiment assisted Moynihan (2006) in understanding how variation in interpretation between analysts employing reasonable and logical warrants can still lead to dissimilar conclusions. The results show various ways in which different individuals can examine the same programme and, using logical warrants, reach different conclusions regarding performance and future funding needs.

Moynihan and Pandey (2010) then argue that the availability of information is a key performance driver. Thus, the development of structures and procedures that enable and further strengthen interpretative processes and organisational learning based on performance information is intrinsically important (Rashman et al., 2009). According to Laihonen and Mäntylä (2017, p. 416), “In performance dialogue, performance data act as a stimulus for shared discussion, and this performance dialogue is expected to result in an actionable solution, which preferably has a positive impact on performance”. Their empirical findings also support Wouters and Wilderom (2008), who propose that the formalizations of the dialogue, which depend on contingent factors, are experienced based, allow experimentation and utilise professionalism, which refers to employees’ orientation toward learning. In general, Laihonen and Mäntylä (2017) provide a new
perspective on management control and organisational learning in public administration based on performance dialogue. Their research builds on a qualitative case study that explains how the focus of measuring and managing performance has changed in the city of Tampere, Finland, from designing and implementing new performance measures to providing a holistic culture and using performance information in a dialogic and integrative way.

These studies introduce the notion of dialogue as an important success factor in the context of performance management. However, they fail to provide specific advice on how to carry out performance dialogue in such a way that information is not only collected, but also used within the sense-making process, and for the creation and dissemination of knowledge. Most articles go only as far as presenting general guidelines for an effective performance dialogue. More specifically, the importance of a structured communication around meetings is mentioned or highlighted by most authors previously cited as playing a critical role, yet this concept has not been further developed. In addition, the more prescriptive articles are essentially focussed on the public sector. Our study adds to the existing literature as it is more specific, addresses the importance of structured meetings and is focussed on the private sector (namely, the pharmaceutical sector), while being performed over the course of several years with the full collaboration of the management team of the company under investigation. This paper introduces important findings into existing performance management research on organisational dialogue and communication, more specifically on communication models.

Organisational dialogue and communication

The idea that a communication process is a linear transfer of information from the sender of a message to the recipient through a medium is reflected in early models and theories of information, which originate in the 1940s (e.g. Shannon and Weaver, 1949). Some versions of these early models of communication are still presented in articles and books regarding communication (Fiske, 1982; Carey, 1989; Osborn and Motley, 1999; Higgins et al., 2001; McQuail and Windahl, 2015), but they also received serious criticism by scholars since they fail to emphasise the sense-making process involved in communication, in receiver or meaning-centred models (Gerbner, 1956; Craig, 1999; Herrmann and Kienle, 2004; Mengis and Eppler, 2008), where meaning is actively constructed by communication partners through activities of selecting, contextualising, interrelating and appropriating (Mengis and Eppler, 2007). Meaning creation is crucial for performance dialogue, because selecting, questioning, challenging, and interpreting raw data, and discussing the links between non-financial and financial measures ensure that employees’ actions are in line with the company’s objectives.

For the purposes of this research, Mengis and Eppler’s (2008) framework for conversation management was selected to analyse dialogue on performance. This is a recent framework in the organisational communication literature, which emphasises key variables of conversations that should be managed by participants to co-construct knowledge when interacting. Mengis and Eppler’s (2008) model for conversation management is a simple, management-oriented structure that refers to the interactive group sense of conversations. The authors suggest six main elements that should be used when creating and making sense of a message. These are: message, conversation process, conversational intent, mental models of the participants, group dynamics and conversational background. This framework is structured along the lines of what the case company was trying to achieve, which was a crucial deciding factor. Indeed, it is observed that the key principles of the framework (presented in Figure 1) were those implicitly valued by the management team of the studied organisation (although the team, before its intervention, did not clearly see how to practically apply those principles).
The message includes all the signs that are exchanged between the participants. From a knowledge perspective, the message should provide enough and appropriate cues so that the participants can interpret it adequately and should be well-rooted in facts. The conversation process refers to the time element of the conversation and describes the flow of the dialogue in time. There are several opinions on how structured the dialogue could be; some authors support the natural flow (Bohm, 1996), while others recommend a more formal procedure for effective conversations (Isaacs, 1993; Argyris, 1996; Harkins, 1999). Therefore, to successfully manage this area, the overall conversation flow should be structured in such a way to allow focus and synthesis, as well as exploration (Mengis and Eppler, 2008).

Conversational intent refers to the common and individual goals that are pursued in a conversation, which have a direct effect on how people will interpret the message. Often participants do not have the same or even compatible goals, and individual intentions remain hidden to others. Obscure goals or intentions could lead to a lack of commitment. Therefore, the communicational intent should be explicitly shared by all participants and should be oriented towards co-creation of meaning (Mengis and Eppler, 2008). Group dynamics are the result of the interaction among participants. This element plays a central part in the collaborative sense-making process when groups co-construct meaning within conversations. Dealing with formal and informal structures and coping with relationship conflict becomes, therefore, crucial for an effective conversation.

Mental models are interpretative schemas, which are used by humans to choose new information, interpret it and relate to it (Kim, 1993). They play an active role in the selectivity of the attention, the interpretation of a message and the construction of meaning (Mengis and Eppler, 2008). To manage successful conversations, participants need to be aware of the mental models that appear in a conversation and should be able to question and suspend them. Mengis and Eppler (2008) define the communicative background as all physical, organisational and cultural setting of a conversation, but also social networks and general communicative networks. Conversations shape and structure the communicative background and, at the same time, they are embedded in this larger context. Therefore, the selection of people, allocation of time, choice of physical space and the organisational culture will all play an important role in the creation, integration and sharing of knowledge. In conclusion, these six elements, which are presented in Figure 1, represent the communicative context on which experts and decision makers refer when attempting to integrate knowledge in decision making (Mengis and Eppler, 2008).

The aim of this paper is to extend Mengis and Eppler’s framework for managing conversations to provide guidelines to improve “performance dialogue” by sharing best practices and common ideas, and by elaborating solutions to problems. The underlying

![Figure 1. Mengis and Eppler’s (2008) framework for the management of knowledge-intensive conversations in organisations](attachment://figure1.png)
assumption is that dialogue, under the shape of structured periodical meetings, represents an essential element in the use of performance measures.

To answer the research question, the authors worked closely with the top management of the European and Middle Eastern (EME) veterinary division of a major pharmaceutical company for a period of more than three years. This company was selected as the focus of this study for several reasons. First, it is one of the largest pharmaceutical companies in the world, operating in more than 100 countries with 9,000 employees and 300 product lines. Second, it has been respected as a high-performing organisation over the last few decades, being in constant search of performance improvement with a well-defined and implemented balanced scorecard, key performance indicators (KPIs) aligned to company’s values and individual objectives in line with the company’s major strategy. Despite all these seemingly positive characteristics, the company has encountered major problems in using the balanced scorecard (BSC) as a strategic management tool to improve performance. As a result, the opportunity to study a company known as one of the highest-performing organisations in the world yet struggling to implement performance-enhancing management tools seemed particularly striking. This case-based approach gave the researchers the possibility to work closely with the division’s top management team, addressing managerial issues as well as deriving findings to contribute to the present study.

Methodology
This study follows an action research approach and, more specifically, the clinical field work method. The action research approach was initially developed by Kurt Lewin in the mid-1940s in order to address the problematic issues he identified in social action (Kemmis and McTaggart, 1988, Lewin, 1946). According to Lewin’s initial work, action research requires group decision and commitment to improvement (Dickens and Watkins, 1999), it is represented by the discussion of problems followed by group decisions on how to proceed (Adelman, 1993). Lewin’s original formulation of action research “consisted in analysis, fact-finding, conceptualisation, planning, execution, more fact-finding or evaluation; and then a repetition of this whole circle of activities; indeed, a spiral of such circles” (Sanford, 1970, p. 4). According to Argyris and Schon (1991, p. 86), “Action research takes its cues–its questions, puzzles, and problems – from the perceptions of practitioners within particular, local practice contexts. It builds descriptions and theories within the practice context itself, and tests them through intervention experiments, that is, through experiments that bear the double burden of testing hypotheses and effecting some (putatively) desirable change in the situation”. Schein (1987) builds on both the action research approach and his experience with process consultation to formulate his view of clinical field work, which he differentiates from other ethnographic techniques. According to the clinical perspective, the role of trained professionals who help with individuals, groups, communities or organisations is regarded as highly crucial. This professional “helping” role implies that the typical contribution of the researcher is facilitative (Baskerville and Wood-Harper, 1998). Baskerville and Wood-Harper (1998, p. 102) describe the clinical field work perspective as follows: “The study subjects typically initiate and drive a clinical enquiry, seeking help with an immediate problem. The client expects to pay fees. An action-change study model then operates which is conceptually normative (improve the problem situation) and narrowly focussed on certain detailed data regarding a particular problem. The study is validated by an improvement in the problem situation as a result of the actions. The achievement of this validity means that the study is primarily motivated by both organizational development and scientific knowledge”.

After an initial stage of problem definition with the management team of the organisation, the research involved two series of 20 semi-structured interviews with CEO/Managing Directors and Business Unit/Financial/Country Directors from three business units and eight countries who were directly involved in using the company’s PMS.
These individuals were best placed to observe and report potential effects of performance dialogue and were directly involved in setting and monitoring the organisation’s strategy. The interviews were semi-structured because the researchers wanted to get open answers to open-ended questions to determine the main problems they encountered when using the PMS. The interviews were conducted by two-person teams in which one of the team members was acting as interviewer and the other one as an observer and note taker. The interviews lasted roughly 60 min and the notes were transcribed after the interviews. The interviews were divided into three main parts: an introductive part that focuses on the professional background of the interviewee, the types of decisions that are part of his/her scope of responsibility and the organisational context of his/her work; a second part focusing on the process of performance tracking (challenges and practices of the performance tracking system) and, finally, a third part covering the challenges and practices that characterize the performance dialogue at a management level. In the first series of 20 interviews, typical questions in the interview guide were as follows:

- How is performance addressed and managed at different levels of your business unit?
- How would you envision the implementation of a performance dialogue process in your business unit?
- What are the main challenges you encounter when using PMSs?
- Do you regularly have open and frank conversations about performance?
- How do you share knowledge and best practices?
- From your experience in the company, what do you see as the real drivers for a better management of performance?
- How do you track the performance of your business unit?

The content analysis was then used to analyse the data from the interviews. Whereas content analysis is advocated by Easterby-Smith et al. (2002) as a means of analysing qualitative data, Mostyn (1985) refers to it as “the diagnostic tool of qualitative researchers, which they use when faced with a mass of open-ended material to make sense of”. A simple, open coding was used to categorise the raw data and break down and label the individual elements of information obtained. Tables were then used to further classify the categories (Miles and Huberman, 1984). The open coding was “in-vivo type” whereby the raw material used the wording of the interviewees, as opposed to constructed codes created by the researchers (Glaser and Strauss, 1967). Analysing the data, events were identified with common properties and grouped under the same concept. The data were read through repeatedly, and provisional labels for the blocks of data were created. The analysis was performed line-by-line, and also word-by-word. This line-by-line process of coding is used to build concepts and categories (Khandkar, 2009). In the next step, the researchers attributed the ensued challenges and practices along the framework for knowledge communication (Mengis and Eppler, 2008). This assisted with the depth analysis of the process of performance dialogue from a communication perspective and in the identification of a set of guidelines for improving “performance dialogue”.

Moreover, to verify the impact of performance dialogue, the second series of 20 semi-structured interviews was conducted more than three years after the initial series of interviews, using the same analytical approach of open coding. To complement the data, qualitative communication documents circulating between top management (which ranged
from presentations, brochures, articles, working reports, and the like) were also analysed. This helped the researchers to gather more evidence on the main challenges that were identified during the initial interviews. The analysis mainly focussed on elements such as market growth and market share evolution and involved the comparison of evolution indexes and net promoter scores between the company’s various departments.

**Company and context description**

The case centres on one of the world’s largest pharmaceutical firms involved in the research and development of veterinary drugs. The company entered the animal health industry in the early 1950s and had sales totalling $4.9bn in 2016. Research focussed on the Europe-Middle East division. Historically, the division had been a “financial holding” structure managed from the USA. The countries acted independently, and each was in charge of developing its own strategy. This changed in 2006 when operational headquarters moved to Europe and management decided to leverage the European advantage by unifying communication and targeting strategies across all its European units. Consequently, the division went through a process of tremendous organisational and structural transformation to achieve its vision of becoming a growth leader. It planned to reach this goal by becoming the customers’ clear first choice, an institution admired by competitors and a company where employees were proud to work.

To achieve this transformation, the company implemented a BSC in 2006 that was designed to provide the information required for effective performance monitoring and management decision making. Financial, sales force, technical and marketing KPIs were designed and implemented. These KPIs could be tracked down to all relevant levels of the organisation using an electronic system that provided a deep dive visibility for each KPI. Additionally, the company’s management team launched a cultural change programme to adjust the way the company operated and worked on a day-to-day basis. The new organisational culture programme, as well as the BSC, reflected the company’s strategy led by three goals:

1. to enhance the value proposition to customers to compete effectively against generic and increased commoditization;
2. to support the need for unification within the organisation; and
3. to emphasise execution.

The company invested a great amount of time aligning their strategic objectives with the matching KPIs. As a further step, they devised individual performance indicators aligned to the strategic objectives and KPIs. Even though these measures represented a great step forward, the company felt there was still a need to share best practices and solutions, to explain the strategic objectives and improve organisational communication. However, some felt that a performance-focussed culture was still not being fully embraced by the organisation. According to one CEO “managers [are] focussed on delivering the numbers of the KPI targets first and foremost, and [are] not reflecting on sustainable performance”. The CEO had initiated regular “performance meetings” to discuss business reviews, KPI achievement and performance gaps, with all of these based on the BSC set up by the organisation. Yet, those meetings did not provide any significant improvement to the organisation’s performance due to a lack of dialogue.

**First findings: some key challenges in the performance dialogue**

The first series of interviews revealed an overall need to share knowledge and best practices to find solutions to problems and to improve performance and thus, consolidate the company’s position on the market. “The value of performance dialogue does not only come
from the action plan (corrective actions or best practice sharing). The dialogue is also an opportunity to provide food for thought for the audience. We learn together” (Business Unit Director). This is very much in line with Malina and Selto (2001), who argue that effective communication systems may facilitate the exchange of knowledge and thus a better understanding of the organisation’s status. This may be accomplished by frequent and intense information sharing and dialogue rather than just by reporting to managers. The interviews also revealed a variety of recurrent communicative challenges and practices, which appear to be characteristic of the performance dialogue process. As mentioned earlier, an effective dialogue system, composed of regular meetings, is considered as an essential part of the PMS. These meetings facilitate the exchange of knowledge and shared experiences among individuals and reinforce the organisational culture and strategy required for the success of the organisation (Kaplan and Norton, 2007; Bourne et al., 2000; De Bruijn, 2002; Bruijn and van Helden, 2006). In the following sections, these identified challenges and practices are presented in detail.

Lack of structured process
There was a general agreement among the interviewed managers on the difficulty of successfully using the BSC as a strategic tool to enhance performance. In their views, the difficulty arose from the tendency to focus on measures during meetings. Data accuracy was constantly checked and the KPIs compared and challenged. “In the first meeting, we managed to access the data, in the second one we started to check the data” (Managing Director). Instead, meetings could focus on discussing key strategic issues, on finding solutions to problems and sharing best practices. For this reason, several attempts to use the BSC in business meetings had resulted in failure. “The balanced scorecard was useful for snapshots, but not for strategic decisions” (Managing Director). This pointed to a clear need for thorough preparation ahead of performance reviews to ensure focus on relevant issues and to appropriately guide the feedback and learning agenda. “[…] a robust and analytical part is missing in the discussion and reward in the leadership team” (Business Unit 2 Director). This is consistent with the findings of Kaplan and Norton (2008) and Ross (1994), who suggest setting agendas before each meeting and investing enough time to prepare for the set topics. Therefore, managers would be better off coming to meetings already familiar with the data and topics to be discussed and be prepared to formulate solutions to problems. “There was one test of performance review once in the context of an executive meeting. The test was unsuccessful as it failed to focus on areas of relevance to all. The only way these reviews can work successfully is if they are properly prepared ahead of time in order to draw attention to relevant issues” (Finance Director, Europe).

Lack of defined roles for participants
The interviews also indicated the informal existence of some performance dialogue-like events in different divisions at different levels, thus suggesting that the performance dialogue may develop spontaneously. “[…] I do it (performance dialogue) with my brand leaders, using financial data. I try to institute feedback, but it should be done systematically […]” (Business Unit 2 Director). Unstructured meetings are a good starting point for an eventual dialogue structure, which can evolve “organically” from existing interactions with the right resources. However, “Performance Dialogue should be carried in a structured way” (Head of Country division), it must rely on clearly defined roles to enable effective analysis and knowledge and best practice dissemination. The interviews revealed that the main responsibilities of collecting the data, performing sanity checks, and preparing and validating meeting objectives could be the responsibility of one person, the meeting facilitator. This person could also be responsible for ensuring that root cause analysis is performed by key managers before any review meetings. In addition, other participants
must ensure timely delivery and quality, as well as accuracy of the data and actively contributing to such meetings with solutions and ideas. Furthermore, under- and over-performance gap “owners” must perform the root cause analysis and present it during the meeting together with their vision and action plans. They could also ensure that follow up on these actions is performed after the meeting. The interviews also highlighted that each participant could be given the opportunity to develop a sense of ownership, not just of his/her own area, but of the overall project.

**Insufficient allocated time for conversations and preparation**

Another issue, related to the communicative background (Mengis and Eppler, 2008), is the lack of time to find solutions and improve performance. “We had a performance meeting each month focusing on three performance gaps. That left us with no time to discuss performance improvement” (Business Unit 1 Director). This finding is in line with a communication literature, where creating the time and space for a conversation is endorsed by Gratton and Ghoshal (2002) and Mengis and Eppler (2008) as one success factor for face-to-face knowledge-intensive conversations. Most participants complained of a lack of time, due to the need to provide data and conduct analysis and communication formats, “More time for preparation is needed” (Finance Director). Here an interrelationship between the communication background and the lack of structured process can be identified. Creating the time and space for conversations is strongly dependent on well a structured process that is supported by thorough preparation ahead of any such performance reviews.

**Lack of focus on the important issues**

On the message level (Mengis and Eppler, 2008), managers also voiced skepticism about the focus of the dialogue process on performance issues. The researchers found that the managers feared that meetings would fail to be effective unless a clear focus and deep level of analysis were required and enforced. “[...] it is useful for snapshots, but not for strategic decisions” (Business Unit 1 Director). “Our last executive meeting was a mess. We should focus on two things: why are we performing well or not well? What was well done?” (Country Director); “We should insist on a systematic analysis of the KPIs and BSC” (Business Unit 3 Director).

**Data accuracy**

Data accuracy is another issue that seems to be occurring at the message level (Mengis and Eppler, 2008), as a major challenge for performance dialogue. “The non-financial data is generated from different sources for different countries. There is a critical need for data validation” (Finance Director). “There will be issues of data consistency all the time, because the markets are different, some areas require more flexibility and sales reps are different, too” (Director of Business Technology); “I tried to use the KPI’s system, but the numbers are nonsense, so I stopped using it” (Business Unit 1 Director).

**Relational tensions emerging from differences in communication culture**

On the level of the group dynamics (Mengis and Eppler, 2008), the researchers observed that relational tensions often developed from differences in the communication culture. Typical performance review meetings had involved the full list of KPIs and discussions of their accuracy, variations against the budget and their causes. A move towards clearly identifying proper topics of under- or over-performance gaps was met with some cultural resistance and confusion. “Even though decisions are made by teams of people from all countries, for European projects, when it comes to implementing, people complain that the situation in their country is different” (Business Unit 3 Director). To overcome this issue,
the individuals who were interviewed suggested clearly defining under- and over-
performance gaps.

Lack of common goals
Another finding, which stands in line with previous research on knowledge integration, is
the lack of common goals, which accentuated the problems in performance communication
at the conversational intent level (Mengis and Eppler, 2008). Even though the BSC and the
KPIs system were aligned with the company’s strategy and the new organisational culture
programme, the different business unit leaders seemed to focus on different aspects of
performance. This led to a gap in communication. While the CEO’s main objectives were “to
enhance the value proposition to customers in order to compete effectively against generics
and increased commoditization; to support the need for unification within the organization;
and to emphasize execution because of the organization’s complexity”, other directors’
answers were: “market share growth” (Business Unit 1 Director), “become market leaders in
their market, successfully launch blockbuster, significantly expand the product position on
[…] and achieve sales budget” (Business Unit 2 Director), “increase market share, success
with the new product, customer satisfaction” (Country Director). Furthermore, a tendency to
focus on financial indicators was noticed, which was not in line with the company’s strategy
and BSC. “Performance is measured throughout the organization in terms of sales”
(Europe-Middle East Managing Director). “For finance, there are only three key indicators
that matter: sales, operating expenses and IBT” (Finance Director). “We normally look at %
of growth, productivity margin, sales, marketing margins and IBT” (Head of Country).

Discussion: a framework for successful implementation of performance
dialogue
The analysis of the first series of interviews has identified several important challenges,
which were present during the meetings where performance was discussed (including
during those ineffective “performance meetings” set up by the CEO). These are: a lack of
structured process, a lack of defined roles for participants, insufficient allocated time for
conversation, a lack of focus on critical issues and data accuracy, relational tensions
emerging from differences in communicative culture and a lack of common goals.
To overcome these challenges, a framework was devised to successfully manage
performance dialogue with the company’s experts and decision makers, building on Mengis
and Eppler’s framework for conversation management and on the findings from the case
study. Our framework (and the choice of extending Mengis and Eppler’s framework) was a
result of multiple conversations between us (the researchers) and the management teams of
the company under investigation during this study. In that respect, the framework emerged
from both field-based and theory-informed observations. More specifically, from the
analysis of the interviews, a need for a real dialogue on performance emerged, which also
mirrors the key principles of Mengis and Eppler’s approach. So, in agreement with the
management of the company, we decided at that stage to build upon Mengis and Eppler’s
framework and we jointly devised a more practical framework, which was implemented in
full over the course of three-and-a-half years by the company (from the beginning of our
intervention to our second series of interviews). The following sections discuss the various
components of our framework.

The interview findings at the process level (Mengis and Eppler, 2008) suggested that
performance dialogue would benefit from involving standard processes based on extensive
preparation and clear follow-up to the performance dialogue meetings. The BSC was useful
for snapshots and for providing the raw data needed for discussion, but not for making
strategic decisions. To overcome this, the participants need to be well prepared ahead of the
performance reviews. This preparation will provide the discussion with a robust and analytical segment, which will improve the sharing of best practices and common ideas. While the actual dialogue is obviously carried out during a business meeting, root cause analysis and data collection must be performed prior to the meeting to allow for a solution-driven meeting that leads to strategic decisions. Short of this, the dialogue can easily turn into a discussion on data accuracy and the causes of budget variations. To avoid over-focusing on control aspects when using the BSC, it becomes imperative to spend time prior to the meeting to investigate any potential problems and their root causes, as well as developing solutions to share and discuss with the other participants. In this way, the PMS will be used not only as an information collection and control tool but will also provide the support for sense making, and knowledge creation and dissemination. These findings are in line with Beer and Eisenstat (2004), Harkins (1999) and Bourne et al. (2000).

The implemented process could consist of four separate phases shown in Figure 2 and discussed below:

1. Data collection and identification of the main under- and over-performance gaps: this phase includes data collection and an accuracy check. Data analysis could also be carried out to identify under- or over-performance gaps. This helps to build the case for the meeting’s objectives and agenda. Furthermore, key areas could be identified by senior managers familiar with key strategic issues.

2. Root cause analysis and action plans formulation: in this phase, an analysis is carried out to identify specific root causes of the under- or over-performance gaps. In addition, potential solutions and an action plan need to be developed for discussion in the actual business meeting. The interviews emphasised the challenge of uncovering root causes with the use of benchmarking data from other areas of the organisation. Benchmarking is crucial to understand the real root causes and could be part of the performance dialogue process.

3. Dialogue: during the actual performance dialogue meeting, the main under- or over-performance gaps are presented together with the root cause analysis, and potential solutions and action plans. These serve as the basis for decisions on corrective actions and best practice implementations.

4. Solution implementation and dissemination of best practices: After the performance dialogue meeting, agreed-to actions need to be followed up and best practices and lessons learned disseminated.

Concerning dialogue background (Mengis and Eppler, 2008), performance dialogue could include well-defined roles for participants during the preparation, meeting and follow-up phases of dialogue sessions. Each participant needs to be aware of their own responsibilities.
and allocate the necessary amount of time for the preparation of the root cause analysis. This guarantees data accuracy, required depth of analysis and clear follow-up throughout. This finding is consistent with Kaplan and Norton (2008), Gratton and Ghoshal (2002), Von Krogh and Roos (1995) and Senge (1990) who highlight that the topics to be discussed during meetings will trigger different sets of attendees. Each attendee knows their area of expertise and will contribute accordingly. By thorough advance preparation and contribution in their areas of expertise, attendees will ensure a smooth and efficient performance dialogue process. Figure 3 presents the proposed roles of performance dialogue participants and their responsibilities according to their roles and time.

Regarding the grounded message, performance-focussed sessions could include a review of the KPIs and objective achievements, departing from root cause analyses and leading to action-oriented solutions. Among the corporate objective achievements, it is worth mentioning performance against budget, a focus on selected key areas of under- or over-performance, or those areas of noteworthy relevance, as well as a discussion of the solutions and applicability of the case to other business areas. Kaplan and Norton (2008) stress that performance-focussed sessions should review whether strategy execution is on track, whether problems are occurring in the implementation phase, and who will have responsibility for achieving targets. Beer and Eisenstat (2004) also urge participants to distinguish between identifying problems and giving recommendations.

Data accuracy is another key factor for successful performance dialogue and the lack of it was clearly perceived as a reason for non-constructive discussions, where participants lost time challenging the data and the related sources. This finding from the interviews is in line with previous research in the performance measurement literature (Bourne et al., 2000; Parker, 2000; Franco and Bourne, 2003; Neely et al., 2003; Kennerley and Neely, 2003). This aspect is very much interconnected with the dimension of a standard process. It was

---

### Figure 3: Performance dialogue participants' roles

<table>
<thead>
<tr>
<th>Before the meeting</th>
<th>During the meeting</th>
<th>After the meeting</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chair</strong></td>
<td><strong>Meeting facilitator</strong></td>
<td><strong>Participants</strong></td>
</tr>
<tr>
<td>• Validates objectives and agenda of Performance Dialogue meeting based on the input from facilitator of identified performance gaps</td>
<td>• Leads the discussion towards solutions</td>
<td>• All: Feed data</td>
</tr>
<tr>
<td>• Collects and analyzes data</td>
<td>• Approves corrective actions and best practice implementation and institutionalization</td>
<td>• Ensure timely delivery and quality of data</td>
</tr>
<tr>
<td>• Performs sanity check of data</td>
<td>• Presents the overall performance and highlights the main gaps</td>
<td>• Additional responsibilities for gap-owners:</td>
</tr>
<tr>
<td>• Pre-identifies performance gaps as topics of focus for the agenda</td>
<td>• Facilitates the discussion and presentation of vision and action plans</td>
<td>• Perform root-cause analysis and send results to the meeting facilitator</td>
</tr>
<tr>
<td>• Prepares the case in relation to the meeting objectives and agenda together with the finance and group directors</td>
<td>• Contributes to developing corrective actions and best practice implementation</td>
<td>• Formulate vision and action plans</td>
</tr>
<tr>
<td>• Proposes agenda and objectives to chair</td>
<td>• Keeps participants aware of time against schedule</td>
<td>• All: Active participation and solution contribution to corrective actions and best practice implementation, regardless of issue ownership</td>
</tr>
<tr>
<td>• Sends the pre-identified topics of focus to the &quot;owners&quot; to perform the root-cause analysis, asking specific questions and giving them precise instructions on expectations and type of answers</td>
<td>• Takes minutes</td>
<td>• Additional responsibilities for gap-owners:</td>
</tr>
<tr>
<td>• Presents the root-cause analysis</td>
<td>• Conducts end of the meeting review of discussion</td>
<td>• Present the root-cause analysis including vision and action plans</td>
</tr>
<tr>
<td>• Reviews the selected gaps and compiles possible solutions</td>
<td>• Review progress on approved action plans</td>
<td>• Follow up on actions</td>
</tr>
<tr>
<td><strong>Guests</strong></td>
<td><strong>Meeting facilitator</strong></td>
<td><strong>Meeting facilitator</strong></td>
</tr>
<tr>
<td>• All: Feed data</td>
<td>• Organize follow-up on approved action plans</td>
<td></td>
</tr>
</tbody>
</table>
proposed that the data accuracy could be achieved by following a standard process based on extensive preparation for and clear follow-up to the performance dialogue meetings.

When it comes to group dynamics, interview analysis showed that a clear definition of under- and over-performance gaps is a key to overcome the relational tensions emerging from differences in communication cultures. Managing interpersonal conflict by focusing on facts is also supported by Eisenhardt et al. (1997) and Quinn (1996). Another issue identified during the interviews was that individuals presenting under-performance gaps were being pointed out as a negative example. Relational tensions were thus developed, leading to a tendency to present only over-performance cases. To have a successful performance dialogue, it is recommended that individuals should first encourage positive knowledge sharing. This can be achieved by focusing any initial dialogue on highly successful cases that display a high performance (over-performance gaps), thus leading to the dissemination of best practices. Cases of under-performance should also be addressed, but in a subsequent meeting once the group feels they can openly discuss such issues, as an open and constructive dialogue is in place.

Finally, concerning the common dialogue intent, shared goals should be defined and pursued (Eisenhardt et al., 1997) and the aim of conversation should be bound to the co-creation of knowledge and performance improvement. This was highlighted in the interviews where self- and short-term interests often prevail over the company’s overarching long-term performance and practices. As a result, it was proposed that the meeting chair should always recall that participants’ positions should be temporarily “forgotten” during the meeting, and that everyone is expected to behave as the CEO in action, for the sole benefit of the company (e.g. vs their regional primary goals, etc.). It was, however, expected that the participants’ positions would become re-established and the organisational hierarchy would prevail outside of these meetings. But it was also observed (and confirmed during the second series of interviews) that the topics addressed in the performance dialogue sessions remained “protected” and were addressed with care beyond the meetings as if the naming of “performance dialogue” was somehow placing the topic at a non-hierarchical level (similar, according to some participants, to the mentoring-style discussions that take place at the organisation). According to the participants, this common dialogue intent was also helped by: the allocation of specific next steps, whereby everyone’s contribution was expected at the next meeting, to provide accountability for participants, and the high frequency of the meetings, which sometimes incorporated in-routine management forums. Indeed, the scheduling of the meetings progressively evolved. It appeared that scheduling and conducting frequent and short meetings with a reduced team guaranteed better results than longer and less frequent ones. With regular meetings that cover certain issues of focus, the exchange of opinions about performance dialogue is increased, the potential areas of questioning are identified and the overall performance is enhanced. Our second series of interviews revealed that doing so extended the intent for common dialogue beyond and in-between the performance dialogue meetings.

All the above findings and proposals can be summarised in the framework presented in Figure 4.

**Research impact and implications**

The findings of this research constitute a valuable addition to the pre-existing PMS literature. As highlighted at the beginning of this paper, communication is an important factor to the success of organisational performance and it has been extensively researched and well-explored. Yet, it is apparent that the discussion has been brought one step further by proposing a simple and practical framework that functioned with a high degree of consistency on a non-specific case. The proposed and applied framework is indeed prescriptive enough to allow further testing and research. Overall, this study introduces new dimensions to Mengis and Eppler’s framework. More specifically, the dimension of routine
is introduced, and it is proposed that performance dialogue must be institutionalized in a company in order for it to work and produce satisfactory results. Indeed, by making the original framework of Mengis and Eppler’s more specific and detailed, we’ve developed a concept that is not only more actionable for organisations (which increases its impact as the concept is easier to implement), but we have also produced a more relevant framework as the organisation started to use it as part of its routine business processes, for example, by introducing it in its business reviews. The impact here became significant since our framework contributed to developing a real culture of performance within the organisation, through a self-reinforcing process. More people use it in the first instance (since it is easily actionable, as previously discussed), generating greater dialogue on performance, producing more positive results, leading others to adopt the process to improve their performance and the overall performance of the organisation would therefore increase, and so on. In a nutshell, the co-creation process of our action research approach led to the development of an actionable framework that was thoroughly embraced by the organisation, creating routine behaviours of performance (e.g. including short performance dialogue sessions into normal business reviews), embedding such behaviours into the culture of the company.

When it comes to the case itself, the second series of 20 interviews was conducted three-and-a-half years after the beginning of the intervention when the first series of interviews were conducted. The interviewees were the same people (except for four individuals, who had left the organisation and been replaced) and the same structured approach was applied as the first series. Through these interviews, we sought to understand to what extent our framework was implemented and if so, whether there was any connection to improved performance. A similar open coding analysis indicated that the implementation of the proposed performance dialogue framework and its principles has widely disseminated across the organisation as such, that it had become a normal practice when discussing...
performance and had improved the company’s performance in terms of numbers, practices and by extending the implementation of the framework in many countries. In terms of numbers, the data show a substantial improvement in business performance both in sales and product launches. For example, in product launch efficiency there was an average improvement of 20 per cent after the implementation of performance dialogue and in countries like Germany and France (with already high scores in terms of business practices), the growth in sales doubled (from 10 to 20 per cent in six months). Obviously, these results need to be taken with caution, causality in performance is never clear. As highlighted by Micheli and Mari’s (2014, p. 150), measurement is always a “subjective evaluation”. Bohm (1980, p. 23) also states that “measure is an insight created by man”. Nonetheless, improvement in performance was extensively observed over time and across all geographies.

Most importantly in terms of practices, this second series of interviews revealed that our jointly developed framework is nowadays seen by the managers of the company as a way of doing business and not as a cumbersome and bureaucratic mechanism. Performance dialogue is now regarded as one of the standard practices of the company. Another indication that reveals the impact of performance dialogue in the enhancement of the company’s PMS is that it is no longer identified as a special project for only a few, select countries. Even though senior management has changed since its initial implementation, the PMS is still largely implemented by the company with some small adaptations in countries such as the UK, Germany, France, Italy, Spain, Greece, Ireland, Russia and Nordic countries. The context in which our framework is implemented has evolved and today there are sub-groups of two to three countries in which it is applied to make the whole procedure more practical (as opposed to each individual country implementing in a silo). As we previously discussed, the meetings decreased in length and tended to focus on one topic at a time with limited teams involved, yet maintaining the same overall principles and structure of our framework. Several interviews revealed initial concerns that the performance dialogue process was overly resource intensive. However, these concerns were alleviated by integrating performance dialogue sessions into usual management meetings, at all levels of the organisation. This step contributed to a real cultural change in the company, where performance dialogue was institutionalized, became part of the routine of conducting business, and its success was supported by numerous stories. Some interviewees spoke of a “culture of continuous improvement that generates performance […] without the need to focus on BSC dashboards”.

Figure 5 illustrates the link between the observed problems of the company and the specific practices that became routine, based on the proposed performance dialogue framework.
Regarding the possibility of the results being context-dependent, this research focuses on one industry (which is a potential limitation, although the pharmaceutical industry is not overly specific in many respects). Yet, it is expected that similar results will be found in different contexts (e.g. for different industries, in different locations, when considering different sizes or longevity of the companies under investigation). In the organisation under investigation in this study, data was gathered from eight countries of variable size with different cultures, and various longevities in the company (ranging from less than three years to more than 20 years). Moreover, these results were observed over a three-year period, where the context of the industry evolved considerably (e.g. increased competition, new management teams, etc.).

As far as managerial implications are concerned, this research’s findings should, equally, encourage practitioners to regard and use performance measurement instruments as strategic (rather than control) tools. The communication model will provide practitioners with the necessary guidelines to share best practices and common ideas, which will ultimately improve the decision-making processes and the overall performance of the company. In addition, managers could start the application of performance dialogue on a small scale within two to three countries to keep the procedure under control and achieve the best possible outcomes.

Limitations and objectives for future research

There are some limitations to this study that must be highlighted. One of them derives from the methodology, namely that only one case study has been used. Because the authors’ main intention was to develop theory, rather than to test it on a large scale, their findings are based on interviews and a solitary case. Yet, the use of a solitary case in management research is well-established, it has the capability of employing a range of qualitative and quantitative approaches, such as analysing archives, conducting interviews and using questionnaires (Gummesson, 2000; Yin, 2009). In addition, Laihonen and Mäntylä (2017) also selected a single case study approach in their study of performance dialogue in practice in order to attain detailed information regarding the circumstances, key events and processes related to timely concerns in public performance management.

Furthermore, this study raises two important issues for further research. One is the need to test the framework on a large scale, using more organisations from different industries. The other issue is to perform a quantitative longitudinal study to observe organisations over a period of more than five years and assess their success with performance dialogue implementation. More research needs to be conducted in this area that will benefit both scholars and practitioners. This research also focussed significantly on the dialogue meetings themselves, and less on the dynamic discussions that occurred after and in-between these meetings. While the participants did not mention any major issues with the discussions that took place outside of the dialogue meetings, further research could investigate whether Mengis and Eppler’s framework could also be extended outside of these controlled, dialogue meetings. Finally, the researchers acknowledge the fact that the actual impact of performance dialogue on the enhancement of the effectiveness of PMS is difficult to measure. However, the indications of substantial improvement of the company’s performance in terms of numbers, practices and extended implementation in many countries after the implementation of performance dialogue is worth further examination and certainly constitutes an interesting addition to the existing literature.

Conclusion

While performance management systems have been extensively studied and used by companies for several years, they are still not effectively used to enhance performance.
The literature review found that, despite the abundance of research on performance measurement, there are relatively few studies on the use of performance instruments once they are implemented and that few practical tools are currently available. This paper highlights the importance of acknowledging and establishing the need for performance dialogue after the implementation of a PMS. It also introduces work on organizational communication into the field of PMS, by proposing a communication model for performance dialogue implementation. Furthermore, it addresses the issues many companies face as they attempt to successfully use their PMS. It also proposes a framework with specific prerequisites to be put into practice. Finally, this study offers a different explanation than those currently found in the literature for the failure of PMS, namely, the lack of performance dialogue.

References


Khandkar, S.H. (2009), *Open Coding*, University of Calgary, Calgary.


Further reading


Corresponding author

Nikolaos Sklavounos can be contacted at: sklavou@uom.edu.gr

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Revisiting strategy mapping for performance management: a realist synthesis

Ryan Armstrong
Economía i Empresa, Universitat de Barcelona, Barcelona, Spain

Abstract

Purpose – The strategy map represents a major contribution to the theory and practice of performance management. However, it has failed to realize its full potential due to a lack of theoretical and conceptual development. Therefore, the purpose of this paper is to revisit the theories of strategy maps to better understand how and in what circumstances they benefit performance management.

Design/methodology/approach – The study employs realist synthesis, a method of systematic literature review. A theory on how strategy maps work is extracted from performance management literature, which are subsequently evaluated through a critical examination of empirical studies.

Findings – A theory of how strategy maps are meant work is presented in relation to the generic performance management stages of problem structuring, development and use, where they can serve as a tool for discovery and by stimulating social interactions. Based on the findings, 12 propositions are offered related to the effective use of strategy maps within a performance management framework.

Research limitations/implications – The introduction of the strategy map to performance management represented a breakthrough in how organizational performance could be understood and communicated. This study goes a step further by considering how they work and in what circumstances. In so doing, the study aims to open the way for new and more effective applications of strategy maps within the changing performance management context.

Practical implications – This study provides practitioners with actionable propositions which can help in effectively using strategy maps.

Originality/value – Distinguishing the aims and mechanisms of the strategy map along performance management systems has the potential to greatly increase their effectiveness in practice as a powerful, but underutilized tool. This paper also demonstrates how realist synthesis, currently an uncommon method in management studies, facilitated the creation of a new perspective of strategy maps to fit specifically within performance management.

Keywords Performance measurement, Balanced scorecard, Strategy map, Performance management systems, Realist synthesis

Paper type Literature review

1. Introduction

Only a few years after introducing strategy maps to performance management by incorporating them into the balanced scorecard, Kaplan and Norton (2004) remarked that these were “as big an insight to executives as the balanced scorecard itself” (p. 9). It was a significant observation, given that the balanced scorecard became one of the most widely used frameworks used in practice (Rigby and Bilodeau, 2015). The power of the map as first introduced stemmed from its purported ability to effectively describe strategy in a cohesive and straightforward way, thereby increasing the likelihood of successful strategy implementation (Kaplan and Norton, 2001). Strategy maps could also be used in aiding in formulating strategy, in structuring problems, in defining measures and objectives, and in decision making (Kaplan and Norton, 2004; Kaplan and Norton, 2006; Lueg and Julner, 2014).

However, nearly two decades after introducing the strategy map to performance management, evidence suggests that the impact of strategy maps for performance management, evidence suggests that the impact of strategy maps for performance...
management practice remains limited. There is evidence that few organizations use strategy maps as a part of the balanced scorecard or other performance management framework (Speckbacher et al., 2003; Tapinos et al., 2010), despite these being linked to effective use and satisfaction (Laitinen et al., 2010; Lueg and Julner, 2014). Further, strategy mapping in general often fails to be included in descriptions of the balanced scorecard (see Rigby, 2017), and is seldom used as a standalone tool in practice (Tapinos et al., 2010). In short, it appears that the strategy map and strategy mapping have not realized their potential for performance management.

There are several issues that could explain the lack of impact to date. First, descriptions of the role of strategy maps and how they are meant to work within the balanced scorecard framework have remained vague, often do not specify the outcome intended through their use, or apply overly generalized conceptions of performance (Hoque, 2014; Lueg, 2015; Öllinger et al., 2015). Second, many scholarly works on the strategy maps remain normative (Islam, 2018), or take the limited view of the strategy map as a management control device (Tapinos et al., 2010). Despite a few developments (e.g. the possibility of including time delays), this narrow focus contrasts with an evolving discussion of strategy mapping and its related causal mapping in general in management and operations research (Hodgkinson and Clarkson, 2005), which has not entered mainstream discussions of the tool within performance management. Rather, discussions of the strategy map as it appears in performance management remain bound to the balanced scorecard framework, which, it should be noted, appears to be on the decline (Rigby and Bilodeau, 2018). Therefore, if the map is to reach its breakthrough potential for performance management, it is useful to consider it separately from the balanced scorecard.

Therefore, the aim of this paper is to revisit a major component of performance management, the strategy map, to thoroughly consider the theory of how they work, and further consider this within a unique performance management context. There are two intended contributions through this aim: first, specifying purpose and extracting theory can help practitioners better fit them to purpose and allow maps to be employed more effectively. This synthesis addresses this aim specifically by offering several propositions inferred from the review results. Second, it aims to permit performance management research and practice to be able to adapt, adjust, and expand existing and emerging theory on maps and mapping beyond that offered in the original balanced scorecard framework. In other words, instead of whether strategy maps “work,” the interest of this study is to develop an understanding of the generative mechanisms behind strategy maps:

RQ1. How and in what circumstances do strategy maps contribute to increased organizational performance?

The objective of this paper is to address the research question through a realist synthesis (Pawson, 2006) of empirical studies on the use of strategy maps as a part of a performance management framework. A realist synthesis is a type of systematic literature review that focuses on developing a theory of how a particular tool, framework, program or intervention is meant to work, and then examines the evidence to evaluate the strength of the theory. Because it focuses on theory rather than the tool itself, it is well-suited for evaluating complex interventions like the use of strategy maps, in which there may be multiple, conflicting factors influencing its outcomes. The idea is that by separating the theory from the tool, realist synthesis can facilitate knowledge creation and make it easier to adapt its use to a particular context.

The paper proceeds as follows: first, it explores realist synthesis and the methods of review. Next, results are presented, and then discussed along with implications for research and practitioners.
2. Methodology
Most interest around the strategy map within performance management has maintained Kaplan and Norton's focus on the technical aspects of strategy maps (see Islam, 2018, for a recent review of these) to the detriment of the sensemaking processes that take place around them. Underlying this focus is a common position within performance management studies that the interpretation of performance information is straightforward, linked to positivism (Micheli and Mari, 2014). These assumptions can be problematic when considering the social aspects of performance management (Beer and Micheli, 2018), a criticism that has been applied to strategy maps (Modell, 2012).

Therefore, a potentially fruitful means of understanding how maps work is to also revisit the philosophical assumptions upon which considerations of the strategy map in performance management have been built.

This paper describes a realist synthesis (Pawson, 2006). In practical terms, the method begins with a guiding question: “What works for whom under what circumstances, how, and why?” (Wong et al., 2013). Underlying this question is a realist philosophy of science, which will be briefly discussed in the following paragraphs as a backdrop to the synthesis method.

2.1 Why realism?
Scientific realism developed largely in response to a criticism that traditional research approaches were limited in their ability to provide explanations because they relied on artificially creating or assuming closed experimental conditions (Sayer, 1992). In most cases, experimental closure is undesirable or impossible because reality is fundamentally open (Bhaskar, 1975). This openness quickly comes into conflict with the more commonly employed Humean view of causality which seeks to establish scientific laws by seeking events in succession (Hume, 1967).

Under this empiricist approach, reality is seen as obeying universal laws which can be uncovered through the repeated observation of events. Researchers can then induce the existence of these laws, which can then be tested via statistical methods to establish their validity.

However, scientific practice under the empiricist approach has been criticized because it effectively reduces reality to observable events. In social systems, this position has been cited as especially problematic because it allows for the meaningfulness of social interactions to be completely ignored or greatly reduced (Bhaskar, 1979).

As an alternative, realism adopts a generative view of causality under which cognitive, social and physical entities interrelate to produce events via mechanisms. The primary aim of science under this perspective is to identify these mechanisms and understand their nature in order to improve practice (Bhaskar, 2014, p. v). However, disagreements exist on the meaning of the term "mechanism", which have complicated its application in practice (Dalkin et al., 2015), and so some further clarification is needed.

First, mechanisms are described as the generally unobservable relations between processes, physical and social structures, and ideas that produce outcomes (Astbury and Leeuw, 2010; Mingers and Standing, 2017), which may operate in different contexts in which other mechanisms may be operating simultaneously. Because of the focus on how mechanisms operate in particular contexts to produce outcomes, realist evaluation often reports results in a “CMO” configuration for context, mechanism and outcome (Pawson, 2013). However, several researchers have pointed out continued confusion on what constitutes a mechanism and what does not (Craver, 2009; Dalkin et al., 2015; Mingers and Standing, 2017). This discussion adopts the view of Mingers (2014), in which the mechanism explains the relation between the entities within a system that gives rise to the outcome of interest.

Before illustrating the concept of mechanism used here, it is important to note that from the realist perspective, mechanisms operate in a stratified reality (Astbury and Leeuw, 2010;
Bhaskar and Danermark, 2006). There are a number of ways in which realists conceive of stratification (Bhaskar, 2010), but what is important here is the concept of emergence, i.e. that the properties of an entity cannot be reduced to any one of its components, but rather emerge from their interaction.

An example using a matchstick can help to illustrate these concepts. At one level, the combination of its chemical composition and the friction of the surface create a process of combustion which, given the right conditions (e.g. the presence of oxygen), will produce a flame. Chemical composition and combustion is the mechanism that explains the outcome of the flame but provide part, but not all of the explanation. For example, to achieve the generation of the flame matches generally cannot be lit under water. Neither will the flame be produced if the wrong technique is used: too much pressure, and the matchstick breaks. Too little, and there will not be enough friction for the reaction to take place.

This type of analysis is open to higher-order considerations such as why the match might be struck in the first place, or the systems of production and infrastructure that could explain its existence. It also includes an interest in secondary outcomes: light a match on an airplane, for example, and the interrelation of various social structures will likely result in the person’s arrest – an emergent outcome which cannot be explained through the match’s chemical properties alone and requires understanding how people make sense of the action.

2.2 Why realist synthesis?

Adopting a realist approach to discovery has several implications for how research is carried out and, importantly, how evidence is cumulated and synthesized. Critically, rejecting a view of causality based on events implies that traditional forms of systematic literature review (Tranfield et al., 2003) require revisiting.

Systematic literature review originated in the field of medicine as means of consolidating existing knowledge. These reviews were meant to increase rigor over traditional, narrative reviews through transparency, inclusivity, and a focus on explanation (Denyer and Tranfield, 2009). Realist synthesis adopts many of the elements of these reviews, but requires adapting explanations into the generative view, adopting a more flexible approach to evidence gathering and to collection, and by abandoning the traditional hierarchy of evidence in evaluation. These elements and their implications will be discussed below corresponding with the stages of review, but essentially realist syntheses involve two processes: extracting the theories of how a particular intervention works (the mechanisms) via abductive redescription or abstraction, and evaluating the strength of those theories through a critical examination of the studies uncovered through the search processes.

The following section describes the stages and methods of review, which following Pawson (2006) include identifying a topic, extracting theory, search for literature, selection and appraisal, extraction, analysis and synthesis.

2.3 Identifying the topic of review

The interest of this discussion is in extracting the theory of strategy maps within a performance management context, where with few exceptions, strategy maps are discussed as a part of the balanced scorecard framework. Here, a scoping study revealed generally vague descriptions of how the strategy maps were meant to work, corroborating observations of much literature on the balanced scorecard in general (Hoque, 2014). Therefore, it was thought that a focus on strategy maps would have the greatest potential impact for practitioners and also would benefit performance measurement theory building.
2.4 Extracting the theory of strategy maps within a performance management framework

In a realist synthesis, how an intervention is meant to work often needs to be interpreted or adapted to fit the realist ontology. Even if some research implicitly uses a generative model of causality, few are described initially in such a way (Wynn and Williams, 2012). Others may be useful for evaluating the effectiveness of maps but focus on outcomes whose primary interest is not the direct improvement of organizational performance, e.g. for conflict resolution (Ackermann et al., 2016).

Therefore, a scoping study served to develop an initial classification of potential mechanisms using the foundational texts of the balanced scorecard (e.g. Kaplan and Norton, 2001, 2004, 2006), practitioner resources on the topic (Balanced Scorecard Institute, 2017) and reviews on causal maps and strategy maps (Hodgkinson and Clarkson, 2005; Lueg and Julner, 2014). Theories resulting from the scoping study were refined as the study progressed through a process of abstraction or abductive redescription – in other words, describing how the maps were meant to work in uniform terms to fit performance management.

These were grouped according to their associated performance measurement stage, whether to structure problems, develop, implement or modify a performance management system, or for use as an analysis or communication tool. During the search process, the background section of each study included in the full-text review was evaluated to extract the theory, if present, of how the strategy map or mapping process was meant to work.

The mechanism theory, presented in Section 3, was further divided into hierarchies depending on level, such that the lowest involved largely psychological processes, and the highest considered organizational outcomes. This process and its implications will be explored in the discussion section, but centered on examining how maps could affect organizational properties via the actions of many individuals (Astbury and Leeuw, 2010).

2.5 Search processes

Figure 1 shows an outline of the process for the synthesis. The search for studies to evaluate the propositions began with keyword searches for “performance measurement” in the academic citation databases of Scopus and Web of Knowledge, and later expanded to include “causal map” and “strategy map.” The searches were intentionally broad to increase the likelihood of including relevant articles in the review. That search began with keyword searches of the Scopus and Web of Knowledge academic databases, resulting in 6,583 unique articles. Additional text filters resulted in 4,225 articles for title and abstract review. The review relied heavily on the snowball approach, following Denyer et al. (2008), where references of each selected article were searched for relevant evidence.

2.6 Selection and appraisal of evidence

For the purposes of this review, the definition of performance measurement came from Franco-Santos et al. (2007), who argue that a performance measurement system exists if there are processes of measure design and selection, data capture, and information provision, features performance measures and supporting infrastructure, and has the role of measuring performance. This definition was selected because it encompasses only the necessary conditions of a performance measurement system, and would allow for a wide range of texts to be included.

Selection criteria:

- addresses performance measurement or management in organizations;
- describes an empirical study;
explores the consequences of the use of strategy or causal maps for either structuring problems, developing performance measures, communicating performance or analyzing performance;

journal is included in the Scopus Citations Index or Journal Citations Report;

article is published between 1992 and 2017; and

results in English.

Selection criteria were applied in stages. Titles and abstracts were reviewed separately to exclude only those articles that did not meet the selection criteria. Articles with the possibility of relevance were passed on for further review and were considered relevant if they could be used to evaluate the developing program theory.

Articles that met all the inclusion criteria that were published in peer-reviewed journals were included, though not all impacted the final synthesis to an equal extent. For example, though the study by Cugini et al. (2011) on the application of strategy maps in a university setting provided an example of a successful implementation, the study mainly focuses on describing the resulting strategically linked scorecard, offering little evidence for evaluating underlying causal mechanisms. On the other hand, studies were also evaluated if they were considered to have sufficient rigor and relevance but were not in either citation index, though only one, that of Vo et al. (2005), was included in this fashion.

Application of the selection criteria resulted in 52 studies which were included in the final review. Of these, more than 60 percent were featured in journals with a 2017 SCIMago
Journal Rank in the first quartile, with over a third of the studies in three and four-star journals in the 2018 ABS Academic Journal Guide, both common means of establishing quality (e.g. Franco-Santos et al., 2012).

2.7 Extraction
An extraction form was used to categorize the proposed mechanisms, context, subject, intervention characteristics, and an assessment of relevance and rigor of each of the studies. As it became clear which factors were of particular interest, the extraction form was refined to include the new information, and studies which had been previously examined were examined again to consider any new information. This reflects a recognition that database protocols may need more flexibility in studies on organizations than in the context of evidence-based medicine (Tranfield et al., 2003).

2.8 Analysis and synthesis process
Unlike traditional systematic review, the process of analysis and synthesis takes place alongside assessing relevance and extracting data. Following Pawson (2006) and Wong et al. (2013), full texts were reviewed and analyzed. The logical mode for this process is referred to as abstraction by Pawson (2006) and abductive redescription by Bhaskar (2016), i.e. describing events in a theoretically significant way. The result is an evolving “mechanism sketch” (Craver, 2006), a baseline categorization of the critical features, processes and actors that can explain how strategy maps generate the outcomes of interest.

This baseline, and another key part of the synthesis process, comes from comparing and contrasting findings from the included studies to infer a likely explanation, so that relevant findings could be used to develop specific propositions. Though not discussed specifically by Pawson (2006), the process could be thought of as inference to the best possible explanation (Lipton, 2004). It is important to note that first, the same study may support one proposition while not another. Because the focus is on generative mechanisms, studies may also inform the evaluation of more than one proposition or mechanism. In this way, the findings of these studies were used to evaluate the mechanisms that were derived in the process of abstraction.

3. A theory of maps for performance management
Performance management refers to a wide range of processes which center on setting goals, defining performance measures, reviewing and acting upon performance data, and the activities that surround these, with the ultimate goal to improve organizational performance (Bititci et al., 2018). Strategy maps have been implicated in any number of these activities, but broadly, their use can be seen as addressing three separate but interrelated performance management stages or processes. These stages can be to structure problems, generally in the form of strategy formation, to select, define, modify or develop an existing performance management component or system, or to communicate, analyze or evaluate performance, here referred to as use. It should be noted that studies within performance management rarely distinguish between these different purposes, which, as will be discussed, have complicated research into strategy maps.

The following section explores how maps are seen to drive the desired positive outcomes of each stage. This theory is the result of abstraction described in the previous section, and its purpose is to provide a high-level framework that facilitates the evaluation of results. Alluding again to the match example where combustion provides a baseline explanation for how a match generates a flame, this section aims to find a baseline explanation as to how a strategy map would generate its outcomes.
A summary of the articles included in this review can be found in Table AI which includes the citation, the methodological approach, propositions addressed, research context, the type of strategy map, its complexity, elicitation technique and, if appropriate, the method of its development.

3.1 Strategy mapping for problem structuring

Strategy maps within performance management were originally presented as a way of “describing strategy” in order to understand it (Kaplan and Norton, 2000). This statement highlights that mapping for structuring problems is an active process which aims to facilitate the generation of ideas, gaining a broader understanding, and ultimately pursuing a more effective strategy. Within management studies, mapping has been used to achieve a wide range of ends. Of interest to this review are the mechanisms that explain how the creation of maps work for strategy formation and execution for an individual, in groups, and finally how these can lead to the pursuit of a more effective strategy and increased organizational performance.

3.1.1 The outcome: what is a structured problem?

Broadly, when exploring outcomes for individuals, these studies are concerned with gaining a deeper understanding of an issue. Understanding is discussed as task performance (Öllinger et al., 2015), new knowledge or ideas (Goodier et al., 2010), presenting a diverse range of concepts (Goodier and Soetanto, 2013), or complexity of maps presented (Xu, 2011).

There is also an interest in how participants perceive the strategy or strategy making process, which is often pursued in tandem. For example, mapping can be used for changing how people feel about the strategy itself, whether by allowing their views to be heard, by separating the ideas from the speaker and from the motivational effects these can generate (Ackermann and Eden, 2011). Because of the potential, mapping is used for consensus building and conflict resolution (Ackermann and Eden, 2005; Ackermann et al., 2014, 2016). Ultimately, within performance management, the outcomes discussed above are meant to facilitate the pursuit of a more appropriate or effective strategy (Goodier et al., 2010; Jenkins and Johnson, 1997). A full list of outcomes for structuring found in this review is included in Table I.

3.1.2 How are maps meant to help structure problems?

Figure 2 presents the mechanisms that were found in the literature that would explain how strategy maps can generate learning, motivation, ownership and, ultimately, the pursuit of a more effective strategy—the outcomes sought through their use as a tool for structuring problems. These outcomes correspond to three levels that have been abstracted from the literature: a psychological level whose outcomes are understanding and motivation, a group or social level where, in addition to reaching a shared, broader understanding, there can positive changes in attitude, and finally, the generation and selection of an appropriate course of action at the organizational level.

For the individual, maps are meant to lead to understanding by functioning as a kind of mirror, a process referred to here as actualization. By creating a map, the mapper makes ideas about an issue explicit, and thereby can see and reflect upon them. Eden and Ackermann (2018) refer to the map in this process as a “transitional object.” The nature of the knowledge created and how actualization works have been debated extensively (see Hodgkinson and Clarkson, 2005 for an overview) but remain outside the scope of this paper. What is important is that the node-link structure of causal maps specifically is a key component because it allows seeing, reflecting upon and possibly modifying how ideas relate to one another (Eden, 1988).

Groups can achieve consensus or shared understanding, more holistic views of an issue and have more ideas presented in several ways. First, through the actualization process, participants are able to avoid embarrassment and “save face” (Eden, 2004), participate more, and also perceive the process as fair. As a result, participation, motivation and ownership of the strategy formation process increase. This mechanism is referred to here as inclusion.
<table>
<thead>
<tr>
<th>Source</th>
<th>Summary of findings</th>
<th>Task setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ollinger et al. (2015)</td>
<td>Causal mapping in structuration → improved decision making during subsequent task (NS), E</td>
<td>Individual</td>
</tr>
<tr>
<td>Montemari and Nielsen (2013)</td>
<td>Aiding in the process of creating a causal map → increased understanding of intangible assets (+), C</td>
<td></td>
</tr>
<tr>
<td>González et al. (2012)</td>
<td>Described elicitation using “repertory grid technique” → goal clarity, understanding of organizational goals (+), C</td>
<td></td>
</tr>
<tr>
<td>Aranda and Arellano (2010)</td>
<td>Eliciting a causal map → Understanding of strategy (+), E</td>
<td></td>
</tr>
<tr>
<td>Tegarden et al. (2010)</td>
<td>Anonymity during individual map creation → range of concepts presented, understanding of how to achieve goals (+), P</td>
<td></td>
</tr>
<tr>
<td>Pinch et al. (2010)</td>
<td>Freehand mapping process → revealing issues about which mappers are not aware (NS) Freehand mapping process → understanding of organizational and social context of design thinking (+), P</td>
<td></td>
</tr>
<tr>
<td>Tegarden et al. (2009)</td>
<td>Mapping with framing statements and anonymity → ease of arriving at a shared vocabulary, understanding of strategy (+), C</td>
<td></td>
</tr>
<tr>
<td>Kunc (2008)</td>
<td>Application of systems thinking to develop strategy maps → mental model accuracy (+), E</td>
<td></td>
</tr>
<tr>
<td>Vo et al. (2005)</td>
<td>Involvement in mapping session → subjective assessment of the map for evaluation of performance (+)/*, E</td>
<td></td>
</tr>
<tr>
<td>Hodgkinson et al. (2004)</td>
<td>Pairwise elicitation technique for map creation → map complexity (+), perceived effort (NS) Freehand elicitation technique for map creation → map complexity (+), perceived effort (NS), E</td>
<td></td>
</tr>
<tr>
<td>Jenkins and Johnson (1997)</td>
<td>Complexity of elicited map → firm performance (+), C</td>
<td></td>
</tr>
<tr>
<td>Cossette and Audet (1992)</td>
<td>Elicitation using indirect and direct techniques → learning outcomes of mapping (+), C</td>
<td></td>
</tr>
<tr>
<td>Langfield-Smith (1992)</td>
<td>Elicitation using separate questioning, card sorting and feedback interviews → success in creating map (+), P</td>
<td></td>
</tr>
<tr>
<td>Ackermann and Alexander (2016)</td>
<td>Use of mapping in conjunction with group support software → In group conflict resolution (+), C</td>
<td></td>
</tr>
<tr>
<td>Ackermann et al. (2014)</td>
<td>Use of mapping with group support system → understanding (+), holistic view of the problem (+), P</td>
<td></td>
</tr>
<tr>
<td>Francioli and Cinquini (2014)</td>
<td>Process of creating, reviewing, and discussing strategic linkages → successful development and use (+), C</td>
<td></td>
</tr>
<tr>
<td>Parisi (2013)</td>
<td>Using more than one elicitation technique for mapping → avoidance of confirmatory bias (+), elicitation of tacit knowledge, favorable development outcome (+), A</td>
<td></td>
</tr>
<tr>
<td>Montemari and Nielsen (2013)</td>
<td>Aiding in the process of creating a causal map → actor’s increased understanding of complex network (+), C</td>
<td></td>
</tr>
<tr>
<td>Goodier and Soetanto (2013)</td>
<td>Hand drawn mapping followed by group support system software for map creation → understanding of issues relevant to the mapping session (+), inclusion of viewpoints (+), P</td>
<td></td>
</tr>
<tr>
<td>Gouttenoire et al. (2013)</td>
<td>Causal mapping process with group support system and effective facilitation → self-reflection, understanding of issues, and interest of participants (+), P</td>
<td></td>
</tr>
<tr>
<td>González et al. (2012)</td>
<td>Individual mapping sessions followed by group comparison → aligning managers’ perceptions of organizational strategy (+), C</td>
<td></td>
</tr>
<tr>
<td>Cugini et al. (2011)</td>
<td>Collaborative approach to elicitation → successful development of a strategy map (+), E</td>
<td></td>
</tr>
<tr>
<td>Van den Bossche et al. (2011)</td>
<td>Having a shared mental model → group task performance (+), E</td>
<td></td>
</tr>
<tr>
<td>Xu (2011)</td>
<td>Social interaction → resulting map complexity (+) Feelings of psychological safety → resulting map complexity (+) Modulating effect of gender (NS), E</td>
<td></td>
</tr>
</tbody>
</table>

Table I. Studies addressing problem structuring
Second, the visual mapping process allows participants to “piggy back” (Shaw et al., 2009) off one another’s ideas, and so the process has a self-referential effect. This mechanism is referred to here as reinforcement.

The ideas generated through mapping provide multiple alternatives for action beyond those of other techniques, and so allow decision makers to choose a more appropriate course of action through the increased understanding gained through mapping. This mechanism is referred to here as choice.

Figure 2 also includes a number of components which condition whether and the extent to which actualization will take place. These will be considered further when evaluating the evidence but can be divided roughly into the characteristics of the mapper and their environment, including the nature of the problem. As will be discussed, in groups and for the organization these are especially important for explaining (lack of) outcomes.

### 3.2 Mapping for system development

For the current discussion, “development” refers to processes that aim to alter the state of an existing performance measurement or management system and is meant to include both implementation of a new system and adaptation of existing ones. Within performance

<table>
<thead>
<tr>
<th>Source Summary of findings</th>
<th>Task setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aranda and Arellano (2010) Creating a strategic map as a group → mutual understanding within a management team (+), E</td>
<td></td>
</tr>
<tr>
<td>Goodier et al. (2010) Elicitation through future scenario building → group think during group mapping (−), engaging participants, understanding of relevant issues and implications of decisions, P</td>
<td></td>
</tr>
<tr>
<td>Tegarden et al. (2010) Previous individual, anonymous mapping session → bringing underlying issues to the surface (+), P</td>
<td></td>
</tr>
<tr>
<td>Lucianetti (2010) Using a strategy map → increasing participation of top management in strategy formation (+), S</td>
<td></td>
</tr>
<tr>
<td>Shaw et al. (2009) JOURNEY mapping process with anonymity → broader understanding, inclusion, and synthesis of ideas (+), P</td>
<td></td>
</tr>
<tr>
<td>Niebecker et al. (2008) Use of impact matrix for creation of strategy map → successful building of map (+)*, P</td>
<td></td>
</tr>
<tr>
<td>Vo et al. (2005) Aggregation technique for creating strategy maps → map complexity (+), E</td>
<td></td>
</tr>
<tr>
<td>Ackermann and Eden (2005) Anonymous, software supported mapping process → defensiveness (−), learning (+), inclusion of ideas (+), P</td>
<td></td>
</tr>
<tr>
<td>Craig and Moores (2005) Discusses special difficulties for family firms in initial structuring with strategy maps and uses a scale (F-PEC) to address this difficulty, C</td>
<td></td>
</tr>
<tr>
<td>Shaw (2004) Software supported causal mapping → creating and sharing new ideas (+), understanding of issues (+), P</td>
<td></td>
</tr>
<tr>
<td>Cossette and Audet (1992) Elicitation using indirect and direct techniques → organizational performance through individual action (+), C</td>
<td></td>
</tr>
<tr>
<td>Langfield-Smith (1992) Describes a mapping session in which participants were unable to create a shared map. Cites power dynamics, lack of shared experiences, and design in the mapping process as possible contributing factors, P</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** “+”: signifies a positive relation; “−”: a negative relation; “*”: partially supported; NS: non-significant results. Methods: C: case study; P: participatory workshop(s); E: experimental design; S: survey; A: action research

**Table I.**

**Source:** The author
management, there is clear interest in using maps for system development and in developing maps themselves (Bourne and Bourne, 2011; Kaplan and Norton, 2004).

3.2.1 What outcomes are sought for development? Generally, the outcome sought during development is selecting or creating an “appropriate” measure, or more broadly, creating a more effective performance measurement system. The terms “appropriate” and “effective” are dependent on their context and take on different meanings in the studies in this review but drew on performance management literature. For example, Lucianetti (2010) investigates the use of strategy maps for translating strategy into operational goals, for adopting new performance measures, and for making cause and effect relationships between measures explicit. Drawing on Neely et al. (1995), Montemari and Nielsen (2013) seek measures that are related to specific goals, controllable, have an explicit management purpose, reflect system causality and provide vision. Studies also seek coherence, completeness, a balance of measures (Cugini et al., 2011; Parisi, 2013) or consensus as to the appropriateness of the included measures (Aranda and Arellano, 2010; Francioli and Cinquini, 2014).

3.2.2 How do maps help develop performance management systems? Development generally discussed either as an extension of the structuring process (Aranda and Arellano, 2010; Parisi, 2013). That is, mapping is meant to assist with the selection or measures or with the attribution of value. In effect, strategy maps help answer “what do we measure?” (Montemari and Nielsen, 2013), either by actualizing the idea, or by providing a sufficiently broad vision of the organization, thus increasing the likelihood that appropriate measures are chosen to be developed and included, or that other performance management system components are adapted to align to strategy.

3.3 A theory of strategy maps for use
Within performance management, the potential for maps for communicating and effectively analysis of organizational strategy and performance has been widely discussed (Francioli and Cinquini, 2014; Kaplan, 2012; Norreklit et al., 2012). Rather than centering on the process of mapping, this discussion begins when a map has already been formed and codified. The typical form this takes within performance management is a hierarchical map, sometimes arranged into perspectives following the balanced scorecard, of a limited number
of performance measures (Kaplan and Norton, 2004). The following sections will consider what these reports have been used to achieve, and how they are meant to achieve it.

3.3.1 What outcomes are sought through use? Strategy maps have primarily been discussed within the context of diagnostic and interactive use (Simons, 1995). That is, there is an interest in evaluating the extent to which the organization has been effective or efficient in its pursuit of the strategy (diagnostic), and also in evaluating the extent to which the current strategy is appropriate (interactive). The interest within performance management centers around how maps can lead to better understanding and decision making, and ultimately to increased organizational performance. For an individual evaluating a map-style report, this review is concerned with how strategy maps effectively communicate performance relative to other types of communication.

Operationalized, the aim of using a strategy map for evaluation can be categorized broadly as enabling improved decision making for the individual, and for the organization consensus, collaboration and double-loop learning (Argyris, 2010). A list of outcomes of interest included in this review is included in Table III.

3.3.2 How do maps work for use? How maps are meant to bring about the outcomes described above can be separated into mechanisms explaining improved decisions making at an individual level and the organizational level. For the individual, given the way the mind works, that the node-link structure is appropriate for use, helping to reduce cognitive load and at the same time allowing the inclusion of a more representative depiction of reality (Frederiksen et al., 2011). This mechanism is referred in Figure 3 as processing. There is some discussion that suggests that communicating and analyzing strategy maps facilitates understanding and empowerment, which facilitate organizational learning, consensus and strategic alignment (Kaplan and Norton, 2004; Kaplan and Norton, 2006). Because these discussions revolve around both evaluating the extent to which a given strategy has been achieved and also evaluating the appropriateness of the strategy itself, this mechanism is referred to here and appears in Figure 3 as evaluation.

4. Evaluating the evidence

The previous section has outlined how strategy maps are meant to work within a performance management context. However, in explaining how a match produces flame, what is also needed is to understand key conditioning components that would explain whether a given attempt will produce a flame or not. Therefore, the following section

![Figure 3. Mechanisms and conditioning structures of maps for use](image-url)
evaluates both the strength of the evidence for the mechanisms presented in the previous section, along with the critical conditions, elements and components that determine whether the desired outcome is realized. In total, 12 propositions are inferred from these observations to help researchers and practitioners better fit existing theory on strategy maps and mapping to the needs performance management.

4.1 The evidence: strategy mapping for problem structuring

The previous section puts forth that the process of creating a strategy map works through actualization, inclusion, reinforcement and by offering choice. The articles in Table I address strategy maps or mapping for problem structuring, and these provide the evidence with which the mechanisms can be evaluated, along with observations of conditioning factors.

It should be noted that research is supportive of the potential for mapping for learning purposes, but that positive outcomes are not guaranteed. The elements that condition successful outcomes – the firing of mechanisms, in realist synthesis terms – can be grouped into individual and group characteristics, environment and guidance (also included in Figure 2).

4.1.1 Participant characteristics. First, the characteristics of the person doing the mapping conditions the extent to which learning will occur. Öllinger et al. (2015) provide evidence supporting the idea that creating a map requires a deal of effort, which will be greater for those who lack experience. The properties of the resulting map also appear to be linked to role (Pinch et al., 2010; Tegarden et al., 2009) and industry (Pinch et al., 2010), where, generally, greater familiarity is linked to the development of more complex maps:

\[ P1. \] Mapping will be less effective for learning for those with low subject-matter familiarity.

When undertaken as a group, differences in age, experience, background and resulting in unbalanced power dynamics can significantly affect the mapping process (Goodier et al., 2010; Gouttenoire et al., 2013; Shaw, 2004; Vo et al., 2005; Xu, 2011). Langfield-Smith (1992) cites a lack of shared vocabulary as contributing to a failed group mapping attempt among members of the same profession. Importantly, feelings of psychological safety encourage mappers to present ideas, which can be encouraged through the adoption of various techniques to support inclusion (Ackermann and Eden, 2005; Xu, 2011). Therefore, it appears that the greater the group diversity, differences in power, culture or language, the more difficult it will be to synthesize ideas. These complications are important because the type of social interaction produced in mapping sessions is critical, with evidence of constructive conflict and inclusion of ideas as being particularly important to achieving positive group outcomes (Ackermann et al., 2014; Shaw et al., 2009; Van den Bossche et al., 2011):

\[ P2. \] Diverse groups which view an issue in different ways will have more difficulty achieving consensus.

However, several studies highlight the potential benefit of multiple possibly conflicting viewpoints (Goodier et al., 2010; Gouttenoire et al., 2013). Therefore, if diversity or opposing viewpoints do not result in exclusion of ideas, results can be beneficial:

\[ P3. \] Diverse groups which view an issue in different ways will produce richer, more complete representations of it.

4.1.2 Guidance. It is well established that the process followed will condition successful outcomes (Ackermann et al., 2016; Langfield-Smith, 1992). Despite detailed discussions of the importance of technique, only one study, that of Hodgkinson et al. (2004), compares two techniques directly and finds significantly greater complexity when possible combinations
of ideas are presented together before they are linked. Other studies include a separate opportunity for generating ideas, either using cards or matrices (Langfield-Smith, 1992; Montemari and Nielsen, 2013), framing statements (Tegarden et al., 2009) or previous interviews (Cossette and Audet, 1992). This suggests efforts to elicit ideas prior to linking them may indeed be beneficial, though research is lacking on the size and significance of comparing techniques. Finally, the questions used to elicit and link ideas are critical (Tegarden et al., 2010). It should be noted that the positive learning outcomes, even at an individual level, were obtained in the presence of a highly trained researcher.

A skilled facilitator with the proper technique may be capable of overcoming the barriers mentioned above, even with highly diverse, conflicting groups (e.g. Ackermann et al., 2016), by taking steps to encourage psychological safety, balance participation and asking appropriate questions:

P4. Guidance results in greater learning to the extent that it helps people to understand mapping, provides a structured, fair process and provides an opportunity for fair participation.

Further, to the extent that group outcomes are achieved through fairness and inclusion, any attempt that fails to address these in a session may not only fail to bring about consensus and group learning but may also make things worse. Langfield-Smith (1992) reports simply failing to reach consensus, but the reinforcing effect in groups and the efforts taken by researchers focusing on structuring in this review suggest the following:

P5. In groups that lack initial consensus and without appropriate guidance, mapping will exacerbate existing disagreements.

4.1.3 Environment. The first concern for performance management is understanding whether strategy mapping is more suitable to some problems over others. The diversity of contexts found in this review (see Table AI) suggests that the applications are wide ranging and include small and large organizations, for profit, non-profit, different levels of experience and career level, and inter-organizational contexts. This suggests that:

P6. Mapping will be useful for structuring problems regardless of organizational context or career level.

The most common means are by providing an opportunity for individuals to generate ideas prior to group mapping (Aranda and Arellano, 2010; Goodier and Soetanto, 2013; Goodier et al., 2010). Prior elicitation also can improve learning outcomes by increasing the number of ideas presented in group sessions (Goodier and Soetanto, 2013). Software-assisted mapping, sometimes in combination with individual idea generating sessions, is another means of facilitating anonymity, and can be used in real time (Ackermann et al., 2016; Goodier and Soetanto, 2013; Niebecker et al., 2008; Shaw et al., 2009; Vo et al., 2005).

These studies highlight that the physical space in which mapping is critical to achieving positive outcomes. Here, software can be beneficial in that it allows maps to be more easily edited in real-time compared to other techniques (Ackermann et al., 2016). However, other studies use physical materials and achieve similar outcomes (Goodier and Soetanto, 2013; Hodgkinson et al., 2004):

P7. Environmental conditions such as physical space or software-assisted mapping will condition learning outcomes.

Finally, there is the question as to whether and in what circumstance the positive outcomes of mapping translate into organizations pursuing a more appropriate strategy, a central idea for Kaplan and Norton (2004). Because no study considered this issue directly, this issue will be explored further in the discussions section.
4.2 Strategy maps for development

These paragraphs explore strategy maps for developing and implementing performance measures and performance measurement systems, a key issue in the successful application of performance management.

Studies contributing to the analysis of the role of strategy maps in development are listed in Table II.

When individuals develop performance measures by creating or helping to create maps, then mapping for development appears to be essentially an extension of problem structuring and works in a similar manner, with similar outcomes. That is, mapping draws attention to the most appropriate measures by effectively representing complex issues, which can then reinforce the idea generation process. Like structuring, researchers note that success will depend on the nature of the phenomenon being measured and on the characteristics of the person measuring (Craig and Moores, 2005; Montemari and Nielsen, 2013). Studies also describe similar steps to foment idea generation and participation such as anonymity, providing time for discussion and revision, and techniques to elicit ideas prior to group sessions with a facilitator (Aranda and Arellano, 2010; Cugini et al., 2011; Niebecker et al., 2008; Parisi, 2013):

P8. The elements of effective problem structuring can be extended to include performance management system development.

<table>
<thead>
<tr>
<th>Source</th>
<th>Task setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Montemari and Nielsen (2013)</td>
<td>Development process using previously developed causal map → identification of appropriate measures (+), C</td>
</tr>
<tr>
<td>Tayler (2010)</td>
<td>Involvement in measurement selection → motivated reasoning (+) Involvement in measurement selection + evaluation with BSC-style strategy map → motivated reasoning (−), E</td>
</tr>
<tr>
<td>Francioli and Cinquini (2014)</td>
<td>Development using informal strategy maps with finality relations between measures → avoiding tensions, costs associated with attempting to validate causal links, successful development and use of resulting balanced scorecard report (+), C</td>
</tr>
<tr>
<td>Parisi (2013)</td>
<td>Use of hybrid map development technique and strategic map → selection of the most important, appropriate measures (+), A</td>
</tr>
<tr>
<td>Montemari and Nielsen (2013)</td>
<td>Development process using previously developed causal map → identification and implementation of measures that are related to specific goals, controllable, have an explicit purpose, reflect system causality and provide vision (+), C</td>
</tr>
<tr>
<td>Cugini et al. (2011)</td>
<td>Collaborative approach to development → Developing a more accurate and complete strategy map (+), E</td>
</tr>
<tr>
<td>Aranda and Arellano (2010)</td>
<td>Communicating strategic links during development → consensus between top and middle management, E</td>
</tr>
<tr>
<td>Lucianetti (2010)</td>
<td>Use of strategy map → translating strategy into operational goals (+) Adopting new performance measures (+) Explicating cause–effect relationships (+), S</td>
</tr>
<tr>
<td>Niebecker et al. (2008)</td>
<td>Use of impact matrix → increased transparency across working groups, identifying relevant performance measures (+)*, P</td>
</tr>
<tr>
<td>Craig and Moores (2005)</td>
<td>Notes difficulty in development varies depending on its category (internal, customer, financial or learning), C</td>
</tr>
</tbody>
</table>

Notes: “+”: signifies a positive relationship; “−”: a negative relationship; “*:” partially supported; NS: non-significant results. Methods: C: case study; P: participatory workshop(s); E: experimental design; S: survey; A: action research

Source: The author

Table II. Studies addressing development
While mapping for development appears to work in a similar way to problem structuring, it must be adapted to the challenges of the development context. For example, studies describe using maps as a means for discussion and arriving at consensus prior to investing in performance reporting infrastructure (Aranda and Arellano, 2010; Francioli and Cinquini, 2014; Montemari and Nielsen, 2013), though these descriptions are limited to systems within the financial industry. However, generally studies that take into consideration the complications that arise during implementation, and the role of strategy maps within these, are lacking.

4.3 Strategy maps for use

The following paragraphs evaluate the use of strategy maps to communicate, analyze and evaluate performance (Table III).

Overall, there is little compelling evidence that similar results cannot be achieved through other, less costly means of communication, when the aim is communication of performance in general. However, some experimental tasks found in this study report small positive effects, and so the conditions that might bring these about will be considered. Figure 3 includes properties of the human mind and of the maps themselves as the most critical conditioning elements in the literature for explaining outcomes.

4.3.1 Conditions for processing information in maps. The argument behind using maps for evaluative tasks is that these are effective at communicating complex information in a way that facilitates understanding because the human mind processes the information effectively (Strohhecker, 2016). However, results are mixed for connecting the use of strategy maps for learning, suggesting that the use of maps for evaluation and communication will be limited compared to use for structuring problems. Several studies showed a small positive correlation between use of strategy maps and learning outcomes (Banker et al., 2004, 2011; Cheng and Humphreys, 2012; Farrell et al., 2012; Frederiksen et al., 2011; Hu et al., 2017; Humphreys et al., 2016; Lowe et al., 2011; Mastilak et al., 2012; Tayler, 2010; Vera-Muñoz et al., 2007).

Individual characteristics that were found to be influential were tolerance for ambiguity (Lowe et al., 2011), education and training (Lowe et al., 2011), and prior involvement in developing the strategy map (Aranda and Arellano, 2010; Tayler, 2010). The results of Carmona et al. (2011) draw attention to interaction effects with the reward structure, where these may amplify behavioral effects of using the maps. This highlights a danger noted in previous discussions (Tayler, 2010) that participation in a report’s design can contribute to motivated reasoning.

There is an interest in connecting the properties of maps to decision-making performance. Here, some evidence suggests that the link-node structure may communicate the importance of non-financial issues compared with other forms of performance reporting (Aranda and Arellano, 2010; Carmona et al., 2011; Lowe et al., 2011). However, two experimental studies returned insignificant results and found that participants ignored strategy map communications to some degree (Humphreys et al., 2016; Rompho and Siengthai, 2012; Strohhecker, 2016). Overall, there appears to be a limit on how effective strategy maps can be during timed decision-making tasks. Several authors attribute this limit to the nature of mental models, proposing that strategy maps may help in their formation to a certain extent, after which they will likely be ignored (Frederiksen et al., 2011; Humphreys et al., 2016; Langley and Morecroft, 2004; Rompho and Siengthai, 2012): P9. Strategy maps will be effective for facilitating initial communication of strategy to groups with low subject-matter knowledge, such as across functional areas.

Concerning map styles that lend themselves to analysis, there were several styles of strategy map described in these studies and these appear to influence outcomes to
<table>
<thead>
<tr>
<th>Source</th>
<th>Principal findings</th>
<th>Map type</th>
<th>Map complexity (nodes)</th>
<th>Task setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Handoko and Wehartaty</td>
<td>Performance information communicated via strategy map → reduced motivated reasoning (+)*, E</td>
<td>Hierarchical</td>
<td>&lt; 10</td>
<td>Individual</td>
</tr>
<tr>
<td>Hu et al. (2017)</td>
<td>Use of map for communicating performance, compared to traditional report → understanding of the performance information(+)*, E</td>
<td>Hierarchical</td>
<td>&lt; 10</td>
<td></td>
</tr>
<tr>
<td>Strohhecker (2016)</td>
<td>Use of strategy map to analyze performance → decision making performance (+) NS, E</td>
<td>Hierarchical</td>
<td>&lt; 25</td>
<td></td>
</tr>
<tr>
<td>Humphreys et al. (2016)</td>
<td>Inclusion of time delays with strategy map feedback → improved decision-making*, E</td>
<td>Hierarchical</td>
<td>&lt; 25</td>
<td></td>
</tr>
<tr>
<td>Cheng and Coyte (2014)</td>
<td>Results communicated with strategy map → propensity for knowledge sharing and extra-role behaviors (+)* (only with subjective incentive scheme), E</td>
<td>Hierarchical</td>
<td>&lt; 10</td>
<td></td>
</tr>
<tr>
<td>Cheng and Humphreys</td>
<td>Information presented in strategy map → decision making (+), E</td>
<td>Hierarchical</td>
<td>&lt; 10</td>
<td></td>
</tr>
<tr>
<td>(2012)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rompho (2012)</td>
<td>Information presented in strategy map → decision making (+) NS, E</td>
<td>Hierarchical</td>
<td>&lt; 25</td>
<td></td>
</tr>
<tr>
<td>Farrell et al. (2012)</td>
<td>Narrative links → improved decision making (+), E</td>
<td>Narrative</td>
<td>na</td>
<td></td>
</tr>
<tr>
<td>Mastilak et al. (2012)</td>
<td>Use of a strategy map → perception of controllability of results (+), E</td>
<td>Hierarchical</td>
<td>&lt; 10</td>
<td></td>
</tr>
<tr>
<td>Booker et al. (2011)</td>
<td>Presentation of narrative information → perception of predictive capacity of measure (+), E</td>
<td>Narrative</td>
<td>na</td>
<td></td>
</tr>
<tr>
<td>Banker et al. (2011)</td>
<td>Information presented in strategy map → decision making (+), E</td>
<td>Hierarchical</td>
<td>&lt; 10</td>
<td></td>
</tr>
<tr>
<td>Frederiksen et al. (2011)</td>
<td>Use of map prior to simulation task → task performance (+)*</td>
<td>Cybernetic</td>
<td>&lt; 25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Use of map during simulation task → task performance (−)*, E</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lowe et al. (2011)</td>
<td>Use of integrated map as decision aid → focus on financial performance (−)*</td>
<td>Hierarchical</td>
<td>&lt; 10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Use of compensatory map as decision aid → focus on financial performance (−)*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Moderated by tolerance for ambiguity (effect of integrated map +), financial background (effect of integrated map (+)), E</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carmona et al. (2011)</td>
<td>Pyramid map type (vs silo) → emphasis on financial results (−), moderated by reward structure (NS) and national culture (NS), E</td>
<td>Hierarchical</td>
<td>&lt; 10</td>
<td></td>
</tr>
<tr>
<td>Humphreys and Trotman</td>
<td>Use of maps for communication of performance results → Reduced common measures bias (when all measures are strategically linked) (+), E</td>
<td>Hierarchical</td>
<td>&lt; 10</td>
<td></td>
</tr>
<tr>
<td>(2011)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tayler (2010)</td>
<td>Information presented in strategy map → motivated reasoning (−)*, E</td>
<td>Hierarchical</td>
<td>&lt; 10</td>
<td></td>
</tr>
<tr>
<td>Laitinen et al. (2010)</td>
<td>Perceived causality of between measures → satisfaction with performance measurement system (NS), S</td>
<td>na</td>
<td>na</td>
<td></td>
</tr>
<tr>
<td>Vera-Muñoz et al. (2007)</td>
<td>Communication of performance with strategy map → decision making (+), E</td>
<td>Hierarchical</td>
<td>&lt; 10</td>
<td></td>
</tr>
</tbody>
</table>

Table III. Studies addressing the use of strategy maps (continued)
some degree. These can be hierarchical, display performance drivers or cybernetic, which contain feedback loops. In addition, there is some evidence that a strategy map may be able to communicate certain types of information, such as feedback loops and time delays, which other types of communications will not (Hu et al., 2017; Humphreys et al., 2016; Strohhecker, 2016). As map complexity increases, satisfaction with the map appears to decrease (Vo et al., 2005), in line with research on information overload and suggestions to limit the complexity in communications (Aranda and Arellano, 2010). There are some limited supports that the combination of the categories of the balanced scorecard together with the strategy map leads to improvement in learning outcomes (Carmona et al., 2011; Lowe et al., 2011): P10. Strategy maps are effective at communicating complex information such as time delays and feedback loops, over traditional performance reports, but overly complicated reports will confuse and possibly frustrate evaluators.

4.3.2 Strategy maps in strategic evaluation. Importantly, the results discussed above come largely from experimental settings where the idea performance is unproblematic, i.e. operationalized and interpreted by the researcher, often operationalized as task

<table>
<thead>
<tr>
<th>Source</th>
<th>Principal findings</th>
<th>Map type</th>
<th>Map complexity (nodes)</th>
<th>Task setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wong-on-wing et al. (2007)</td>
<td>Use of maps for communication and analysis → reduced bias in evaluating performance (+), E</td>
<td>Hierarchical</td>
<td>&lt;10</td>
<td></td>
</tr>
<tr>
<td>Dilla and Steinbart (2005)</td>
<td>Communication of tabular displays or graphs → improved decision making, consensus, and consistency (all NS), E</td>
<td>na</td>
<td>na</td>
<td></td>
</tr>
<tr>
<td>Vo et al. (2005)</td>
<td>Map complexity → satisfaction (−), E</td>
<td>Hierarchical + cybernetic</td>
<td>&lt;25</td>
<td></td>
</tr>
<tr>
<td>Banker et al. (2004)</td>
<td>Communication of performance with strategy map → decision making (+), E</td>
<td>Hierarchical</td>
<td>&lt;10</td>
<td></td>
</tr>
<tr>
<td>Langley and Morecroft (2004)</td>
<td>Strategic map decision aid → long-term learning (NS), E</td>
<td>Hierarchical</td>
<td>&lt;10</td>
<td></td>
</tr>
<tr>
<td>Francioli and Cinquini (2014)</td>
<td>Use of BSC strategy map → strategy execution, communication (+), C</td>
<td>Hierarchical</td>
<td>&lt;25</td>
<td>In group</td>
</tr>
<tr>
<td>Lucianetti (2010)</td>
<td>Use of strategy map for performance analysis → Improving internal communication among people (NS) → Aligning action with strategy (+) → Building consensus around the organization’s vision and strategy (+) → Enhancing time and efforts on strategic-related issue (+) → Making strategy everyone’s day job (+), S</td>
<td>na</td>
<td>na</td>
<td></td>
</tr>
<tr>
<td>Aranda and Arellano (2010)</td>
<td>Communicating strategy through a map, with time for discussion of content and relevance with peers → consensus about strategy (+), effect more pronounced for non-financial performance, E</td>
<td>Hierarchical</td>
<td>&lt;25</td>
<td></td>
</tr>
<tr>
<td>Malina et al. (2007)</td>
<td>Validity of causal relations → improved decision making (NS), C</td>
<td>Hierarchical</td>
<td>&lt;10</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** “+”: signifies a positive correlation; “−”: a negative correlation; “#”: partially supported; NS: non-significant results. Methods: C: case study; P: participatory workshop(s); E: experimental design; S: survey; A: action research. Only variables related to strategy maps are included in this table

**Source:** The author
performance. The maps themselves are generally simple — fewer than ten nodes — conflicting with the more complex maps developed in non-experimental settings (Aranda and Arellano, 2010; Malina et al., 2007) and with those created during problem structuring. Here, the evaluative mode reflects interactive use, where strategy itself can be questioned and adjusted. Given that this type of use mimics the “piggy-back” mechanism of structuring, it is likely that strategy maps would be helpful when used in this manner:

**P11.** Strategy maps will be effective for evaluation when used as a basis for problem structuring and interactive use.

In contrast to interactive use, diagnostic use is periodic or exceptions based, and used primarily for control purposes (Simons, 1995; Tessier and Otley, 2012). In these cases, there is little evidence found in this review to suggest that strategy maps are well suited for this purpose. And yet, most studies on evaluation in this review focus on this type of use. The theoretical foundation of many of these, that the links represent valid causal relations, has been questioned (Norreklit, 2000), but most importantly, no study in this review reported diagnostic control outside of experimental settings in the form of evaluating the validity of links, whether causally or as means–ends relations, and two discuss significant barriers to carrying these out (Francioli and Cinquini, 2014; Malina et al., 2007). Rather, these studies highlight the activity centering around the development and discussion of strategy map reports, in which the causal relations go untested. Further, some authors have suggested (Frederiksen et al., 2011) that strategy maps will be most useful for evaluation if they are processed before they are needed for decision making. These results in combination with longitudinal studies (Aranda and Arellano, 2010; Francioli and Cinquini, 2014) suggest that maps will be ill-suited to the demands of frequent diagnostic use, especially when strategy changes frequently:

**P12.** Strategy maps will be effective for diagnostic use only in environments where strategic change is low, while in dynamic environments they will be overly restrictive.

5. Discussion
The results suggest that the process of creation is what lends the map its power, through its abilities as a tool for reflection, and the learning that can take place when they are created, discussed and revised. Performance management research generally maintains a restrictive view of the strategy maps that presents maps and mapping as a relatively uncomplicated way to achieve strategic alignment and organizational performance, without sufficient consideration of the sensemaking processes around their creation and evaluation. These take place within complex processes of performance management, where the chance of failure is high. Therefore, the following session will discuss the findings to consider how what we know about strategy maps could help reduce the likelihood of failure.

5.1 Strategy maps for problem structuring
Performance measurement begins with forming an idea of what to measure and manage, and strategy mapping seems to present an ideal way to represent and learn about generally complex organizational strategies. This review concentrates on several mechanisms that might bring about this learning, and reflecting on these can guide practice and future research.

The first concerns how mapping can lead to learning for the individual, i.e. the mechanisms that take place largely within the mind when creating a series of nodes and links. Generally, this review supports the observation of Öllinger et al. (2015) that within the performance measurement literature theoretical discussions are underdeveloped.
This synthesis suggests that the benefits can be explained through actualization, inclusion, reinforcement and finally by offering choice.

These are worth considering within the context of performance measurement for the possible benefits of using strategy maps in combination with other elements of performance measurement. For example, the original balanced scorecard report and its requirement for measures to come from multiple categories appear to complement the strategy map building exercise, as the requirement to have a variety of measures from different categories can yield richer, more complete representations of strategy (Hodgkinson et al., 2004). Further exploration of these synergies could result in interesting new lines of research. For example, there is little discussion of possible pitfalls of the strategy mapping process apart from one reported failure caused by disagreement (Langfield-Smith, 1992). Are there situations in which these could do more harm than good?

The discussion of elicitation, setting and group dynamics goes well beyond the typical treatment the process of creating strategy maps receives in performance measurement literature, which often presents maps as if their creation is unproblematic. Researchers and practitioners should therefore be aware of the difficulties in creating strategy maps. Otherwise, they may result in more instead of less disagreement.

5.2 Strategy maps for system development

If development can be explained using the mechanisms of problem structuring as suggested by this synthesis, then the key gap in researching strategy maps for development is studies that consider the many potential sources of failure within the development process, such as those enumerated by Van Camp and Braet (2016). For example, the development process is complex, often includes multiple actors, can take years (Craig and Moores, 2005; Franco-Santos and Bourne, 2003) and may be the most likely stage of failure (Neely et al., 2000). Generally, evidence was supportive of the potential of a strategy map to promote successful development outcomes, but there were few descriptions of the process (Aranda and Arellano (2010) and Francioli and Cinquini (2014) are notable exceptions), whether that was using strategy maps to develop performance measures for use, developing strategy map-style communications, or both.

The evidence in this review suggests that when strategy maps are used as a continuation of the strategic dialog begun during problem structuring, then it is more likely to result in better performance measurement systems. Reviewed texts were generally favorable the effects of participation in development, which coincide with other studies in performance measurement literature (Groen et al., 2012). Participation in development holds the risk of leading to biases (Taylor, 2010). In theory, at least, using the strategy map as a tool for fomenting debate could prevent these biases from unbalancing the measurement system. But researchers and practitioners should be aware that if the benefits of mapping are explained in large part through inclusion, and that this is brought out in part to the extent that the mapping process is seen as fair, then care is needed in how the ideas are implemented so as not to bring about dysfunctional effects (Franco-Santos and Otley, 2018).

Therefore, there is an interesting opportunity for studies that observe strategy maps in the processes of development and implementation specifically to learn more about how they can or cannot help navigate the complex development process, especially for aligning the various elements of performance management systems to organizational strategy.

5.3 Strategy maps for use

Two mechanisms are presented to explain how strategy maps can lead to better decision making and organizational performance. For individuals communicating or analyzing
maps, the power of the map has been described as resting in its ability to show causal
relations and relevant information and so facilitate processing.

The results of this review suggest that more research building on links to cognitive
psychology in the line of Dilla and Steinbart (2005) and Cheng and Humphreys (2012) could
help develop a theory of when strategy maps will be most effective for communication and
evaluation, especially when used diagnostically. As it stands, both the theoretical
explanation and studies demonstrating a map’s practical adequacy are lacking.

More importantly, this review highlights that experimental studies with tightly defined
notions of performance and short time limits are a poor reflection of how maps are used for
communication in organizations, though admittedly this could be due to selection bias or
limitations of the review. Nevertheless, studies focusing on decision-making contrast with
descriptions in the field (Aranda and Arellano, 2010; Francioli and Cinquini, 2014), where the
strategy map serves as a tool for ongoing discussion over long periods of time, and in which
manger participants had the opportunity to analyze, question and importantly to refine the
strategy maps presented to them. It appears again that the strength of maps is not primarily
in their ability to communicate, but rather in their suitability for structuring problems and
developing a balanced, complete measurement system. In this way, they do appear to serve
as a medium for achieving “double-loop” learning and can result in the kind of
transformative outcomes described in Kaplan and Norton (2004), subject to the constraints
and difficulties described for the previous stages.

6. Conclusions
Two decades after its introduction to the field, the strategy map has the potential to represent a
major contribution to contemporary performance management. This review suggests that
separating the strategy map from the balanced scorecard could help it realize its potential as a
breakthrough theory within performance management. Doing so allows the identification of
mechanisms that explain how strategy mapping can facilitate strategy formation, performance
measurement system development, and strategy evaluation and communication, which can
further lead to the development of more effective applications of the concept.

Realizing the potential of the strategy map will require addressing a mismatch between
research focus to date and organizational reality. To fully utilize strategy maps within
performance management, researchers will need to better understand how these feature with
other performance management components. Doing so will require shifting focus from
evaluative tasks for diagnostic use – representing the majority of research on evaluation – to
observing how these function in organizations and how they can support the overall strategic
dialog. Experimental research is helpful for better understanding the behavioral effects of
these maps, and yet they often neglect the difficulty in developing and implementing them for
use in organizations, generally operating in conditions of frequent strategic change (Porporato
et al., 2017). Therefore, a major contribution of this review is to highlight the importance of
differentiating these processes in order to analyze how maps work in organizations.

The second contribution of this review is that it begins to separate the theory of strategy
maps from any particular tool or framework, which in performance management is
generally the balanced scorecard. Through the realist synthesis process, the review offers a
“mechanism sketch” (Craver, 2006), a baseline categorization of the critical features,
processes and actors that can explain how strategy maps generate the outcomes of interest.
Given the realist assumption of openness, the exact way that these features interrelate will
vary from situation to situation, but the mechanism should remain constant.

Further, 12 propositions are offered on how strategy maps will work, for which purpose,
and in what circumstances. Future research within performance management can build
upon these to develop a unique theory of maps that is specific to and useful for the field.
More research is needed to understand, for example, how the use of strategy maps for
Evaluation might lead to unintended, potentially negative impacts when they are combined with existing incentive structures (Cheng and Coyte, 2014; Mastilak et al., 2012), but there is also a need to explore interactions with target setting, defining KPIs, information flows and other performance management components. Doing so opens the possibility of discovering new applications of strategy maps and mapping within performance management.

Separating the theory from the tool is also important because it can help to explain and address failures at different levels. Distinguishing level could help explain why, for example, strategy maps could effectively improve communications across groups, but lead to poor decision making in an individual evaluative task. The view offered here is that understanding the two requires a consideration of largely different levels, one primarily cognitive, the other situated in and conditioned by organizational-level elements. Perhaps most importantly, a focus on how can help the strategy map to establish its own place within performance management study, and to evolve in the rapidly changing organizational context (Bititci et al., 2012).

The review represents one of very few realist syntheses in management studies, though recent calls for more reviews of this type highlight their perceived potential (Jones and Gatrell, 2014). By focusing on the underlying theory of how strategy maps are meant to work, these types of reviews open new lines of questioning that could be of interest to performance measurement and management.

Although the findings are encouraging, the review is limited in several ways. Perhaps most importantly, by taking a broad view of strategy maps across three stages of performance management, nuance has been sacrificed in the analysis of each. While maintaining sufficient breadth is useful for considering strategy maps within performance management at a high level, future studies will be needed to better establish particular configurations of elements that generate outcomes. This is not a call for lists in the form of context, mechanism and outcome, but rather for continued focus on building nuanced explanations of strategy maps.

The findings of this paper are important for practitioners using or considering adopting the use of strategy maps. First, it highlights that creating strategy maps is a highly accessible activity for achieving shared understanding of what organizations do and how they do it, even among diverse groups of stakeholders. What is significant, and distinct from recent reviews (e.g. Islam 2018), is that the process of creation is what drives much of the benefits to be had from the strategy map, and further one that likely requires significantly less investment than many elements of the performance management system. For example, simply attempting to create a strategy map as a group can be a useful exercise that can generate consensus. These benefits can be carried over to develop or implement appropriate performance measures, where they serve as a focus point for discussion to link measures to strategy. Conversely, practitioners should proceed with caution before investing in strategy map-style reports for communicating performance for diagnostic use. Not only are there multiple challenges to developing such reports, but they may also have unintended effects on behavior or simply be ignored.

The original purpose of the strategy map was to describe strategy at a time when intangible assets were being recognized as central to gaining sustainable competitive advantage. In the current global context, characterized by an increasing rate of change, the introduction of disruptive technology and societal shifts, organizations that effectively address complexity will have an advantage over those which cannot (Kelly, 2015). This review suggests that the strategy map is particularly well-suited to addressing this need because of its ability to support consensus building and learning, and therefore could support critical performance management aims in ways that have to date not been fully explored. By considering the theory of how strategy maps work and in what circumstances, both researchers and practitioners alike can move toward realizing the full potential of strategy maps in performance management.
References


Bhaskar, R. (1979), The Possibility of Naturalism: A Philosophical Critique of the Contemporary Human Sciences, Harvester Press, Hassocks.


**Corresponding author**
Ryan Armstrong can be contacted at: hello@ryanarmstrong.me
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Year</th>
<th>Journal</th>
<th>Methodology</th>
<th>Stage</th>
<th>Participants</th>
<th>Map complexity</th>
<th>Map type</th>
<th>Development method</th>
<th>Elicitation technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Handoko and Wehartaty</td>
<td>2017</td>
<td><em>Polish Journal of Management Studies</em></td>
<td>Experimental</td>
<td>USE</td>
<td>Students</td>
<td>Basic</td>
<td>Hierarchical</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hu <em>et al.</em></td>
<td>2017</td>
<td><em>European Journal of Operational Research</em></td>
<td>Experimental</td>
<td>USE</td>
<td>MBA students</td>
<td>Mid</td>
<td>Hierarchical</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ackermann and Alexander</td>
<td>2016</td>
<td><em>Group Decision and Negotiation</em></td>
<td>Case study</td>
<td>STR</td>
<td>Organizations in negotiation</td>
<td>Mid</td>
<td>Hierarchical</td>
<td>Workshop</td>
<td>Direct</td>
</tr>
<tr>
<td>Humphreys <em>et al.</em></td>
<td>2016</td>
<td><em>The Accounting Review</em></td>
<td>Experimental</td>
<td>USE</td>
<td>Graduate students</td>
<td>Basic</td>
<td>Hierarchical</td>
<td></td>
<td>Indirect</td>
</tr>
<tr>
<td>strohecker</td>
<td>2016</td>
<td><em>Journal of Management Control</em></td>
<td>Experimental</td>
<td>USE</td>
<td>Students</td>
<td>Mid</td>
<td>Hierarchical</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Öllinger <em>et al.</em></td>
<td>2015</td>
<td><em>Educational Technology Research and Development</em></td>
<td>Experimental</td>
<td>STR</td>
<td>Undergraduate students</td>
<td>Varied</td>
<td>Hierarchical</td>
<td></td>
<td>Direct</td>
</tr>
<tr>
<td>Ackermann <em>et al.</em></td>
<td>2014</td>
<td><em>European Journal of Operational Research</em></td>
<td>Participatory</td>
<td>STR</td>
<td>Senior managers</td>
<td>Mid</td>
<td>Cybernetic</td>
<td>Workshop</td>
<td>Direct</td>
</tr>
<tr>
<td>Cheng and Coyte</td>
<td>2014</td>
<td><em>Management Accounting Research</em></td>
<td>Experimental</td>
<td>USE</td>
<td>Graduate students</td>
<td>Basic</td>
<td>Hierarchical</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td>Francioli and Cinquini</td>
<td>2014</td>
<td><em>Journal of Accounting and Organizational Change</em></td>
<td>Case study</td>
<td>STR,</td>
<td>Managers</td>
<td>Mid</td>
<td>Hierarchical</td>
<td></td>
<td>Direct</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DEV, USE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goodier and Soetanto</td>
<td>2013</td>
<td><em>Journal of Maps</em></td>
<td>Participatory</td>
<td>STR</td>
<td>UK construction industry experts</td>
<td>Mid</td>
<td>Hierarchical</td>
<td>Workshop</td>
<td>Direct</td>
</tr>
<tr>
<td>Gouttenoire <em>et al.</em></td>
<td>2013</td>
<td><em>Agronomy for Sustainable Development</em></td>
<td>Participatory</td>
<td>STR</td>
<td>Organic farmers</td>
<td>Complex</td>
<td>Cybernetic</td>
<td>Workshop</td>
<td>Direct</td>
</tr>
<tr>
<td>Montemari and Nielsen</td>
<td>2013</td>
<td><em>Journal of Intellectual Capital</em></td>
<td>Case study</td>
<td>STR,</td>
<td>Network of organizations in the</td>
<td>Mid</td>
<td>Cybernetic</td>
<td>Workshop</td>
<td>Indirect</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DEV</td>
<td>same value chain</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parisi</td>
<td>2013</td>
<td><em>Studies in Managerial and Financial Accounting</em></td>
<td>Action</td>
<td>STR,</td>
<td>Pharmaceutical company</td>
<td>Complex</td>
<td>Cybernetic</td>
<td>Aggregate + congrigate</td>
<td>Indirect</td>
</tr>
<tr>
<td>Cheng and Humphreys</td>
<td>2012</td>
<td><em>The Accounting Review</em></td>
<td>Experimental</td>
<td>USE</td>
<td>Graduate students</td>
<td>Basic</td>
<td>Hierarchical</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Year</th>
<th>Journal</th>
<th>Methodology</th>
<th>Stage</th>
<th>Participants</th>
<th>Map complexity</th>
<th>Map type</th>
<th>Development method</th>
<th>Elicitation technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Farrell et al.</td>
<td>2012</td>
<td><em>Journal of Management Accounting Research</em></td>
<td>Experimental</td>
<td>USE</td>
<td>Undergraduate business students</td>
<td>Basic</td>
<td>Narrative</td>
<td></td>
<td>Indirect</td>
</tr>
<tr>
<td>González et al.</td>
<td>2012</td>
<td>Total Quality Management &amp; Business Excellence</td>
<td>Case study</td>
<td>STR</td>
<td>Business executives</td>
<td>Mid</td>
<td>Hierarchical</td>
<td>Aggregate</td>
<td>Hybrid</td>
</tr>
<tr>
<td>Mastilak et al.</td>
<td>2012</td>
<td><em>Journal of Management Control</em></td>
<td>Experimental</td>
<td>USE</td>
<td>MBA students</td>
<td>Basic</td>
<td>Hierarchical</td>
<td></td>
<td>Indirect</td>
</tr>
<tr>
<td>Rompho</td>
<td>2012</td>
<td><em>Measuring Business Excellence</em></td>
<td>Experimental</td>
<td>USE</td>
<td>MBA students</td>
<td>Mid</td>
<td>Hierarchical</td>
<td></td>
<td>Indirect</td>
</tr>
<tr>
<td>Banker et al.</td>
<td>2011</td>
<td><em>International Journal of Accounting Information Systems</em></td>
<td>Experimental</td>
<td>USE</td>
<td>Graduate management students</td>
<td>Basic</td>
<td>Hierarchical</td>
<td></td>
<td>Indirect</td>
</tr>
<tr>
<td>Booker et al.</td>
<td>2011</td>
<td><em>Advances in Accounting</em></td>
<td>Experimental</td>
<td>USE</td>
<td>Graduate management students</td>
<td>Mid</td>
<td>Narrative</td>
<td></td>
<td>Indirect</td>
</tr>
<tr>
<td>Carmona et al.</td>
<td>2011</td>
<td><em>Advances in Accounting</em></td>
<td>Experimental</td>
<td>USE</td>
<td>Executive MBA students</td>
<td>Basic</td>
<td>Hierarchical</td>
<td></td>
<td>Indirect</td>
</tr>
<tr>
<td>Cugini et al.</td>
<td>2011</td>
<td><em>Public Money and Management</em></td>
<td>Experimental</td>
<td>STR, DEV</td>
<td>University department</td>
<td>Mid</td>
<td>Hierarchical</td>
<td>Aggregate</td>
<td>Hybrid</td>
</tr>
<tr>
<td>Frederiksen et al.</td>
<td>2011</td>
<td><em>Learning and Instruction</em></td>
<td>Experimental</td>
<td>USE</td>
<td>Undergraduate Students</td>
<td>Mid</td>
<td>Cybernetic</td>
<td></td>
<td>Indirect</td>
</tr>
<tr>
<td>Humphreys and Trotman</td>
<td>2011</td>
<td><em>Journal of Management Accounting Research</em></td>
<td>Experimental</td>
<td>USE</td>
<td>MBA students</td>
<td>Basic</td>
<td>Hierarchical</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lowe et al.</td>
<td>2011</td>
<td><em>Accounting and Business Research</em></td>
<td>Experimental</td>
<td>USE</td>
<td>Executive MBA students</td>
<td>Basic</td>
<td>Hierarchical</td>
<td></td>
<td>Indirect</td>
</tr>
<tr>
<td>Van den Bossche et al.</td>
<td>2011</td>
<td><em>Instructional Science</em></td>
<td>Experimental</td>
<td>STR</td>
<td>Undergraduate students</td>
<td>Mid</td>
<td>Hierarchical</td>
<td></td>
<td>Indirect</td>
</tr>
<tr>
<td>Xu</td>
<td>2011</td>
<td><em>Gender in Management</em></td>
<td>Experimental</td>
<td>STR</td>
<td>Undergraduate business students</td>
<td>Unknown</td>
<td>Unknown</td>
<td></td>
<td>Direct</td>
</tr>
<tr>
<td>Aranda and Arellano</td>
<td>2010</td>
<td><em>Journal of Management Accounting Research</em></td>
<td>Experimental</td>
<td>STR, DEV</td>
<td>Savings bank employees</td>
<td>Mid</td>
<td>Hierarchical</td>
<td></td>
<td>Direct</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Year</th>
<th>Journal</th>
<th>Methodology</th>
<th>Stage</th>
<th>Participants</th>
<th>Map complexity</th>
<th>Map type</th>
<th>Development method</th>
<th>Elicitation technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goodier et al.</td>
<td>2010</td>
<td>Futures</td>
<td>Experimental</td>
<td>STR</td>
<td>Practitioners, various CEOs</td>
<td>Mid</td>
<td>Hierarchical</td>
<td>Hybrid</td>
<td>Hybrid</td>
</tr>
<tr>
<td>Laitinen et al.</td>
<td>2010</td>
<td>International Journal of Accounting, Auditing and Performance Evaluation</td>
<td>Survey</td>
<td>USE</td>
<td>Mid</td>
<td>Mid</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>Lucianetti</td>
<td>2010</td>
<td>International Journal of Business Performance Management</td>
<td>Survey</td>
<td>STR,</td>
<td>Managers</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>Pinch et al.</td>
<td>2010</td>
<td>Geoforum</td>
<td>Participatory workshop(s)</td>
<td>STR</td>
<td>Design organizations</td>
<td>Mid</td>
<td>Hierarchical</td>
<td>Workshop</td>
<td>Direct</td>
</tr>
<tr>
<td>Tayler</td>
<td>2010</td>
<td>The Accounting Review</td>
<td>Experimental</td>
<td>DEV,</td>
<td>MBA students</td>
<td>Basic</td>
<td>Hierarchical</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td>Tegarden et al.</td>
<td>2010</td>
<td>Accounting Education</td>
<td>Participatory workshop(s)</td>
<td>STR</td>
<td>Academics</td>
<td>Mid</td>
<td>Cybernetic</td>
<td>Workshop</td>
<td>Direct</td>
</tr>
<tr>
<td>Shaw et al.</td>
<td>2009</td>
<td>International Journal of Management and Decision Making</td>
<td>Participatory workshop(s)</td>
<td>STR</td>
<td>Managers</td>
<td>Mid</td>
<td>Hierarchical</td>
<td>Workshop</td>
<td>Direct</td>
</tr>
<tr>
<td>Tegarden et al.</td>
<td>2009</td>
<td>Group Decision and Negotiation</td>
<td>Case study</td>
<td>STR</td>
<td>Executives</td>
<td>Mid</td>
<td>Hierarchical</td>
<td>Hybrid</td>
<td></td>
</tr>
<tr>
<td>Kunc</td>
<td>2008</td>
<td>Management Decision</td>
<td>Experimental</td>
<td>STR,</td>
<td>Students</td>
<td>Mid</td>
<td>Cybernetic</td>
<td>Congregate</td>
<td>Direct</td>
</tr>
<tr>
<td>Niebecker et al.</td>
<td>2008</td>
<td>International Journal of Managing Projects in Business</td>
<td>Participatory workshop(s)</td>
<td>DEV</td>
<td>Managers</td>
<td>Mid</td>
<td>Hierarchical</td>
<td>Workshop</td>
<td>Direct</td>
</tr>
<tr>
<td>Malina et al.</td>
<td>2007</td>
<td>Contemporary Accounting Research</td>
<td>Case study</td>
<td>USE</td>
<td>Fortune 500 company</td>
<td>Basic</td>
<td>Hierarchical</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td>Vera-Munoz et al.</td>
<td>2007</td>
<td>Contemporary Accounting Research</td>
<td>Experimental</td>
<td>USE</td>
<td>Public accountants</td>
<td>Basic</td>
<td>Hierarchical</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td>Wong-owing et al.</td>
<td>2007</td>
<td>Accounting, Organizations and Society</td>
<td>Experimental</td>
<td>USE</td>
<td>MBA students</td>
<td>Basic</td>
<td>Hierarchical</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ackermann and Eden</td>
<td>2005</td>
<td>Group Decision and Negotiation</td>
<td>Participatory workshop(s)</td>
<td>STR</td>
<td>Managers</td>
<td>Mid</td>
<td>Hierarchical</td>
<td>Workshop</td>
<td>Hybrid</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Year</th>
<th>Journal</th>
<th>Methodology</th>
<th>Stage</th>
<th>Participants</th>
<th>Map complexity</th>
<th>Map type</th>
<th>Development method</th>
<th>Elicitation technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vo <em>et al.</em></td>
<td>(2005)</td>
<td><em>Causal Mapping for Research in Information and Technology</em></td>
<td>Experimental</td>
<td>STR, USE</td>
<td>Practitioners, various</td>
<td>Mid</td>
<td>Hierarchical + cybernetic</td>
<td>Aggregate, congregate, and workshop</td>
<td>Hybrid</td>
</tr>
<tr>
<td>Hodgkinson <em>et al.</em></td>
<td>(2004)</td>
<td><em>Organizational Research Methods</em></td>
<td>Experimental</td>
<td>USE</td>
<td>MBA students</td>
<td>Basic</td>
<td>Hierarchical</td>
<td>Indirect</td>
<td>Hybrid</td>
</tr>
<tr>
<td>Jenkins and Johnson</td>
<td>(1997)</td>
<td><em>British Journal of Management</em></td>
<td>Case study</td>
<td>STR</td>
<td>Business owners</td>
<td>Mid</td>
<td>Hierarchical</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td>Cossette and Audet</td>
<td>(1992)</td>
<td><em>Journal of Management Studies</em></td>
<td>Case study</td>
<td>STR</td>
<td>Owner/Manager</td>
<td>Mid</td>
<td>Cybernetic</td>
<td>Hybrid</td>
<td></td>
</tr>
<tr>
<td>Langfield-Smith</td>
<td>(1992)</td>
<td><em>Journal of Management Studies</em></td>
<td>Participatory</td>
<td>workshop(s)</td>
<td>Firefighters</td>
<td>Mid</td>
<td>Hierarchical</td>
<td>Direct</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** STR: problem structuring; DEV: performance management system development; USE: use for communication, analysis and evaluation. *aComplexity is judged by the number of nodes: basic: ≤10 nodes; mid: between 10 and 25 nodes; complex: > 25 nodes*  
**Source:** The author
Competency based superior performance and organizational effectiveness

Sateesh V. Shet
NMIMS University, Mumbai, India
S.V. Patil
School of Management Studies and Research,
KLE Technological University, Hubballi, India, and
Meena R. Chandawarkar
Department of Commerce, Akkamahadevi Women’s University, Bijapur, India

Abstract
Purpose – The purpose of this paper is to explore the relationship between competency-based performance management and organizational effectiveness (OE). It signifies the importance of developing competency-based performance concept in organizations. Since conventional performance management systems (PMSs) are diminishing and as organizations are looking for breakthrough PMSs, this research attempted to fill the gap from stakeholder’s perspective – employee, manager and organization in devising new approach in PMS.
Design/methodology/approach – The research design involved developing scale for “competency-based superior performance” and validating scale for “organizational effectiveness.” The data for this survey are collected from 292 respondents through structured questionnaire. Hypotheses depicting aforementioned relationships were empirically tested in the context of competency-based performance practices in organizations based in India. Structural equation modeling (SEM) technique was used for data analysis.
Findings – The empirical results provide methods to accelerate the performance management initiatives based on a leadership competency model (LCM), which are necessary for building performance culture in the organization. The paper contributes by developing a new scale for measuring competency-based performance practices. The scale for OE is revisited. A positive relationship between competency-based superior performance and OE with productivity, adaptability and flexibility has been empirically confirmed using SEM.
Research limitations/implications – The paper limits the performance measurement concept using leadership competencies.
Practical implications – The developed model will act as a building block for performance measurement in organizations. This paper promotes LCM to be applied in creating a performance-based culture.
Originality/value – This is a unique attempt to test the relationship between competency-based performance management and OE.
Keywords Performance management, Organizational effectiveness, Competency model, Competency-based performance management, Leadership competencies, Organizational performance and productivity, Scale development, Structural equation modelling, PMS

1. Introduction
The last two decades have evidenced a majority of organizations using competency-based people practices in organizations (Vos et al., 2015; Audenaert et al., 2014; Campion et al., 2011). Of them, many used the competency framework for high potential identification using the assessment center development center (AC/DC) followed by learning and development (L&D) plans based on competency gaps (Rodriguez et al., 2002). However, only a few organizations integrated leadership competency model (LCM) with other HR processes, including performance management system (PMS) using critical incident and behavioral event interview (Spencer and Spencer, 1993; Boyatzis, 1982; Gorman et al., 2017).

Since the time of writing Meena R. Chandawarkar has since retired.
The Development Dimensions International study by William Byham found that, for organizations to be effective, integration of a goal-based system and competencies is essential. Byham queried how individuals can increase productivity or cut costs if they are already trying as hard as they can, unless they change their behavior. He concluded that the failure of the goal-based system can be attributed to non-consideration of the competencies and organizations implementing goal-based appraisals communicated the achievement (or lack of achievement) of goals, but failed to communicate why they achieved or failed to achieve them. The dissatisfaction with performance appraisal is at an all-time high (Adler et al., 2016). McGregor (1972) advocated that performance management should be futuristic and not based on the past. Most corporate PMSs do not work nowadays because they are rooted in the models for specializing and persistently optimizing discrete work tasks. These models date back to more than a century ago, to Frederick W. Taylor (Ewenstein et al., 2016). The routine employee conversations on performance and development change the focus to build the workforce competitiveness both today and tomorrow, which is lacking in goal-based evaluations. Cappelli and Tavis’ (2016). McKinsey’s (2014) report on building capabilities found that 35 percent of the organizations believe that leadership capabilities contribute to business performance and organizations are considering capability building as a top strategic priority. In 2016, TCS, one of the world’s largest IT company, abandoned the bell curve system and began focusing on employee feedback and development. This move was started later in 2014, with companies such as Accenture, Microsoft, Google, Adobe, KPMG discarding the annual appraisal process. As a result, organizations introduced a competency-based PMS. The goal of introducing a competency-based approach is to drive business fundamentals with a new performance paradigm and building capability in an accelerated way in working professionals and assisting the organizations in addressing business challenges (Athey and Orth, 1999; Bersin, 2007). However, integrating LCM demands the capability of the HR personnel and the organizational process maturity to build a performance-oriented culture. Therefore, assessing the impact of such competency-based PMS in relation to organizational effectiveness (OE) is vital because LCM is time-consuming, costly and resource-oriented. This research attempts to study the development of the competency-based PMS (as an independent variable) and OE (as a dependent variable).

Goal-based systems do not possess the required capabilities; an increasing number of organizations are inclined to adopt competency-based PMS, as it focuses on sustainability of the achievements based on individual capabilities linked to organizational capabilities. Organizations are concentrating on developing the required competencies that will facilitate the competent personnel in delivering organizational results in a sustained mode, both during progression and recession time. Based on this view, corporations like Capgemini, Infosys, BP switched to the competency-based PMS that provided timely constructive feedback to individuals, thereby developing the requisite competencies that consecutively drive organizational capabilities, which are superior to event-based goal achievement. European countries, such as the UK (Spangenberg et al., 1999), the USA (Boyatzis and Kolb, 1995) and Australia (Cornford and Athanasou, 1995), have demonstrated a significant surge in the use of competency-based approaches.

Hence, competency-based performance management gains high demand and significance because corporations using these practices call for evaluating the procedures. The purpose of this paper is to focus on organizations implementing LCM for performance management practices using competencies. Along with the OE scale, the study developed a new scale for measuring competency-based performance management. Business executives, HR, consulting and academic practitioners will be motivated to speed-up their competency-based performance initiatives for the required organizational competencies.

The ontological stand of this research is that the LCM (used in organizations) can be effectively applied to performance building and measuring its impact with OE.
This empirical study uses an inductive logic for data gathering and structural equation modeling (SEM) for analysis. For developing the scale to measure competency-based superior performance and studying the theoretical framework (Figure 1), positivism as an epistemology for data collection and SEM for analysis is used.

Various methodologies such as goal setting, skill profiling, performance feedback, management by objectives (MBO), 360-degree feedback, ACs, behaviorally anchored rating scale (BARS), balanced scorecard are being used to develop performance measurement at the individual and organizational levels. Competency-based performance management points to the measures involving LCM used in organizations using various aspects of performance building, such as competency-based self-assessment, coaching, recognizing competency gaps, providing feedback, raising the performance bar, L&D plans and competency-based PMS policy and processes. The advantages of competency-based performance management in organizations include:

- raises the performance bar of the employees;
- fosters training using essential capabilities required for that particular role/level;
- develops employees’ proficiencies using a common approach;
- standardizes the required capabilities across geographies;
- aids in structured identification of training needs;
- helps identify the competency gaps, thus formulating L&D plans; and
- defines the leadership capability required at different leadership levels.

Thus, organizations focus on the “foundation of human competence” to enhance employee performance.

2. Research objectives

Competent personnel are the building blocks for nurturing a performance-oriented culture in any organization. Therefore, owning a workforce having the pre-requisite capabilities is vital for effective performance. However, making such an assessment on capabilities is not an easy task. Companies seek to develop measurable, observable systems and tools through which employees’ performance can be assessed and developed (Lucia and Lepsinger, 1999). In other words, a competency-based performance approach should enhance employee
evaluation and development. As a result, both private and public organizations have shown much interest in using competencies as a tool for management. Hence, the objectives of this research are:

- to analyze the impact of competency-based performance management practices on superior performance; and
- to establish the correlation between the competency-based superior performance and OE.

3. Theoretical background
The term “competencies” was first introduced in the 1970s by David McClelland (Boyatzis, 1982; Spencer and Spencer, 1993; Spencer et al., 1994). Competency models consist of knowledge, skills, abilities and other traits necessary for effective performance in an employment position (Pickett, 1998; Lucia and Lepsinger, 1999; Shipmann et al., 2000; Rodriguez et al., 2002; Martone, 2003; Hollenbeck et al., 2006; Levenson et al., 2006; Catano et al., 2007; Vito and Taylor, 2012; Sutton and Watson, 2013; Long et al., 2013). Campion et al. (2011) asserted that a competency model is an important innovation that helps organizations focus on job-related information and personnel skills for managing employees. Parry (1996) and Olesen et al. (2007) stated that these models seek to segregate top performers from average performers. The concept of competency has many facets, but with one common factor: competencies aim at enhancing an individual’s performance at the workplace (Hoffman, 1999).

Employee competencies influence HRD interventions for OE (Potnuru and Sahoo, 2016). Lo et al. (2015) differentiated between strategic and functional HR competencies and believed that strategic HR competencies are superior to functional HR competencies for successful performance in the HR domain. Veliu and Manxhari (2017) concluded that the relationships between independent managerial competencies were observed to be positively significant for business performance. Young and Dulewicz (2009) advocated the approach on leadership selection and development by identifying four “supra-competency” clusters associated with high performance in the British Navy. Halabi et al. (2017) developed and proposed a validated scale to measure intrapreneurial competencies in an organization. Shet et al. (2017a) proposed a framework for review on leadership competencies covering different dimensional aspects of competency studies, including performance management in human resource management. Regardless of the scant evidence in their favor, consultants, employers and strategy-makers continue to have an optimistic view of competencies.

4. Hypotheses development
4.1 Competency-based superior performance (CBSP)
Most studies demonstrate that performance management is the most exploited and controversial HRM system in the organization. This system has created more despondency and loss of faith and confidence than it has given any morale, direction and L&D opportunities to employees. Several attempts have been made by organizations for improving or revising the shortcomings of appraisal to enable employee development, rather than lower the employee morale. Organizations that claim moderate success in their PMS fail to justify how this helps in enhancing performance, rather than simply limiting to assessing and understanding the performance level of employees. Systems like bell curve, key result area dictionaries and frequent interval appraisal also have failed to yield the desired results, as these systems viewed PMS from a stand point of extracting work from employees, rather than making them passionate about their work. By contrast, competency-based performance management can offer an integrated approach where performance measurement is a minor factor, and building
capabilities for sustainable continual superior performance is the key. Superior performance as an outcome can be evaluated through the competency-based PMS (Boyatzis, 1982; Abraham et al., 2001). Several problems associated with the conventional performance appraisal system are also addressed by LCM. Using competencies ensures agreement on the performance criteria (Abraham et al., 2001), accomplishments and gathering of relevant and adequate data. The “gap analysis” approach is used to identify competency needs and further estimate the strength of such managerial competencies in predicting job performance (Lakshminarayanan et al., 2016). Moreover, competency models offer supervisors with the opportunity to observe and analyze behaviors (Winterton and Winterton, 1999), specificity and concreteness in discussions about performance deficiencies, besides assisting in handling huge data in a structured manner (Draganidis and Mentzas, 2006). Thus, the quality of the performance appraisal discussion is enhanced through competency models. The managerial competencies are also associated with developmental opportunities (such as assessment strategies, development approaches and potential gains in organizational commitment and managerial performance) in the manager’s current task and group environments (Frost and Wallington, 2013).

A competency model clearly defines the skills, knowledge and characteristics necessary for the successful implementation of a specific role, thereby providing a feedback platform to the supervisor for discussing the focus areas. Integrating LCM with the performance appraisal process strikes a perfect balance between what gets done and how it gets done because the performance appraisal process concerns results, as well as the behavior and method of attaining those outcomes. Hence, LCM depicts a to-the-point and precise representation of what is relevant and important for effective performance (Martone, 2003). Competency models aligned with the organizational objectives help define the performance criteria to measure the effectiveness and success in a particular role (Spencer and Spencer, 1993), team effectiveness (Maduka et al., 2018) and firm performance (Asree et al., 2010; Tognazzo et al., 2017). Competency, as a measurement tool, facilitates identifying performance-related behavioral factors of a particular job (Whiddett and Hollyforde, 1999). The performance management process becomes robust when employees are appraised on both the objectives (what) and behavioral performance (how) of their role. This approach is referred to as the “mixed model,” which develops a shared understanding of what will be monitored and measured, thereby ensuring an understanding of how the work gets done, in addition to knowing what gets done (Sengupta et al., 2013; Draganidis and Mentzas, 2006). In general, competency believes to be related to the job done and is considered as the least level of achievement to perform that job efficiently (Garavan and McGuire, 2001). The gap between the expected and actual level of competencies can predict future performance and measure individual efficacy and human capital adequacy of a department/organization, thus leading to individual performance (Sebastian and Kumar, 2018). The empirical study by Gabriela (2014) found that, if an employee’s competency is higher than the other employee, the former’s performance is assumed to be higher by 7 to 12.5 percent. Assessing job performance is a complex construct (Campbell et al., 1993), with its definition varying across contexts and roles in the organization. The number of methods for evaluating job performance is almost similar to the number of job roles (Aguinis, 2013), such issues are addressed with a competency-based PMS. Matching employee competencies and job requirements is claimed to improve employee and organizational performance, as well as lead to increased satisfaction (Spencer and Spencer, 1993).

There are also studies that claim there is no relationship between competency development and performance variables. Yazdanfar et al. (2014) and Sparl et al. (2013) confirmed the same issue; however, they advised for further examination on this phenomenon. Several organizations do not appraise the competencies in their performance appraisal system (Abraham et al., 2001). Ineffectively executed competency models sometimes prove to be a critical expense (Shippmann et al., 2000).
Competency-based appraisal provides an individual to assess their competencies with critical incidents, thus allowing the appraiser to endorse the competency level. Because the onus of initiating the process lies with the appraisee, he/she is likely to take more effort in meeting the competency requirements of the role/level/organization. BARS assists the employees to know their proficiency level against each competency. It confirms whether the employee meets the current level and pushes an individual to attain the next level of proficiency, thus raising the performance bar (American Compensation Association (ACA), Hay Group, Hewitt Associates LLC, Towers Perrin, William M. Mercer Inc., 1996; Hollenbeck et al., 2006).

The goal of the competency-based feedback system is to capture an early opportunity for providing feedback on the employee and to assist in improvement, rather than critical reporting of events in which employee’s non-performance and weaknesses are highlighted as reprimand in conventional PMS. The conventional appraisal system has failed to assist senior-level employees, where competency-based performance (may be with use of 360-degree feedback) assists the employee in coaching. Liske and Holladay (2016) in their research on leadership coaching showed that individuals participating in the coaching program, in comparison with those who did not, demonstrated significantly improved leadership competencies and considerably higher retention rates one-year post-program. In addition, managers become coaches to their subordinates in developing the required level of competencies. A competency model distinguishes excellent performers from average performers by differentiating the proficiency levels against each competency or group of competencies required for the particular role or level. Organizations are integrating competencies in the PMS, as it paves the way for feedback to an individual on the behavioral aspects. By integrating competencies in the PMS policy and process, employees will not only focus on goals but also on the capabilities that drive the goals. Spencer and Spencer (1993) expressed matching the competencies of an employee and job needs to enhance employee and overall organizational performance and enhanced satisfaction. Thus, we hypothesize:

H1. Competency-based performance management practices (self-assessment, BARS, appraisal feedback, coaching, performance differentiation and competency-based PMS policies) have a significant positive influence on superior performance.

4.2 Organizational effectiveness (OE)
According to Oghojafor et al. (2012), OE is a complex and ambiguous concept without a set definition (Rodsutti and Swierczek, 2002). Although there is no consensus on “what” and “how” effectiveness is measured, some studies consider effectiveness as focusing on goals, stakeholders’ satisfaction and connecting with external stakeholders. Cameron (1981) and Lachman and Wolfe (1997) identified organizational goals and objectives to be complex. Hitt et al. (1982) supported the need to measure OE for both organizational executives and researchers, as no common approach measures OE. Inappropriate measures will lead to incorrect conclusions; hence, a conceptual link between organization and management theories cannot proceed without OE measures. Thus, there is no consensus on the definitions of and measurements for OE (Bluehorn, 1980; Bourgeois, 1980; Cunningham, 1977; Molnar and Rogers, 1976; Steers, 1975; Tsui, 1981). Business practitioners use financial measures such as ROE, ROI, EPS, ROA, capital asset pricing for measuring effectiveness, whereas some focus on time-based indexes with industry and inflation as control variables (Hitt et al., 1982; Rumelt, 1974). Cameron (1981) recommended system resource model, goal model, participant satisfaction model and process model besides other approaches of OE (Steers, 1975, 1976; Cameron, 1980). OE is a debated topic in both behavioral and social sciences (Coulter, 1979). Effectiveness relies on what is considered effective; it depends on one’s perspective of the world. Lado and Wilson (1994) claimed that organizations must create and nurture knowledge and intellectual capital. Before the 1960s, organizations’
success sometimes depended on particular management techniques and conventional Tayloristic techniques. Consecutively, internal and external factors formulated the managerial and leadership competencies, which were then adopted by organizations. As competent and smart employees drastically affect organizations’ effectiveness and efficiency, skilled and talented workers are deemed as an asset for the success of any entity, public or private (Snell and Dean, 1992).

According to Mott (1972), the quality, quantity of the output produced, organization’s ability to adapt to the changes in external and internal issues define OE. Thus, OE is the capability of any organization to activate its power centers to adapt and/or act. The criteria of effectiveness of Paul Mott’s OE scale are summarized as: organizing power centers for everyday production (productivity), organizing power centers to change day-to-day activities (adaptability) and organizing power centers to handle short-term impulsive work overloads (flexibility).

Although most organizations use MBO, key performance indicator, KRAs, balanced scorecard-based goal systems, CEOs feel the PMS is not effective. To check whether the competency-based performance management influences OE, thus we hypothesize the following:

H2. Competency-based superior performance has a significant positive influence on OE.

5. Research methodology
The extant literature review on “LCM” and “OE” enabled to propose theoretical model for hypothesis testing between “competency-based superior performance” and “organizational effectiveness.” The theory development on “CBSP” further assisted in developing scale to measure the competency-based performance management practices leading to “superior performance” in the organization. The research method involved 5 phases as elaborated in the Figure 2. The scale for “OE” was revisited in context to this research. These constructs were integrated in scale development, contextualization and validation process. The research design highlighting this study is based on the soft positivism approach (Seddon and Scheepers, 2006) using a quantitative cross-sectional survey to study the impact of CBSP (as independent construct) on OE (as dependent construct.) The collected data enables hypothesis testing via statistical tools, which results in the determination of causality between dependent and independent constructs (Neuman, 2006) (Figure 3).

5.1 Scale/instrument development
Measures for this research are “competency-based superior performance” and “OE.” As there is no scale for “competency-based superior performance” (referred herein as CBSP), this research evidenced the scale development process for CBSP. This research adopted (Table I) construct development, item selection and development, face/content validity, expert opinion, pilot data analysis, full-scale survey, EFA/CFA and scale establishment for the scale development.

The literature on LCM convinced to have 14 items to measure CBSP and 8 items for OE as face validity. The instrument was sent to chosen experts (academia, industry and consulting, who are experienced professionals using LCMs in their domain area) for critical review and remarks using Lawshe’s content validity formula (CVR) comprising (useful, relevant and average) for content validity. Based on the suggestions, changes were made to the instrument. To check whether the theoretical factors resemble statistical factors, a pilot data analysis was done with 127 respondents; EFA shows one item having the factor score below 0.50 in CBSP; hence, it was excluded from the scale, making the final scale with six items on this construct. Reliability analysis of the pilot data was significant. While this research used scale development for CBSP items (Table I), for OE, as it was a valid scale, only EFA was conducted during pilot data analysis because the respondents
were based out of India. The items from the original scale have been slightly modified because Mott’s scale was developed in 1972, and the language was made more focused and crisp addressing all kinds of respondents. Mott’s scale on OE has been widely used in literature. The results yielded perfect factors with significant values. CBSP as an independent variable is used to measure the LCM-based performance management practices with six items, and OE as a dependent variable with eight items was measured using Paul Mott’s scale comprising three factors – productivity, flexibility and adaptability (Table II). The theoretical framework (Figure 1) is tested for empirical

<table>
<thead>
<tr>
<th>Research Conceptualization</th>
<th>Review on extant literature on LCM and OE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Theoretical Framework and Hypothesis Development</td>
</tr>
<tr>
<td>Research Operationalization</td>
<td>Construct development</td>
</tr>
<tr>
<td></td>
<td>Item selection and development</td>
</tr>
<tr>
<td></td>
<td>Face/content validity</td>
</tr>
<tr>
<td></td>
<td>Expert opinion</td>
</tr>
<tr>
<td></td>
<td>Pilot data analysis</td>
</tr>
<tr>
<td></td>
<td>Scale establishment</td>
</tr>
<tr>
<td>Research Execution</td>
<td>Full-scale survey</td>
</tr>
<tr>
<td></td>
<td>Data Screening and Examination</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>Internal Consistency</td>
</tr>
<tr>
<td></td>
<td>Exploratory Factor Analysis</td>
</tr>
<tr>
<td></td>
<td>Convergent and Discriminant Validity</td>
</tr>
<tr>
<td></td>
<td>Nomological and Unidimensionality Tests</td>
</tr>
<tr>
<td></td>
<td>Confirmatory Factor Analysis</td>
</tr>
<tr>
<td></td>
<td>Common Method Bias</td>
</tr>
<tr>
<td></td>
<td>SEM and Hypothesis Testing</td>
</tr>
<tr>
<td>Findings and Discussions</td>
<td>Analysis of Hypothesis</td>
</tr>
<tr>
<td></td>
<td>Implications</td>
</tr>
<tr>
<td></td>
<td>Limitations</td>
</tr>
</tbody>
</table>

**Figure 2.**
Stages in research methodology
correlation. Example of items includes “An individual in this organization does self-assessment using leadership competencies,” “Considering the overall organizational scenario, productivity of employees is perceived to be generally high here.” A five-point Likert scale having values from 5 (always true) to 1 (not true) was applied to the instrument.
5.2 Sample
The demographic variables included questions on the sector of business, turnover of the company (in INR), number of employees, tenure of the competency model being in practice (in years), type of the company – Indian/MNC and public/private/government.

The questionnaire was sent to target samples via e-mail, social media, local and national HR forums. This study is based on a non-probabilistic sampling strategy. For selection of the survey respondents, purposive sampling and snowball sampling were combined. A sample frame was created by listing the companies with LCM. The HR members of those companies represent the sample. Only HR members who held designations such as the head of learning and organizational development, HR vice president, HR director, head of talent management, HR manager, head of training and development, principal consultant and consultant of leadership practice were approached. The reason for this selection was that they had the required experience in implementing LCM in their earlier and present organizations. To eliminate self-biased response error, the authors ensured anonymity of the respondents (Table III).

In total, 292 responses were received. The highlights of the demographics were information technology/information technology-enabled services dominate the respondents (21.5 percent), manufacturing/capital goods/engineering (12.3 percent) and automobile/auto-ancillaries (11.6 percent). In total, 74 percent of the companies had annual turnover over Rs 500 INR crores ($750,00K approximately). Over 62 percent of the companies had LCM being practiced for more than five years.

6. Data analysis
6.1 Internal consistency of the scale
Although demographic variables were entered as control variables in the first step, but because the results were not significant to our study, the authors excluded them from the final analysis. The “internal consistency” method is used to ascertain the reliability of instruments. The coefficient $\alpha$ varies between 0.880 and 0.924, making the scales reliable.

6.2 Exploratory factor analysis
Principal component analysis with varimax rotation was used for the EFA, which resulted in four factors – one factor for competency-based superior performance and three for OE. These four factors showed 79.93 percent of variance in the examined items. The Kaiser–Meyer–Olkin (KMO) test of sampling adequacy and Bartlett’s test of sphericity were used to determine data appropriateness; KMO was found to be 0.858, which is significant (Tabachnick and Fidell, 2007); with Bartlett’s test of sphericity was $p < 0.001$, which is fairly significant. Thus, it can be hypothesized that the correlation matrix may be considered an identity matrix, whereas outcomes indicate significance and vice versa. All communalities ranged from 0.677 to 0.874.

6.3 Convergent and discriminant validity
Convergent validity was examined using Fornell and Larcker’s (1981) method to calculate the average variance extracted (AVE) and composite reliability (CR). Hair et al. (2010) suggested CR measures should be more than 0.70. In this study, all measures of OE and CBSP surpassed the threshold limits. The results indicate high level of convergence with the items and their related constructs. In addition, CR > AVE and AVE exceeds 0.50, Hair et al. (2010). All CBSP and OE factors (Table IV) have AVE values above 0.50, thereby supporting convergent validity. According to Fornell and Larcker (1981), for discriminant validity, AVE should be greater than the maximum shared variance (MSV < AVE) and average shared variance (Hair et al., 2010). The results show that the shared variance between the computations was less than the AVEs of the CBSP and OE individual measures, which confirms discriminant validity. Hence, all items contribute to their pertinent scales and indicate good construct validity.
### Descriptive and Reliability Statistics

<table>
<thead>
<tr>
<th>Classification</th>
<th>Number of respondents</th>
<th>% to total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sector of business</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Automobile/auto ancillary</td>
<td>34</td>
<td>11.6</td>
</tr>
<tr>
<td>Capital goods/heavy engineering/manufacturing</td>
<td>36</td>
<td>12.3</td>
</tr>
<tr>
<td>Aviation/hospitality/travel</td>
<td>8</td>
<td>2.7</td>
</tr>
<tr>
<td>Pharmaceuticals</td>
<td>28</td>
<td>9.56</td>
</tr>
<tr>
<td>FMCG</td>
<td>26</td>
<td>8.87</td>
</tr>
<tr>
<td>Chemicals/fertilizers/agro</td>
<td>8</td>
<td>2.73</td>
</tr>
<tr>
<td>Power/energy/electrical</td>
<td>11</td>
<td>3.75</td>
</tr>
<tr>
<td>Banking/finance/insurance</td>
<td>27</td>
<td>9.22</td>
</tr>
<tr>
<td>Infrastructure/construction</td>
<td>8</td>
<td>2.73</td>
</tr>
<tr>
<td>IT/ITES</td>
<td>63</td>
<td>21.5</td>
</tr>
<tr>
<td>Media/entertainment</td>
<td>6</td>
<td>2.05</td>
</tr>
<tr>
<td>Textiles</td>
<td>5</td>
<td>1.71</td>
</tr>
<tr>
<td>Telecom</td>
<td>12</td>
<td>4.1</td>
</tr>
<tr>
<td>Others</td>
<td>21</td>
<td>7.17</td>
</tr>
<tr>
<td><strong>Mode of competency model creation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internally created</td>
<td>32</td>
<td>10.92</td>
</tr>
<tr>
<td>Externally created</td>
<td>260</td>
<td>89.08</td>
</tr>
<tr>
<td><strong>Turnover of the company</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than INR 100 crores</td>
<td>5</td>
<td>1.71</td>
</tr>
<tr>
<td>INR 101-500 crores</td>
<td>71</td>
<td>24.23</td>
</tr>
<tr>
<td>INR 501-1,000 crores</td>
<td>127</td>
<td>43.34</td>
</tr>
<tr>
<td>INR 1,001-5,000 crores</td>
<td>69</td>
<td>23.55</td>
</tr>
<tr>
<td>Above INR 5,001 crores</td>
<td>21</td>
<td>7.17</td>
</tr>
<tr>
<td><strong>Competency model being practiced for</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than a year</td>
<td>2</td>
<td>0.68</td>
</tr>
<tr>
<td>1–2 years</td>
<td>27</td>
<td>9.22</td>
</tr>
<tr>
<td>3–4 years</td>
<td>74</td>
<td>25.26</td>
</tr>
<tr>
<td>5–6 years</td>
<td>136</td>
<td>46.76</td>
</tr>
<tr>
<td>7 years and above</td>
<td>53</td>
<td>18.09</td>
</tr>
<tr>
<td><strong>Type of company</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indian</td>
<td>218</td>
<td>75.09</td>
</tr>
<tr>
<td>Multi-national</td>
<td>73</td>
<td>24.91</td>
</tr>
<tr>
<td><strong>Respondent level in management hierarchy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Top management</td>
<td>67</td>
<td>23</td>
</tr>
<tr>
<td>Middle management</td>
<td>140</td>
<td>48</td>
</tr>
<tr>
<td>Junior management</td>
<td>85</td>
<td>29</td>
</tr>
</tbody>
</table>

Table III. Demographics Profile

<table>
<thead>
<tr>
<th>Classification</th>
<th>Number of respondents</th>
<th>% to total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Descriptive statistics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>2.77</td>
<td>0.804</td>
</tr>
<tr>
<td>SD</td>
<td>0.924</td>
<td>0.098</td>
</tr>
<tr>
<td>Cronbach’s α</td>
<td>0.924</td>
<td>0.695</td>
</tr>
<tr>
<td>Average variance explained</td>
<td>74.004</td>
<td>0.932</td>
</tr>
<tr>
<td>KMO</td>
<td>74.004</td>
<td>1</td>
</tr>
<tr>
<td>Construct composite reliability</td>
<td>0.405**</td>
<td>0.373**</td>
</tr>
<tr>
<td>CBSP</td>
<td>0.346**</td>
<td>1</td>
</tr>
<tr>
<td>OEP</td>
<td>0.430**</td>
<td>0.433**</td>
</tr>
<tr>
<td>OEA</td>
<td>0.405**</td>
<td></td>
</tr>
<tr>
<td>Flexibility</td>
<td>3.06</td>
<td>0.956</td>
</tr>
<tr>
<td>OEF</td>
<td>0.884</td>
<td>0.500</td>
</tr>
<tr>
<td>Note: <strong>Correlation is significant at 0.01 level (two-tailed)</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table IV. Descriptive and reliability statistics
6.4 Test of unidimensionality and nomological validity
Unidimensionality was measured by ensuring that each item reflected only one underlying construct (Bollen, 1989; Gerbing and Anderson, 1988). The $\alpha$ coefficient ranged from 0.880 to 0.924 and CR ranged from 0.868 to 0.932, which is acceptable (Fornell and Larcker, 1981; Nunnally and Bernstein, 1994). The correlation computations for validation were measured (Arnold and Reynolds, 2003), which revealed that superior performance and OE are positively correlated, thus affirming the nomological validity.

6.5 Confirmatory factor analysis/measurement model
The measurement model was created on all set of items with the maximum likelihood procedure. The model indices emerged as $C_{min}/df = 2.689$, RMSEA = 0.076, NFI = 0.937, CFI = 0.959, GFI = 0.919, IFI = 0.960, RMR = 0.04 and $p = 0.001$, which are acceptable and significant. Hence, CFA asserts the complete fit of the measurement model as good (Maruyama, 1997; Hair et al., 2010).

6.6 Structural equation modeling
For hypothesis testing for the proposed research framework, this study used the SEM approach with SPSS Amos version 20.0. SEM is effective when path analysis includes latent constructs having multiple items. Bollen’s (1989) approach was used for the overall model fit index. The study used the $\chi^2$ (with significance level $p > 0.001$) and normed $\chi^2$ (for mediocre, good or excellent fit), root mean squared error of approximation (RMSEA) (with critical level not more than 0.05, 0.08 or 0.10 to indicate excellent, good or mediocre fit, respectively), goodness of fit index (GFI) (with critical level not lower than 0.80 or 0.70), normed fit index (NFI) (with critical level not lower than 0.90), comparative fit index (CFI) (with critical level not lower than 0.90), incremental fit index (IFI) (with critical level not lower than 0.90) and root mean squared residual (RMR) (with critical level not more than 0.08) (Hair et al., 2010).

The hypothesis (Figure 3) has been tested by developing the measurement model and SEM using the maximum likelihood procedure (Hair et al., 2010; Maruyama, 1997). The overall model fit is calculated by analyzing $C_{min}/df = 2.348$, RMSEA = 0.068, NFI = 0.942, CFI = 0.966, GFI = 0.926, IFI = 0.966, RMR = 0.046, $p = 0.001$). Thus, these tests confirm that the model presents an overall good fit. The results in Figure 3 indicate positive and significant correlations between all structural constructs, thus supporting the hypotheses of the study.

6.7 Common method bias
For examining the common method bias (Podsakoff et al., 2003), Harmon’s one factor test was used to test the variables; however, the test turned to be negative, with no single factor accounting for the variance.

7. Discussions, implications and future research
Although tremendous research and practice over the years has been conducted to improve the performance appraisal and performance management processes in organizations, dissatisfaction with the process is still observed. More than 90 percent of the managers, employees and HR heads are of the opinion that their performance management processes fail to deliver the expected results; many believe the process to be inaccurate and/or ineffective (CEB Corporate Leadership Council, 2012). This study discusses whether competency-based performance meets the expectations of all stakeholders – employee, manager, top management, HR in achieving their objective of performance measurement.

7.1 Analysis of hypotheses testing
This section summarizes the major findings of this study. First, this research tested the “competency-based superior performance” construct with the items self-assessment on
competencies ($H_{1a}$), raising the performance bar of the employee with next-level behaviors ($H_{1b}$), providing feedback to employees on the expected behavior ($H_{1c}$), facilitating coaching during performance discussion ($H_{1d}$), distinguishing excellent performers from others ($H_{1e}$) and lastly, integrating LCM in the performance management process ($H_{1f}$); these were included in the research model for hypothesis testing. The SEM assessment shows LCM compatibility with all items, resulting in a positive significant effect to use LCM in organizations to create superior performance. These findings support the views of Shippmann et al. (2000), Martone (2003) and Draganidis and Mentzas (2006). But, these previous studies were restricted to “mere assumption,” but not evidenced empirically. Also, studies were restricted to “competencies” or “influence of competencies” or “integration of competencies in HR practices” but did not discuss the impact of LCM on performance and OE. Besides, the authors tested whether CBSP influences OE. The results (Figure 3) conclude that there is a significant correlation between CBSP and OE.

The concept of self-assessment in competencies provides an opportunity for an individual to introspect his/her performance on the defined level. LCM thus motivates an individual to meet the expected level of competencies by demonstrating the behaviors, which in turn lead to individual and organizational performance. This study confirms that organizations that provide such opportunity are in a better position to shift the onus of managing performance to an individual. In development-oriented culture organizations, employees, especially high potential employees and young generation workforce, manage their own development and proactively seek for feedback on the growth areas and improvement. Competencies in performance appraisal aid in self-assessment. BARS in LCM motivates further to demonstrate behaviors of the next level for the deserving potential employees; thus, such employees turn to be potential talents for career and succession planning, as well as for promotion opportunities. This also enhances the organizational capabilities in those competencies as employees are consistently demonstrating best-in-class behaviors, providing an organization a competitive edge. By using competency-based performance management, the organization describes the criteria for performance for every level so that the workers/employees can understand which competencies are required for their advancement (Pickett, 1998). The results ensure that organizations having BARS in LCM practice identify such talents. Competency-based performance management provides feedback to employees on which competencies the employee is weak and strong, thus creating areas for improvement on its own for such employees. The performance feedback in this context becomes more open and transparent as the employee is aware of his/her strengths and weakness. The empirical results of this study confirm that organizations providing feedback on performance-based competencies influence the individual performance. The science of behavioral change based on the control theory also elucidates that some discomfort, little feedback on the gap between the “now” state and the desired future state is required to stimulate behavior change and development (Carver et al., 2000). The performance feedback employees receive is often inconsistent and unreliable in goal-based appraisals. One of the core assumptions of virtually all performance management programs is that, if individuals receive feedback on their performance, they will be engaged to improve (assuming that appropriate organizational supports and incentives are in place to support this improvement). This assumption is fulfilled in the competency-based performance process. LCM assists in performance dialogue with coach–coachee’ relations. The coach is in better position to bring change in the employees’ competencies as LCMs are uniformly accepted by the employees, and as the organization has integrated LCM in other people practices, they have better credibility to use. The value of active, ongoing, constructive feedback, performance strategy and coaching conversations between managers and employees
has been long recognized (Meyer et al., 1965). This study validates the influence of performance coaching using competencies. LCM distinguishes the excellent and average performers (Dainty et al., 2004; Hollenbeck et al., 2006; Ryan et al., 2009); thus, excellent performers are differentiated with their compensation, career plans and L&D opportunities. The results of this study validate such exemplary behaviors from competencies and provide an opportunity to retain such employees. By integrating LCM in the PMS policy and processes, the organization ensures all its employees are channelized toward performance using competency-based performance practices (Shet et al., 2017b). The results of this study confirm that, by integrating LCM in PMS, the organization ensures all dimensions of performance are influenced.

It is not sufficient that the organizations use competency-based performance management, such systems should provide better competitiveness for the organization. This research tested CBSP with OE dimensions such as productivity, adaptability and flexibility. The results confirm that CBSP influences OE.

Usage of LCM establishes a significant relationship that will enhance the performance culture in the organization, as the organization moves from a conventional appraisal system to a competency-based PMS, thus ensuring sustainable superior performance. This positive relationship indicates that companies operating in India accept competency-based performance practices and are determined to create a competency culture in organizations. The research further advanced by correlating “CBSP” with “OE.” The positive significant effect of SEM concludes that competency-based performance practices influence OE.

7.2 Theoretical implications
The research correlates LCM influencing superior performance with validated results. This research model will further advance the study either in specific organization, industry or geography or specific intervention such as effectiveness of performance management to a group or individual and so on widely. Evaluating the research model confirms its reliability and validity for replication in other contexts. The correlation of CBSP with OE is again a contribution to the HR domain. This research offers an empirical confirmation for the influence of CBSP on OE. Till date, the literature lacks studies investigating this relationship with empirical evidence. In addition, existing studies often neglected correlations with different significant dimensions of CBSP, which has been discussed in this research. This research is an extension of the previous research on performance management based on the goal system, which has been extensively criticized with researchers saying “it doesn’t work, as it lacks consideration of competencies.” Hence, this research on competencies developed a theory followed by scale development and empirical evidence. To the best of the authors’ knowledge, this study is the first to show how LCM practices such as self-appraisal, competency feedback, BARS as competency level differentiate an excellent performer from an average performer and integrating LCM in the PMS policy and process as a model for superior performance. This research adds to the body of performance management with the CBSP approach and competency literature from a performance management dimension.

7.3 Managerial implications
These findings are important to organizations to build credibility to a competency-based PMS, as the research empirically validates the findings. While the organizations use LCM for many other aspects, the foremost important criterion is organizational performance, where as a goal-based system does not provide the desired results. This research supports the influence of LCM on performance aspects of the organization (Figure 4).
A fundamental message of the competency-based practices is that capability development and performance are the two essential dimensions of implementing LCM, as they improve OE, even when the organizations are operating in the recession period. LCM's consistency and distinctiveness are the main indicators of a strong "behavior enabler," as they affect organizational performance both during the growth and recession periods. These messages demonstrate that the design and implementation of LCM practices in line with the business strategy will make an individual and organizations effective. Consequently, this study contributes to both LCM as content and execution as a process that influences individual and organizational performance. Therefore, these fundamental messages convey that organizations can benefit significantly by focusing on the CBSP dimensions considered in this study.

The findings of this study indicate the methods through which practitioners can reap additional benefits from using LCM-based performance practices; in particular, the findings indicate that, to enhance OE, organizations should apply proactive people practices for improving individual and team performance. The authors recommend the practitioners to focus more on such practices for enabling employee performance (communication of competencies, agreement and acceptance on competency levels by employees, helping to benchmark with proficiency levels against each competency, continuously coaching on competencies, recognizing exemplary behaviors in competency demonstration and finally, suitably rewarding the competent employees with compensation, career plan, succession plans, L&D plans) aiming to improve employee performance. LCM-based performance also offers job engagement and satisfaction, as the focus is on individual capability rather than task accomplishment. LCM acts as a pull strategy for the organization to upgrade the competencies from one level to another; however, this requires continuous communication and creating a culture. By contrast, the authors believe that a proper use of these practices in organizations will provide an opportunity to maximize the benefits from their investment in LCM.

It is not only LCM as a content that develops the shared meanings among employees about performance practices, but it is the approach delivered to employees. We advise that practitioners focus more on the features of performance practices, such as distinctiveness (e.g. legitimacy, visibility, understandability and relevance), consistency (e.g. instrumentability, validity, consistency of performance messages) and consensus (e.g. agreement among the stakeholders and fairness in the system) and making CBSP attractive to employees so that they appreciate their value. The outcomes are the credibility to the
system, transparency in appraisal, shifting the onus to an individual to drive his/her performance and high-performance work culture.

The research findings offer crucial implications for managers. This research developed a scale for measuring competency-based performance management, using the scale development and validation approach. Organizations can use this scale to examine the effect of LCM on superior performance. This model can assist HR practitioners in creating awareness about the required managerial intervention in using competency-based HR practices. In this empirical study, the relationship between LCM-based performance practices and OE is developed; thus, this study can be considered as the one that has furthered the knowledge and understanding of competency-based HR practices and linking them to OE.

The aspects of LCM influencing performance management practices in organizations can be easily determined by the managers, and accordingly alter strategies, if required. This is the first empirical study that evaluates the relationship between competency-based performance management and OE; thus, our research can be considered as a stepping stone for understanding how competency-based HR practices influence OE indices.

7.4 Limitations
This study focused on organizations based in India having LCMs and implementing competency-based performance practices and their correlation with OE. Although adequate literature is available on OE as a variable, this study used Mott’s (1972) scale. Another limitation is the approach used for data gathering; the authors gathered one-time data for the current research, whereas longitudinal analysis will provide consistent and valid explanations to the causal influence of the constructs and to determine whether this hypothesized model would change over time.

7.5 Future research
The findings of this study need to be verified in more general settings: e.g. over larger, more heterogeneous samples. Longitudinal analyses are needed to assess how the findings of this study evolve with time. The effect of competency performance management needs to be understood using before- and after-studies of such interventions. A comparative study in the form of a case study of an MNC having LCM practices can be studied to know which country’s impact of LCM is effective in relation with OE indices. A qualitative method can also be triangulated with the research findings to deepen the understanding of the influence of competency-based performance on OE. Moderation analysis with the maturity of competency-based PMS practices can be investigated. Mediation analysis with “management support” and “organization culture” can be reflected. Different moderators can be used to study the relationship of LCM with different dependent variables. The study can be extended on a comparative basis to determine the industries in which LCMs are more matured and factors responsible against the companies where LCMs are not so much matured and the barriers in those organizations. Demographic characteristics can be incorporated as mediators of the factors considered in this study or can be assessed as independent determinants of acceptance. Studies can also relate whether it is a senior manager, middle manager or junior manager’s competencies that influence organizational performance. Business segment-wise studies can be made in a vast country like India.

8. Conclusion
This study provided a theoretical and empirical test of the underlying assumption in the competency literature that competency-based performance practices can enhance effectiveness in organizations. The theory building on the competency-based performance management
assisted in creating a scale for measuring the impact of CBSP and OE. The present study has implications for future research when the goal-based appraisal system is diminishing from the workplace. Particularly, when organizations implement competency-based performance practices such as self-assessment on competencies, creating BARS to raise the performance bar of the employee to the next level, providing feedback to employees on the expected behavior, facilitating coaching during performance discussion, distinguishing excellent performers against others and integrating LCM in the performance management process. All these practices need to be measured by the organization to understand which aspects of LCM are influencing performance of an individual and in totality of the organization. LCM, on its own cannot drives performance, unless the above practices are built and measured, not just process measures, but by linking to OE measures like productivity, adaptability and flexibility to get a better picture of the effectiveness of the competency-based PMS. In summary, this multidimensional research concluded that competency-based performance practices influence OE. This new approach on competency-based performance management will continue to advance the domain of the performance management.

References
American Compensation Association (ACA), Hay Group, Hewitt Associates LLC, Towers Perrin, William M. Mercer, Inc. (1996), Raising the Bar: Using Competencies to Enhance Employee Per-Formance, American Compensation Association, Scottsdale, AZ.


**Corresponding author**

Sateesh V. Shet can be contacted at: svshet@hotmail.com

---

For instructions on how to order reprints of this article, please visit our website: [www.emeraldgrouppublishing.com/licensing/reprints.htm](http://www.emeraldgrouppublishing.com/licensing/reprints.htm)

Or contact us for further details: permissions@emeraldinsight.com
Objectivity in performance ranking of higher education institutions using dynamic data envelopment analysis

Anup Kumar
Department of Operations and Business Analytics,
Institute of Management Technology Nagpur, Nagpur, India, and
Rajiv R. Thakur
Department of Strategy, Birla Institute of Management Technology, Noida, India

Abstract

Purpose – There has been a persistent debate on measures of efficiency and ranking procedures of higher education institutions (HEIs). Deriving absolute efficiency measures and their ranking provide a critical input for the society to choose the appropriate educational institute. The purpose of this paper is to evaluate the relative performance of institutions in management education in different locations in India and propose a holistic efficiency measurement which can be applied to HEIs in general.

Design/methodology/approach – This study uses dynamic data envelopment analysis (DDEA) as the primary methodology of analysis. Multiple measures of inputs and output have been defined to assess efficiency in institutions of management education. Some of the output variables used for measuring relative effectiveness are: the number of students placed, number of entrepreneurs, median CTC of placed students, total number of students passed, number of research publications, number of students and faculty who have participated in international exchange, input variables used, student intake, faculty profile, resource allocation on the development of student, faculty and staff, industry linkages, alumni network. The institutions under study are in three different locations in India, having distinct characteristics. The multiple measures of inputs and outputs defined have been used to measure efficiency, following which DDEA was used to rank the efficiency measures.

Findings – Various agencies use their framework to evaluate and rank HEIs; however, they are either subjective or less researched methodologies. The proposed method acts as a new researched and objective methodology for ranking of HEIs operating across regions with different societal, economic and political contexts. Efficiency in education is of high relevance today for various stakeholders such as students, parents, industry, policy-makers and government. An objective, such as the one proposed in this paper, would be helpful in satisfying the needs of various stakeholders. Furthermore, the government has policies of allocating funds, in case of public-funded institutions, based on efficiency levels in HEIs. The measure using DDEA suggested in this study provides a better measurement of efficiency.

Research limitations/implications – This research is based on the extension of DDEA with slight modification to the denominator portion of efficiency calculation. The modification is accentuated by taking an industry benchmark or government benchmark. This may lead to slight difficulty in the appropriation of input parameters. Hence, selection of appropriate input and output parameters is the key limitation. To demonstrate capabilities of the proposed approach, this framework is implemented for performance evaluation of institutions of higher education in India. Some helpful policy-making and managerial insights are derived from the numerical results.

Originality/value – The uniqueness of this research is that it adds a well-researched methodology based on DDEA to measure efficiency and rank HEIs for effective assessment and benchmarking. The frameworks used so far have been either subjective or less researched methodologies.

Keywords Equal opportunities, Efficiency, Higher education, Performance measures, Dynamic data envelopment analysis

Paper type Research paper

1. Introduction

Globally, the economics of higher education has propelled countries to develop world-class institutions to attract the attention of foreign students. “Some countries make very large amounts of money from international students, but it is a cutthroat market, and perceived...
status and reputation are important marketing tools. The arrival of two major global ranking lists which have attracted much public attention has provided added impetus to the competition for international students. These are the Shanghai ‘Academic Ranking of World Universities’ and the Times Higher Education Supplement of ‘World University Rankings,’ first published in 2003 and 2004, respectively” (Harvey, 2008).

Six Sigma is a process improvement tool that may be applied to improve the ranking of institutions. A well-crafted state-of-the-art literature review was conducted by Biju (2013) for various strategies to improve the performance of higher education institutions (HEIs) in India (Biju, 2013; Biju and Nair, 2009). There have been instances where various tools like Six Sigma and process orientation tools like DMAIC are applied for the measurement of quality and performance improvement of HEIs (Biju and Nair, 2017). Ranking and reshaping of the higher education system have been considered important for some time. In India, for example, National Institutional Ranking Framework (NIRF) was started by the Government of India in 2015. In China, the government has four categories of elite HEIs. The first and largest of these is Project 211, which was established in 1995 to strengthen research standards in China’s top universities; universities that exceeded a threshold were included in this group and received significantly increased funding. As of 2018, 116 HEIs are members of Project 211 (Times Higher Education Ranking, 2017). The second, established in 2015, is the Double First Class University Plan to create 42 world-class universities by 2050. A third, more selective group, is Project 985, established in 1998. The Chinese Government included 39 universities in Project 985 and capped membership to these 39 in 2011. The final and most selective group is the C9 League, established by the Chinese Central Government on May 4, 1998, as part of Project 985 with the goal of advancing Chinese higher education through formalizing an elite group of universities to foster better students and share resources. Nine universities were selected and allocated funding, and on October 10, 2009, the relationship between these nine universities was formalized into China’s C9 League. In the USA, the quality of graduate institutes has been evaluated through staff surveys since the 1920s, and US News and World Report started publishing “America’s best colleges ranking” in 1983. In the UK, ranking tables appeared in the 1990s (Bowden, 2000). However, rankings of higher education have gained prominence in the past decade. “Within only a few years, rankings have become an unavoidable part of academic life, for better or worse” (Harvey, 2008).

World media organizations have been predominant in the publication of rankings such as the Times Higher Education Supplement (THES) (first published in The Times, October 1992), the Financial Times and The Sunday Times (UK/Ireland), Der Spiegel (Germany), Maclean’s (Canada) and Reforma (Mexico). There is growing ambition with the standing and trajectory of the top 100, which are less than 1 percent of the world’s institutions. Similar to the ranking of restaurants or hotels, no one wants to be at the bottom of the ranking list. In recent years, government and accreditation agencies and higher education organizations have developed systems for evaluating and ranking institutional performance: for example, CHE (Germany), AQA (Austria), CIEES, CACEI, CNEIP and CONEVAL (Mexico), NAAC, NBA (India), Higher Education Council and TUBITAK (Turkey), the Commission on Higher Education and Philippine Accrediting Association of Schools, Colleges and Universities (the Philippines) and the Higher Education Evaluation and Accreditation Council of Taiwan (HEEACT). In addition, there are a variety of commercial college “guide” books and websites, for example, the Good Universities Guide (Australia), Bertelsmann Stiftung (Germany) and Research Infosource Inc. (Canada). As higher education became globalized, the focus has shifted to worldwide university rankings, for example, the Shanghai Jiao Tong Academic Ranking of World Universities (ARWU), Times Higher Education – QS Top Universities and Webometrics (Deem et al., 2008; Hazelkorn, 2009, 2013, 2015).
Major concerns of the ranking systems are methodological, pragmatic, moral and philosophical. Therefore, the ranking systems have proved to be controversial. Critique coming from the methodology side is generally based on: selection of indicators and its validity, data availability, dealing with missing values, weighting of the indicators, reliability, the equation for final rank and statistical significance (Bowden, 2000; Harvey, 2008; Taylor and Braddock, 2007). The concept of ranking of HEIs is dominated by the THES system and the Jiao Tong system (Taylor and Braddock, 2007).

THES ranking system is based on the following indicators and their respective weights:

1. 40 percent of a university’s total score is determined by peer review;
2. 10 percent depends on ranking by major (mainly international) graduate recruiters;
3. 20 percent is fixed by citations (per capita) of published academic papers;
4. 20 percent is determined by ratios of the teaching staff; and
5. 10 percent is determined by “international orientation.”

The Jiao Tong system is based on the following indicators and their respective weights:

1. 10 percent of the score is awarded on the basis of Nobel Prizes and Fields Medals won by alumni;
2. 20 percent is awarded on the basis of Nobel Prizes and Fields Medals won by faculty members;
3. 20 percent is awarded on the basis of the number of highly cited researchers in a wide range of disciplines;
4. 20 percent is awarded on the basis of the number of papers published by staff in the journals Science and Nature;
5. 20 percent is awarded on the basis of the number of papers published by staff in a wide range of academic journals; and
6. 10 percent is awarded on the basis of academic performance with respect to the size of an institution.

Clearly, the major difference between the above two ranking systems focuses on diverse indicators. THES focuses on infrastructure and public perception, whereas the Jiao Tong system largely focuses on research. Therefore, it is evident that there is a need to develop an objective ranking system that incorporates the concerns of critics of the above-ranking systems.

Quality of higher education has always been the top priority since the initial stages of awareness about the performance of HEIs. The initial focus was on quality assurance for teaching–learning processes and research (Alexander, 2000; Deem et al., 2008; Dill et al., 1996; Massy, 1997). Efficiency of educational institutions has not been of interest for some reason or the other until recent times, but growing competition and quality conscious society have forced researchers to develop some efficiency measurement framework for educational institutes. Quality in HEIs is a complex philosophy that has been researched by many researchers and is well-defined and measured (Harvey and Williams, 2010; Tam, 2001; Tasopoulou and Tsiotras, 2017). Tam, 2001 had discussed the quality-based performance measurement framework. However, the input–output-based efficiency of HEIs and its measurement is perhaps less researched or is almost untouched.

This paper is positioned to develop an input–output-based efficiency measurement framework for HEIs. Keeping in mind different segments of higher education needs of the society, public and private players have developed various educational institutes.
However, the input and output expectations are diverse due to demographic diversification as well as for temporal and spatial reasons. Therefore, the challenge before researchers and promoters is how to measure holistic efficiency of these diverse educational institutes. Comparison of efficiencies of a well-located old and renowned institute with relatively new and adversely located institute is perhaps not a good proposition. To avoid these biases, a holistic methodology should be developed to measure and compare diverse HEIs.

Efficiency measurement of HEIs has recently been in the limelight among researchers. “Efficiency in education is a topic of intense debate among politicians, teachers, and other educational stakeholders. In addition to increased awareness of public sector efficiency, the increasing cost of education might be a reason for the enhanced interest about efficiency in education. On an average, education becomes more expensive than other commodities” (Andersson et al., 2017; De Witte and López-Torres, 2017). Technical efficiency and productivity of HEIs have been measured using data envelopment analysis (DEA) framework (Andersson et al., 2017; Jauhar et al., 2017). However, traditional DEA has some logical issues regarding the use of the input and output data. Logically, any input intervention has some lag to actually convert into output deliverables; therefore, this paper has used a modified version of DEA, that is known as dynamic data envelopment analysis (DDEA), to overcome this logical issue (Kumar et al., 2015; Mariz et al., 2018). In traditional DEA, we assume the decision-making unit (DMU) as a black box and simply infer that the input is directly converted to output without knowing the processes inside DMU, and “fails to illustrate the relationship between variables and the mapping of the underlying production process. In reality, decision-making units (DMUs), here universities, often perform several different functions and readily divide these functions into different components, in a series or in parallel and/or in a more complex form of network type. This suggests that outputs produced in a certain series may become intermediate inputs in a subsequent production stage. To determine the appropriate production process that best represents university research production, we employ the network DEA (NDEA) approach” (Lee and Worthington, 2016). The DDEA applied in this paper is quite different from NDEA. Lee and Worthington (2016) argued that it is essential to know the underlying functions inside DMUs which transform inputs into outputs. Knowing the process inside DMUs does not perhaps help the goal of developing an appropriate method based on input and output, which captures the actual transformation of inputs into outputs over a period; NDEA has missed out the time factor which plays an important role to convert inputs into outputs. Therefore, the major gap in NDEA methodology is that the method ignores the time factor. Hence, this study is positioned to incorporate time factor into consideration by applying DDEA.

Why DEA?
Higher educational organizations are not-for-profit organizations. Gauging the efficiency of a non-for-profit organization is always a nonparametric type of situation; DEA is a well-established nonparametric method to compare and rank various DMUs. DEA, as a methodology, has been used to measure the research performance of Chinese HEIs, monitor and assess the institutional performance of HEIs (Abbott and Doucouliagos, 2003; Alexander, 2000; Chu Ng and Li, 2000; Johnes and Li, 2008). DEA is used to measure organizational effectiveness for HEIs (Cameron, 1978); cost-efficiency in public HEIs is also measured using DEA (Robst, 2001). Technical and scale efficiencies of Australian universities and HEIs of Sweden are also measured using DEA (Abbott and Doucouliagos, 2003; Andersson et al., 2017; Avkiran, 2001).

A comparative analysis is provided using differential evolution and DEA to measure the sustainable performance of an Indian HEI educational supply chain (Jauhar et al., 2017). The social perception of HEI is a typical parameter that is used by various rating agencies and is difficult to gauge. A set of measures, derived from Twitter data, to quantify the
effectiveness of higher education services, have been used for performance measurement in the public sector (Agostino and Arnaboldi, 2017). DEA has been modified for its multidimensional data handling capability. It has also been redefined for handling fuzziness in data. DEA has been widely applied in measuring the efficiency of banking units and supplier efficiency measurement in a supply chain of various commodities (Azadi et al., 2015; Fukuyama and Matousek, 2017; Stewart et al., 2016; Zhou et al., 2016).

Applicability and capability to measure efficiency for multidimensional nonparametric data have prompted us to use DEA to measure the efficiency of HEI. This paper is positioned to propose an alternative method for ranking of HEI (Andersson et al., 2017).

Taking into account the above analysis and the need to develop a scientific efficiency measurement and ranking methodology for HEI, this paper has the following objectives:

1. to develop a methodology to measure efficiencies of HEIs considering the time factor using DDEA; and
2. to provide a holistic approach to rank HEIs in the current competitive educational world.

This paper provides a holistic approach to rank HEIs with an example from business schools in India. Section 1 is the introduction, Section 2 explains data and methodology, Section 3 elaborates on the example, followed by Section 4 (results and implication) and Section 5 (conclusion and future direction).

2. Data and methodology

Students, parents and the industry are three major stakeholders in the field of management education, and meeting their expectations is the key objective of the institutions. In meeting the expectations of these key stakeholders, quality is a must in institutions as far as input, processes and outputs are concerned. Rising expectations and competitive forces create an urge among institutions toward qualitative excellence which can be seen in their individual communications through various channels. However, there are agencies that take an independent outsider’s view and rank these institutions on certain parameters. Some such rankings for management education in India are done by agencies such as National Human Resource Development Network – Cinque, Competition Success Review – GHRDC, C fore-Business World and MDRA-BT. Generally, across all the rankings, the methodology adopted is to use the average of the scores of different parameters. A few of the common parameters found in these rankings are faculty strength and profile, research output, student intake, curriculum and teaching-learning, placements, management development programs and consultancy, internationalization and infrastructure. The Ministry of Human Resource Development, Government of India, which started NIRF rankings in 2016, besides the above parameters also based its ranking on perceptions of key stakeholders of the institutions. However, these rankings are not without limitations. For example, the ranking models put the inputs going in and output coming out of the institution together in the value creation process which is a serious limitation. Second, the sample used for ranking constitutes institutions across different types such as government-run autonomous institutions, Indian Institute of Management, university-affiliated institutions offering MBA, and autonomous private institutions offering Post Graduate Diploma Program. These institutions vary in the basic characteristics of their foundation, governance, resources, curriculum and program design. Moreover, the sample institutions are all combined into one, despite their huge age differences and significant differences in geographic dispersion, and their connection with the industry. These add to the limitations. The current ranking methods and results of these rankings are therefore not very objectives, leading to ambiguity and confusion in the minds of stakeholders.
Consider the following case taken from the literature. A study of 638 public secondary schools in Texas was undertaken by a consortium of three universities, in collaboration with Texas Education Agency. The study intended to try to develop improved methods for accountability and evaluation of school performances (Cooper et al., 2006).

Consider, for instance, the State-mandated Excellence Standards for Texas, recorded in Table I. These represent statements of desired goals, and schools are rewarded (or not rewarded) on the basis of their achievements. Nothing is said about the quantities (or varieties) of resources used.

Hence, it should be no surprise that only one of the excellent-rated schools included in the study was found to be efficient. All the other schools rated as excellent by the State of Texas had expended excessive resources in achieving these goals. Contrarily, many schools that failed to achieve excellence were nevertheless found to be efficient in producing the desired outputs under very difficult conditions. Therefore, some way of recognizing this type of achievement is needed if “efficiency” and “effectiveness” are to be rewarded. “The rationale of performance funding is that funds should flow to institutions where performance is manifested: ‘performing’ institutions should receive more income than lesser performing institutions, which would provide performers with a competitive edge and would stimulate less performing institutions to perform. The output should be rewarded, not input” (Herbst, 2007).

Therefore, an alternative objective and more scientific model is attempted here in the form of an input–output model. Selection of input and output data is an important part of efficiency measurement and ranking. Performance of HEIs is directly related to the quality of higher education. There are various models of assessing the quality of HEIs; these schools of thoughts focus on various aspects of HEIs’ quality parameters. The main models are listed below (Srikanthan and Dalrymple, 2002):

1. A transformative model which mainly focuses on the quality of learning (Harvey and Knight, 1996).
2. An engagement model of program quality which focuses on the quality of pedagogy and courses (Haworth and Conrad, 1997).
3. University of learning model which focuses on teaching research and community development (Marton and Bowden, 1998).
4. A model for a responsive university – “The model is based on the premise that the public would judge the university in terms of the quality of relationships and the quality of outcomes. Quality relationships [are] characterized by mutuality and equality. Therefore, to survive and thrive, universities will have to be responsive” (Tierney, 1998).

<table>
<thead>
<tr>
<th>Outcome indicator</th>
<th>State-mandated excellence standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Graduation rate</td>
<td>99% of graduating class</td>
</tr>
<tr>
<td>2. College admission tests</td>
<td>35% of graduates scoring above the criterion score, which is equal to 25 on the ACT and 1,000 on the SAT</td>
</tr>
<tr>
<td></td>
<td>70% of graduates taking either the ACT or the SAT</td>
</tr>
<tr>
<td>3. Dropout rate</td>
<td>Less than or equal to 1% of total enrollment</td>
</tr>
<tr>
<td>4. Attendance</td>
<td>97% of total enrollment in the school</td>
</tr>
<tr>
<td>5. Texas Assessment of Academic Skills (TAAS) test</td>
<td>90% of students passing all standardized tests</td>
</tr>
</tbody>
</table>

Notes: ACT, American Collegiate Tests; SAT, Scholastic Aptitude Tests

Table I. State-mandated excellence standards on student outcomes
The above-said models set the guidelines to choose appropriate input and output parameters for efficiency measurement and ranking of HEIs. The Department of Higher Education, Government of India has segregated HEI indicators into various categories for ranking (National Ranking Framework). These categories are as follows:

1. Teaching, learning and resources:
   - student strength including doctoral students;
   - faculty–student ratio with emphasis on permanent faculty (FSR);
   - combined metric for faculty with PhD (or equivalent) and experience (FQE); and
   - total budget and its utilization (CBTU).

2. Research and professional practice:
   - combined metric for publications (PU);
   - combined metric for quality of publications (QP);
   - IPR and patents: filed, published, granted and licensed (IPR); and
   - Footprint of Projects and Professional Practice and Executive Development Programs (FPPP).

3. Graduation outcomes:
   - combined percent for placement, higher studies and entrepreneurship (GPHE);
   - metric for university examinations: GUE;
   - median salary;
   - metric for graduating students admitted to top universities (GTOP); and
   - metric for a number of PhD students graduated GPHD.

4. Outreach and inclusivity:
   - percent of students from other states/countries (region diversity RD);
   - percentage of women (WF) + (WS) + (WA);
   - economically and socially challenged students (ESCS); and
   - facilities for physically challenged students (PCS).

5. Perception:
   - peer perception: employers and research investors (PREMP);
   - peer perception: academics (PRACD);
   - public perception (PRPUB); and
   - Competitiveness (PRCMP).

In the above categories, perception is very difficult to quantify and is fuzzy in nature. Therefore, in this paper, we have omitted perception and taken all other parameters segregated as input and output parameters, as listed in Tables II and III.

Dynamic DEA
To incorporate the lag effect while measuring efficiency in various industries, a good estimate of lag is required. Lag time period is estimated using VAR (\(p\)). The VAR model has attractive properties in providing consistent and asymptotically normal estimates.
<table>
<thead>
<tr>
<th>X1</th>
<th>X2</th>
<th>X3</th>
<th>X4</th>
<th>X5</th>
<th>X6</th>
<th>X7</th>
<th>X8</th>
<th>X9</th>
<th>X10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of student intake all programs (no.)</td>
<td>Total number of faculty with PhD (no.)</td>
<td>Number of faculty with more than 5 years of industry experience</td>
<td>Total number of support staff (no.)</td>
<td>Frequency of change in curriculum and program design in last three years</td>
<td>Number of titles books and journals</td>
<td>Number of software and database</td>
<td>Student development fund</td>
<td>Faculty development fund</td>
<td>Staff development fund</td>
</tr>
<tr>
<td>X11</td>
<td>X12</td>
<td>X13</td>
<td>X14</td>
<td>X15</td>
<td>X16</td>
<td>X17</td>
<td>X18</td>
<td>X19</td>
<td></td>
</tr>
<tr>
<td>Alumni network (no.)</td>
<td>Industry linkage (no.)</td>
<td>International linkages (no.)</td>
<td>Number of executive training programs</td>
<td>Average cutoff percentile for admission</td>
<td>Median salary for faculty</td>
<td>Average tenure of director</td>
<td>Median tenure of faculty</td>
<td>Total number of training programs and workshops for students</td>
<td></td>
</tr>
</tbody>
</table>

**Table II.**

Input parameters

Objectivity in ranking of HEIs
<table>
<thead>
<tr>
<th>Y1</th>
<th>Y2</th>
<th>Y3</th>
<th>Y4</th>
<th>Y5</th>
<th>Y6</th>
<th>Y7</th>
<th>Y8</th>
<th>Y9</th>
<th>Y10</th>
<th>Y11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of students placed</td>
<td>Number of entrepreneurs</td>
<td>Median CTC of placed students</td>
<td>Number of international placements</td>
<td>Total number of students passed</td>
<td>Number of research publications in indexed journals</td>
<td>Number of cases published in referred journals/ Ivey or Harvard</td>
<td>Number of trained executives through training programs</td>
<td>Revenue generated through executive training</td>
<td>Number of research/ consultancy projects</td>
<td>Number of students and faculty underwent international exchange (incoming and outgoing)</td>
</tr>
</tbody>
</table>

Table III. Output parameters
regardless of possible nonstationary, unit roots or co-integration in the input or output series. First, VAR estimation provides information about the dynamic interrelation among inputs and outputs. After that, the information is used to derive output responses due to shocks in inputs, which can be used as estimators of lag parameters (Kumar et al., 2015).

Here, in this research, in HEIs policy interventions and investments are mostly assumed to be effective with a lag of three years. Therefore, three-year, exponential smoothing filter has been used to conduct the DDEA model for ranking of HEIs (Figure 1).

Based on the static production framework, this Section sets up the DDEA model which can be used to evaluate efficiencies under the influence of lag effects. The DDEA model captures the intertemporal relationship between inputs and outputs in dynamic production processes. We will first define lag parameters and then continue to formulate the DDEA model.

Consider a dynamic production process observed from time period $t_0$ to $t_N$. For an arbitrary time period $t_a$, the $m$-period lag model depicts a production process in which inputs used in period $t_a$ can contribute to output production in periods $t_a$, $t_{a+1}$, …, $t_{a+m}$, with $a + m \leq N$. Formally, we represent the productive effects of a specific input $j$ of DMU$_j$ (Figure 2).

To convey key ideas of the dynamic model more easily, we simplify the notation by assuming that the dynamic parameters can be considered invariant with respect to their starting times and across different DMUs. In various types of situations, the way of capturing the lag effect might differ, for example, in the case when pharmaceutical retail efficiency (Kumar et al., 2015) suggested to calculate the lag length using Schwarz and Hannan and Quinn’s loss matrices (Hannan and Quinn, 1979; Quinn, 1980, 1989). Moreover, the lag length may be verified using test results for no seasonal variations. Augmented Dickey–Fuller test was used to test the null hypothesis conveyed that each variable series is nonstationary. Johansen’s trace test (Hansen and Johansen, 1999; Johansen and Juselius, 1992) was used to test the co-integration.

![Figure 1. Steps of evaluating dynamic efficiency](source:kumar-et-al-2015)

![Figure 2. Illustration of the $m$-period lag model](source:kumar-et-al-2015)
In the HEI setting, it is assumed that policy interventions are generally annual and discrete in nature, and the effect of this policy intervention is gauged annually. Almost all ranking and accreditation agencies generally collect annual data and they look for three years data for continuous improvements in the analysis. The question is what should be the ideal lag length in an educational setting. One might argue to use the methods discussed above, but they might give vague and unrealistic results in the case of HEIs. It is always better to look for a method that incorporates the partial contribution of each year data, and higher weight should be given to the previous year. Generally, in India, HEIs have a three-year period to implement and check the output. Therefore, the ideal time to check the effect of policy could be taken like three years.

In any HEI ranking, the objective is to track the trend for continuous improvement. We have taken three-year smoothing method to forecast the dynamic parameters, where the highest weight is given to the base third year. This is also called low-pass filter in time series analysis. A low-pass filter is used to track the lower order frequency trends, which is required for the educational setup. The following low-pass filter is used to track trends in the chosen input and output parameters:

\[
Sm(X_t) = \sum_{j=0}^{n} a(1-a)^j X_{t-j},
\]

(1)

where \( n \) is number of years taken for analysis:

\[
Sm(Y_t) = \sum_{j=0}^{n} a(1-a)^j Y_{t-j},
\]

(2)

where \( 0 < a < 1 \); the method is known as exponential smoothing technique and it plays a crucial role in many empirical studies. Literature suggests that \( \alpha \) is chosen between 0.1 and 0.3 (Kendall and Ord, 1990).

We have suggested using an exponential filter (Figure 3) to track the trend in data and modified the step suggested by Kumar et al. (2015) (Figure 1).

The modified DDEA model would be:

\[
\text{Max } v = \sum_{r \in k} u_r Sm(Y_{t_r}),
\]

(3)

where \( r \) is output index total number of output parameter are \( k \) and \( j \) is DMUs index total number of DMUs are \( N \).

Subject to:

\[
\sum_{i \in h} v_i Sm(X_{t_i}) = 1,
\]

(4)
where \( i \) is input index total number of input parameters are \( h \), and:

\[
\sum_{r \in k} u_r Sm(Y_{ir}) - \sum_{i \in h} v_i Sm(X_{ij}) \leq 0
\]

\[ u_r, v_i \geq 0 \text{ and } r \in k, i \in h, j \in N. \quad (5) \]

It was noted that if one uses Equation (4) to calculate efficiency, the majority of the DMUs may be efficient, and therefore in the case of HEIs, it is very difficult to differentiate them. Therefore, we suggest that in an environment like HEIs, where government policies are dominant, one should take minimum benchmark input resources required. Input variables may be a benchmarked government minimum requirement or maybe industry average.

Here in the example, we have taken Equation (4) in a special situation as follows. Subject to:

\[
\sum_{i \in h} v_i (X_{\text{industry average}_{ij}}) = 1, \quad (6)
\]

where \( i \) is input index total number of input parameters are \( h \).

The above modification compares the output of each of the HEI with a benchmark input parameter.

3. Data collection and preparation with an example

The above-proposed methodology is illustrated using one example of a private educational group operating at various locations in India. We have collected primary data for four consecutive years for three educational institutes of the group located in the north, north-west, and central India. The three institutions, though belonging to the same group and running similar programs in management, are different in age, scale and location, which broadly affect the competitive scenario, profile of students, faculty, connect with the industry and general functioning of the institution. After data collection, we have applied a low-pass filter to gauge the trend in data, assuming \( n = 4 \) (for four years; see Equations (1) and (2)). After filtering the data for input and output at each of the location (DMUs), we applied input-based Cranes, Cooper and Rhodes to calculate efficient DMUs (Charnes et al., 1978).

Filtered data with the low-pass filter are shown in Tables IV–VI for each of the DMU. For all DMUs, input and output parameters after applying low-pass filter are tabulated in Tables VII and VIII, respectively.

In this study, we have proposed a novel type of methodology that is expected to fill an essential gap in the ranking process of HEIs. Old methods were survey-based and they were unable to incorporate input–output appropriation. In the suggested method, we have modified the process to appropriate input and output parameters which are an essential part of performance ranking. The following steps have been taken to demonstrate the proposed methodology:

- Step 1. We collected appropriate input and output data points at three different HEIs (the input and output parameters may be government standards and listed parameters).
- Step 2. After collecting data, we identified an industry benchmark for each of the input parameters.
- Step 3. After filtering the input and output variables, industry averages for input parameters are as shown in Tables IX and X.
- Step 4. Finally, we applied DDEA as discussed above and demonstrated in one of the DMUs below.
The linear equation for DMU 1 is as follows.

Max: 

\[ U_1 = 37.9852 + U_2 1.21 + U_3 2.81608 + U_4 1.84 + U_5 39.5035 + U_6 4.7338 + U_7 2.827 \\
+ U_8 13.978 + U_9 2.3755 + U_{10} 1 + U_{11} 1.21, \]

subject to:

\[ V_1 = 127.714 + V_2 8.1246 + V_3 4.4573 + V_4 13.0316 + V_5 1.3766 + V_6 218.95 \\
+ V_7 7.4848 + V_8 14.069 + V_9 8.95305 + V_{10} 2.6891 + V_{11} 233.8249 \\
+ V_{12} 3.8455 + V_{13} 5.998 + V_{14} 5.9539 + V_{15} 27.3256 + V_{16} 16.2474 \\
+ V_{17} 2.33945 + V_{18} 3.3954 + V_{19} 8.1701 = 1, \]
and:

\[
\sum_{r \in k} u_r \text{Sm}(Y_{r_0}) - \sum_{i \in h} v_i \text{Sm}(X_{i_0}) \leq 0
\]

\[u_r, v_i \geq 0 \text{ and } r \in k, i \in h, j \in N.\]

Linear equations for DMUs 2 and 3 would be formulated, based on the above illustration. The rank based on the output maximization (efficiency) is shown in Table XI. Here, we may encounter efficiency more than 1 because we have benchmarked the input parameters and taken industry average that causes efficiency to go beyond 1, as listed in Table XI.
4. Results, implications and limitations

We collected data for the stated input and output variables for each of the DMU. We also figured out the benchmarked input variables taking the average of the each of the DMU input variables (it may be industry benchmark or government standard). In the original DEA methodology, absolute input variable values are used to calculate efficiency, but here we have taken the newly introduced average values of input variables (maybe industry benchmarked in other cases) for calculation of efficiency which leads to a clear ranking, presented in Table II.

From Table II, it is clear that DMU 3 has the highest efficiency compared with DMUs 1 and 2. Calculation of DEA efficiency based on benchmarked input parameters could lead to a better indicator for efficiency ranking than absolute ranking on individual DMU input–output efficiency. Furthermore, it will normalize the input parameters on the basis of standards governed by regulators. The proposed method is better than the existing ranking
<table>
<thead>
<tr>
<th>HEIs/DMUs</th>
<th>X1</th>
<th>X2</th>
<th>X3</th>
<th>X4</th>
<th>X5</th>
<th>X6</th>
<th>X7</th>
<th>X8</th>
<th>X9</th>
<th>X10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total number of student intake all programs (nos)</td>
<td>Total number of faculty with PhD (nos)</td>
<td>Number of faculty with more than 5 years of industry experience</td>
<td>Total number of support staff (nos)</td>
<td>Frequency of change in curriculum and program design in last three years</td>
<td>Number of titles books and journals</td>
<td>Number of software and database</td>
<td>Student development fund</td>
</tr>
<tr>
<td>1</td>
<td>56.19</td>
<td>6.063</td>
<td>1.92</td>
<td>9.698</td>
<td>1.46</td>
<td>2.304</td>
<td>4.679</td>
<td>17.63</td>
<td>2.998</td>
<td>1.638</td>
</tr>
<tr>
<td>2</td>
<td>56.19</td>
<td>7.649</td>
<td>2.13</td>
<td>11.63</td>
<td>2.13</td>
<td>1.999</td>
<td>8.783</td>
<td>17.63</td>
<td>7.69</td>
<td>2.352</td>
</tr>
<tr>
<td></td>
<td>X11</td>
<td>X12</td>
<td>X13</td>
<td>X14</td>
<td>X15</td>
<td>X16</td>
<td>X17</td>
<td>X18</td>
<td>X19</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Alumni network (nos)</td>
<td>Industry linkage (nos)</td>
<td>International linkages (nos)</td>
<td>Number of executive training program</td>
<td>Average cutoff percentile for admission</td>
<td>Median salary of faculty</td>
<td>Average tenure of director</td>
<td>Median tenure of faculty</td>
</tr>
<tr>
<td>1</td>
<td>107.9</td>
<td>1.67</td>
<td>4.893</td>
<td>2.134</td>
<td>28.46</td>
<td>38.64</td>
<td>2.005</td>
<td>3.123</td>
<td>1.924</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>260.2</td>
<td>7.693</td>
<td>6.725</td>
<td>1</td>
<td>24</td>
<td>4.503</td>
<td>2.737</td>
<td>3.867</td>
<td>14.68</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>333.3</td>
<td>2.174</td>
<td>6.376</td>
<td>14.73</td>
<td>29.52</td>
<td>5.603</td>
<td>2.277</td>
<td>3.197</td>
<td>7.903</td>
<td></td>
</tr>
</tbody>
</table>

Table VII. Input parameters after low-pass filter

Objectivity in performance ranking of HEIs.
<table>
<thead>
<tr>
<th>HEIs/DMUs</th>
<th>Number of students placed</th>
<th>Number of Entrepreneurs</th>
<th>Median CTC of placed students</th>
<th>Number of international placements</th>
<th>Total number of students passed</th>
<th>Number of research publications in indexed journals</th>
<th>Number of cases published in referred journals/Ivey or Harvard</th>
<th>Number of trained executives through training programs</th>
<th>Revenue generated through executive training</th>
<th>Number of research/consultancy projects</th>
<th>Number of students and faculty undertaken international exchange (incoming and outgoing)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>37.9852</td>
<td>1.21</td>
<td>2.81608</td>
<td>1.84</td>
<td>39.5035</td>
<td>4.7338</td>
<td>2.827</td>
<td>13.978</td>
<td>2.3755</td>
<td>1</td>
<td>1.21</td>
</tr>
<tr>
<td>2</td>
<td>49.636</td>
<td>3.0706</td>
<td>2.99899</td>
<td>1</td>
<td>52.5277</td>
<td>7.1236</td>
<td>1</td>
<td>27.5755</td>
<td>13.7302</td>
<td>3.4108</td>
<td>2.9278</td>
</tr>
<tr>
<td>DMUs</td>
<td>X1</td>
<td>X2</td>
<td>X3</td>
<td>X4</td>
<td>X5</td>
<td>X6</td>
<td>X7</td>
<td>X8</td>
<td>X9</td>
<td>X10</td>
<td>X11</td>
</tr>
<tr>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
</tbody>
</table>

Table IX: Input parameters after DDEA

Objectivity in performance ranking of HEIs
method in cases where the government or private group wants to fund the institution among a single-tier institute or in the same league of institutes. We have illustrated our methodology using a group of private B-schools in India. However, the methodology may be applied across the globe, as we have taken global indicators with the proposed model. There may be cases where a few indicators may change slightly based on particular demographics, but the model is generic and with slight modifications in the choice of indicators, the model may be applied across the globe. Ranking factors considered for medical, engineering and management schools are different and must be considered so by choosing appropriate inputs and outputs.

Efficiency in HEIs is of primary importance today, more important in developing countries like India where the numbers of such institutions have grown exponentially over the years; however, efficiency and quality levels in these institutions are highly questionable. Nevertheless, on the positive side, there is increasing consciousness about improving efficiency and quality measure in these institutions, whether it is the government, policy-makers, private promoters, students, parents or the industry and society at large. While the government, private promoters or policy-makers are aspiring to establish institutions of excellence, benchmarking global institutions, students and parents are looking for institutions which are efficient and deliver quality. Similarly, industry and the society look for institutions which create relevant value in efficient ways.

Our attempt in this study is to use the methodology of DDEA in developing an objective measure of efficiency and ranking, with the objective of providing far-reaching implications in the higher education arena and society. Since results from the cases discussed in the paper of three DMUs operating in India in three different contexts show objective variations in efficiency levels across the three DMUs, it implies that the government or the policy-makers can use this methodology for effective planning and strategic interventions, identifying thrust areas across DMUs. In fact, very recently in India, the government has decided to allocate funds to institutions for development and growth depending on how efficiently the institution has performed. The methodology discussed in the paper has effective implications for fund disbursements and to identify efficient institutions. This methodology overcomes the limitations of present ranking methods and frameworks, which contrary to the input–output variables approach of the framework discussed in this paper, use a mixed approach where the input and output variables are clubbed together and are scored to get ranks for institutions and therefore are found to be less objective and create plenty of ambiguities in the minds of stakeholders. With DDEA methodology, there shall be a more objective and scientific ranking of educational institutions, helping students, parents, society and the industry to have the right judgment about the institution.

<table>
<thead>
<tr>
<th>DMUs</th>
<th>Y1</th>
<th>Y2</th>
<th>Y3</th>
<th>Y4</th>
<th>Y5</th>
<th>Y6</th>
<th>Y7</th>
<th>Y8</th>
<th>Y9</th>
<th>Y10</th>
<th>Y11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>37.9852</td>
<td>1.21</td>
<td>2.81608</td>
<td>1.84</td>
<td>39.5035</td>
<td>4.7338</td>
<td>2.827</td>
<td>13.978</td>
<td>2.3755</td>
<td>1</td>
<td>1.21</td>
</tr>
<tr>
<td>2</td>
<td>49.636</td>
<td>3.0706</td>
<td>2.99899</td>
<td>1</td>
<td>52.5277</td>
<td>7.1236</td>
<td>1</td>
<td>27.5755</td>
<td>13.7302</td>
<td>3.4108</td>
<td>2.9278</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DMUs</th>
<th>Rank/efficiency (DDEA)</th>
<th>Rank in NIRF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.439952</td>
<td>Not included</td>
</tr>
<tr>
<td>2</td>
<td>0.439952</td>
<td>Not included</td>
</tr>
<tr>
<td>3</td>
<td>2.120096</td>
<td>42nd</td>
</tr>
</tbody>
</table>

Table X. Output parameters after DDEA

Table XI. Ranking of efficiencies of DMUs
promoters of private institutions, the present ranking methods provide misplaced perceptions about their institute ranking and therefore their strategic focus is found to be misplaced. A more scientific way of looking at efficiency levels using the input–output variables, method and the DDEA can help promoters realistically understand critical thrust areas to improve efficiency levels.

5. Conclusion and future direction
Expectations of the stakeholders from HEIs put a high premium on their efficiency levels. Increasing opportunities globally on the one hand and competitive forces on the other, along with the expectations of stakeholders, have established an urge among institutions toward qualitative excellence and increased efficiency. These realities have triggered initiatives of measuring efficiencies and ranking of institutions with the objective of providing a basis for decision-making by various stakeholders such as government, private promoters, students, parents, faculty, industry and others. In an emerging country like India, where the education sector is growing in leaps and bounds, quality and efficiency are huge concerns and therefore, measures of efficiency are highly required. No doubt, there are agencies that take an independent outsider’s view and rank these institutions on certain parameters, but such ranking or measurement methodologies are found to be less objective and ambiguous. Across all the rankings, the methodology adopted is to use the average of the scores of different parameters. For example, the ranking models put inputs going in and output coming out of the institution together in its value creation process which is a serious limitation.

Efficiency measurement of a non-for-profit organization is a nonparametric situation. HEIs are not-for-profit organizations. DEA is a well-established nonparametric method to compare and rank various DMUs. In HEIs' setting, it is assumed that policy interventions are generally annual and discrete, and the effect of this policy intervention is gauged annually. It is more scientific and realistic to look for a method that incorporates the partial contribution of each year data, and higher weight should be given to the previous years. The DDEA model has therefore been used to evaluate efficiencies under the influence of lag effects. The DDEA model captures the intertemporal relationship between inputs and outputs in dynamic production processes, thus providing a more objective and realistic setting.

In this paper, the outcome of the methodology DDEA is used to lead to a better indicator for efficiency ranking than the absolute ranking on individual DMU input–output efficiency. This paper has developed a methodology to measure efficiencies considering the time factor using DDEA, and the paper has been positioned to provide a holistic approach to rank the HEIs with an example from business schools in India.

The system of performance evaluation, especially when agencies disburse performance funding, should have a holistic approach to appropriate the various evaluating parameters. Older methods of performance evaluation of HEIs were based on subjectivity and had various biases such as perception and demographics. The proposed method can nullify subjectivity and demographic biases. The method appropriates the input and output parameters. In the future, policy-makers may use the proposed method for rewarding various HEIs. In addition, future extension of the method may include sensitivity analysis of the method, and further modification in input and output variable types and various ways to benchmark input parameters.

References


Corresponding author
Anup Kumar can be contacted at: anunewin@gmail.com
High performance work practices and organizational performance—mediation analysis of explanatory theories

Naval Garg
University School of Management and Entrepreneurship, Delhi Technological University, Delhi, India

Abstract
Purpose – Although high-performance work practices (HPWPs) are considered to have a strong influence over organizational performance, researchers are not unanimous about the exact mechanism through which the impact of HPWS transcends to organizational performance. The purpose of this paper is to explore two explanatory theories (job characteristics theory and psychological impact theory) of HRM and examine their possible mediation effect on the relationship between HPWPs and organizational performance.

Design/methodology/approach – Structural equation modeling was used to examine the mediation effect.

Findings – Four constituents of job characteristics theory (autonomy, feedback, skill variety and task significance) and two constituents of psychological impact theory (job satisfaction and organization citizenship behavior) reported partial mediation.

Originality/value – The paper is based on primary data collected by author.

Keywords Organizational performance, Job characteristics

Introduction
Since the concept of high-performance work practices (HPWPs) has emerged in 1990s (e.g. Arthur, 1994; Robinson et al., 1994; Huselid, 1995; Becker and Gerhart, 1996; Delaney and Huselid, 1996; Delery, 1998), a large number of researchers have explored the linking pin between HPWPs and organizational performance (Obeidat et al., 2016). The association has been studied with the help of various conceptual approaches (Tsai, 2006). These frameworks include: resource-based view, expectancy theory, ability-motivation-opportunity model, social exchange theory, behavioral theory, resource dependency theory, labor process theory, human capital theory, strategic management, etc. (Delaney and Huselid, 1996; Guest, 1997; Appelbaum et al., 2000; Jiang et al., 2012; Darwish and Singh, 2013; Obeidat et al., 2016; Ogbonnaya and Valizade, 2018). Majorly, the theories link HPWPs with both firm-level outcomes and employee-level outcome (Tsai, 2006) through their impact on employees' skills, knowledge, values, ability, opportunities, motivation, communication, participation, relevance and flexibility (Arthur, 1994; Delaney and Huselid, 1996; Guest, 1997; Ichmiowski et al., 1997; Whitfield, 2000; Obeidat et al., 2016).

Researcher found significant empirical evidence of the linkages between HPWPs and a wide range of indicators of organizational performance (Wright and Kehoe, 2008). These indicators include financial performance (Guthrie, 2001; Shih et al., 2006; Sun et al., 2007), total quality management capabilities (Nasim et al., 2014), employee productivity (Huselid, 1995; Guthrie, 2001; Garg and Punia, 2017), different types of employees commitment (Robinson and Morrison, 2000), employee turnover rates, employees participation in decision-making process (Tsai, 2006; Wright and Kehoe, 2008), organizational citizenship behavior (OCB) (Zhao et al., 2007; Garg, 2013a), labor productivity (Datta et al., 2005), employee competence, attitudes and motivation (Huselid, 1995; Jiang et al., 2012; Punia and Garg, 2012), voice suggestions for productivity and quality improvements, employee
subjective well-being (Fan et al., 2014), innovative work behaviors (Fu, 2013), empowerment climate (Garg, 2015b) and lower employee turnover (Selden and Sowa, 2015) and competitive advantage (Wright and Gardner, 2000).

Despite extensive empirical and theoretical research works, few theoretical questions remain unanswered in the domain of HPWPs (Obeidat et al., 2016). These theoretical challenges are absence of a widely accepted definition of HPWPs, the conceptualization of HPWPs to identify the HR practices that will form part of a wider high-performance work system, finding an empirical answer to the debate of integrationist vs isolationist approach of HPWPs, etc. There is no consensus over the conceptualization of HPWPs (Becker and Gerhart, 1996; Wood, 1999; Boselie et al., 2005; Punia and Garg, 2012; Obeidat et al., 2016). Thoroughly researched and well-evidenced components of HPWPs are required to explore the magnitude of individual and combined impact of HR practices on organizational performance (Obeidat et al., 2016; Garg and Punia, 2017). The lack of conceptualization also limits comparative studies as different authors adopt different set of HPWPs ( Purcell and Kinnie, 2007). Further, few recent scholars (e.g. Takeuchi et al., 2009; Chuang and Liao, 2010; Chang and Chen, 2011; Obeidat et al., 2016) suggested the exploration of mediation effect of employee-related outcomes to unlock the “black box” of the linkage between HPWPs and organizational performance. Again, such studies require an appropriate modeling of HPWPs. Conceptualization challenge also restricts answer to debates of universal vs contingency and also of integrationist vs isolationist.

The present study tends to contribute to existing literature by proposing two multi-dimensional models of HPWPs and organizational performance. The first model conceptualizes HPWPs and organizational performance according to five dimensions of job design- autonomy, feedback, skill variety, task identity and task significance (job characteristics theory). On the other hand, the second model looks at conceptualization with the help of two dimensions – job satisfaction and OCB (psychological impact theory). Moreover, it has been argued that the most of the research studies have investigated the linkage between HPWPs and organizational performance in the firms operating in USA ( Boxall and Macky, 2009). Generalization of linkage requires authentication in Asian companies ( Wright et al., 2005; Shih et al., 2006; Singh, 2004). Therefore, the present study chooses Indian companies for the appropriate conceptualization of HPWPs.

The following sections of the paper extensively reviewed existing literature in the field of HPWPs and organizational performance in light of job characteristic theory and psychological impact theory. Research methodology has been outlined before results, discussion and conclusion of the study.

Theoretical framework
The linkage between HPWPs and organizational performance has been advocated by different schools of thoughts, namely configurational, contingency, universalists, mutual gain and integrationist (Delery and Doty, 1996). Universalistic perspective favored the positive relationship between HPWPs and organizational benefits regardless of organizational context. HR practices create organizational value in a different business context. In nutshell, the perspective supports the fact that an HR practice is always beneficial for the organization irrespective of organizational values (Pfeffer, 1994; Huselid, 1995). Integrationist argued that organizational benefit of HR practices multiplies when the practices are implemented in complementary bundle. Contingency perspective advocates the aligning of HPWPs with strategic choice of the organization. They favored a vertical fit between HRM practices, organizational strategy and environmental characteristics to achieve competitive advantage in the market (Youndt et al., 1996). Further, configurational perspective demanded both vertical and horizontal fits of HR practices. Horizontal fit refers to the internal consistency of the practices. Although different lines of thoughts explored
the domain of HPWPs, there is little consensus over the HR practices to be included in a model of HPWPs.

The researchers’ interest in undertaking studies on linkage between HPWPs and organizational performance is adversely affected by conceptualization problem (Purcell and Kinnie, 2007). The issue remains a major roadblock (Obeidat et al., 2016), as one of the pioneer researchers in the field of HPWPs, Guest (1997, p. 274) stated that “only when we make progress in measuring the independent and dependent variables can we begin to give full attention to the way in which they are linked.” The inconsistencies in conceptualization could be clearly understood from diverse definitions of HPWPs. “HPWPs are a set of complementary work practices covering three-broad categories which include; high employee involvement practices, human resource practices, reward and commitment practices” (Sung et al., 2005, p. 4). Patel and Conklin (2012, p. 210) defined HPWPs as those “set of employee management practices that positively affect employees’ attitudes, motivation, and performance.’” Different terminology for HPWPs, high-commitment management (Gould-Williams, 2004), innovative workplace practices (Askenazy, 2001), high-involvement management (Guthrie, 2001; Mohr and Zoghi, 2008) and alternative work practices (Godard, 2001) also reflects the intensity of conceptualization challenge.

Hackman and Oldham’s (1976) job characteristics theory explains work conditions which keep employees intrinsically motivated while fulfilling organizational responsibilities. “Job characteristics are the systems or situational factors affecting the psychological and attitudinal condition of employees” (Johari and Yahya, 2016, p. 555). The theory supports that the enriched and significant job characteristics lead to motivating, positive and stimulating emotional, psychological and cognitive job conditions. The five dimensions included in job characteristics are skill variety, task significance, task identity, feedback and autonomy. Job characteristics affect three psychological conditions of the employees: responsibility, feedback and meaningfulness of work (Hackman and Oldham, 1976; Morgeson and Campion, 2003; Wood et al., 2012; Ling and Toh, 2014; Zhao and Ghiselli, 2016).

Skill variety is one of the most important job characteristics that lead to motivated workforce. The factor could lead to enhanced employee performance as it allows employee to utilize a variety of skill set at workplace (Johari and Yahya, 2016). Indulgence in different activities and thereby using various skills and talents may lead to positive attitude and behavioral outcomes (Krasman, 2012; Ghosh et al., 2015). Skill variety is associated with positive job performance (Sulea et al., 2012).

Researchers of job design domain have conceptualized “task significance” as a pivotal objective of the work itself (Steers and Mowday, 1981). Task significance promotes sense of meaningfulness as employees could recognize their work as relevant, important and beneficial for achieving organizational goals (Hackman and Lawler, 1971; Hackman and Oldham, 1975). A growing number of research works argue that the manipulation of task significance ultimately increases job performance. Task significance promotes job performance through promoting low attrition rate (Rentsch and Steel, 1998), high level of job and organizational engagement, and increased tendency of pro-social behavior (Grant, 2007).

Hackman and Oldham (1975, p. 161) defined “task identity” as “the degree to which the job requires completion of a ‘whole’ and identifiable piece of work – that is, doing a job from beginning to end with a visible outcome.” The factor advocates whole, complete and holistic work profile that provides an opportunity of engagement in meaningful and representative work (Park, 2017). It has a direct impact on employees’ sense of responsibility, locus of control and loyalty (Johari and Yahya, 2016). Employees having an identifiable large job design are more likely to perform better than employees confined to a narrow work flow (Uruthirapathy and Grant, 2015).
Job autonomy is the degree of freedom and discretion provided to employees to make work-related decisions at workplace (Shirom et al., 2006; Park and Searcy, 2012). Job autonomy leads to the optimum utilization of human capital and hence it is a vital factor for better performers (Park and Searcy, 2012). Higher job autonomy develops positive feelings related to work and such constructive perception results into a higher level of desirable and effective work behavior (Bontis et al., 2011; Krasman, 2012). Previous researchers have found empirical evidence of the positive relationship between job autonomy and employee and organizational-related outcomes e.g. “job satisfaction” (Dodd and Ganster, 1996), “individual performance” (Piccolo and Colquitt, 2006), “desirable work behavior” (Hassan, 2014), “organizational citizenship behavior” (Chiu and Chen, 2005), “employee commitment” (Parker, 2003; Ahuja et al., 2007), “in-role performance” (Bakker et al., 2004) and “better mental health” (Park and Searcy, 2012; Parker, 2003).

Hackman and Oldham (1975, p. 162) defined feedback as “the degree to which carrying out the work activities required by the job results in the employee obtaining direct and clear information about the effectiveness of his or her performance.” Feedback helps employees to develop a deeper understanding of their task. This understanding helps employees in better regulation and monitoring of their work (Hattie and Timperley, 2007). Performance tends to increase with positive feedback (London, 2003; Leslie and Taylor, 2005; Johari and Yahya, 2016).

A number of empirical explorations (Dodd and Ganster, 1996; Bhuian et al., 1996; Bhuain et al., 2001; Bhuain and Menguc, 2002; Garg and Lal, 2016) reported that enriched and significant job engineering allows employees to use different skill set at workplace, nurtures feeling of empowerment through autonomy and ensures adequate feedback. In nutshell, enriched and motivating jobs are associated with positive performance. Based on these studies, following theoretical model is proposed. Author proposes that job characteristics theory mediates the relationship between HPWPs and organizational performance (Figure 1).

Psychological impact theory
Psychological impact theory states that HR practices may lead to a change in employee attitude which may become one of the reasons of organizational performance. Numerous empirical evidences link the implementation of HPWPs with employee attitude measures like job satisfaction, employees’ commitment, OCB, etc. Scholars have found positive links between HPWPs and employee attitude measures such as organizational commitment (e.g. Bryson and White, 2008; Newman et al., 2011). Authors like Mohr and Zoghi (2008), Wood and de Menezes (2011), etc. concluded a positive association between job satisfaction and employee oriented human resource practices. OCB could be elicited from employee through the institutionalization of an appropriate HR system (Paré and Tremblay, 2007).

![Figure 1. Proposed model of job characteristics theory](image-url)
HPWPs provide ample opportunities to employees through which employee starts perceiving oneself as an valuable asset of the organization. This enhances employee participation at work place and consequently employee start utilizing the optimum level of their skill, knowledge and abilities (Wood and de Menezes, 2011; Atkinson and Hall, 2011). Thus, the sense of meaningful work results into positive work-related attitudes and behaviors toward the organization (Wood and de Menezes, 2011). Various constituents of a high performance work system such as information sharing, self-managed teams and incentive-based pay, etc., sponsor OCBs. Moreover, OCB elicit positive reactions from employees (Batt and Valcour, 2003; Evans and Davis, 2005; Nishii et al., 2008; Snape and Redman, 2010). Takeuchi et al. (2009) also reported a positive relationship between a composite measure of HPWPs and employee job satisfaction and affective commitment. Harter et al. (2002) concluded that HPWPs lead to employee satisfaction which influences firm-level outcomes (Snape and Redman, 2010).

Thus, researchers have found conclusive evidence for linkage between HPWPs and job satisfaction and OCB; in the same manner, job satisfaction and OCB have also been found positively associated with organizational performance. Satisfied work force makes organizations more effective than less satisfied employees (Ostroff, 1992). Earnings per share and market performance of an organization increase with the satisfaction level of the employees (Evans and Jack, 2003). Job satisfaction leads to higher “Return on Asset” (ROA) (Schneider et al., 2003). Various other authors also reported a positive association between job satisfaction and organizational performance (e.g. Koys, 2001; Harter et al., 2002; Evans and Jack, 2003; Zohir, 2007; Mafini and Pooe, 2013; Latif et al., 2015).

Further, OCB has been also found to have a positive association with organizational performance. Organizational performance is affected by OCB in seven different domains (Podsakoff et al., 2000). These seven areas include: improvement in work efficiency of employees, supervisors and managers; prioritization of scarce organizational resources to more productive endeavor; reduction in expenditure for routine functions of the organization; better cooperation and coordination between intra and inter group members; better adaption to changing environment; giving more stability to the organization; and helping a firm in talent attraction and retention. Authors like Sadeghi et al. (2016), Bambale (2011), Tai et al. (2012), Kark (2004), Podsakoff, Winer (2001) reported a positive correlation between OCB and organizational performance.

A few studies like Zhang and Morris (2014) suggested that the employee-level outcomes positively mediate the link between HPWPs and organizational performance. In the light of above literature, following hypothetical model was developed. Author proposes that psychological impact theory mediates the relationship between HPWPs and organizational performance (Figure 2).

**Research framework**
The aim of the present study is to explore the conceptualization problem of HPWPs with the help of two theories: job characteristics theory and psychological impact theory. Conceptualization issue has been addressed through investigating meditation effect of the theories in relationship of HPWPs and organizational performance. The research setting
for this empirical analysis is companies located in National Capital Region (NCR), India. The research setting is appropriate as modern organizations and MNCs which are implementing HPWPs are readily available in NCR region. Random sampling was applied to collect data from respondents. As many as 40 companies, banks, insurance educational institute, car, rice, shoe, business processing organizations, etc. were chosen and it was targeted to access information from 20 responses from all 40 companies. Respondents were approached through e-mail, personal interaction and through references (friends or colleagues). Author got 432 responses from 800 approached respondents and thus response rate was comfortably higher than 50 percent. Demographic representation of sample is presented in Table I.

The study is based upon an exploratory-cum descriptive research design. Primary data were collected through a structured questionnaire. Part A of the questionnaire captured the respondents’ demography such as age, gender, work experience and educational qualifications and these demographical variables were analyzed using descriptive statistics. Part B comprised 35 HPWPs taken from exploratory study of Punia and Garg (2012). Punia and Garg (2012) derived a set of nine broader dimensions from 35 HPWPs. These nine dimensions of HPWPs are psycho-strengthening HPWPs, procedural improvement related HPWPs, employee empowerment related HPWPs, social and safety need driven HPWPs, value creating HPWPs, employee engagement related HPWPs, team-oriented HPWPs and traditional HRM practices and reward-oriented HPWPs. These dimensions constitute different number of performance practices (employee empowerment related dimension comprise of employee stock option plan, management by objective and quality circle; procedural improvement related dimension comprise of staff suggestion team and team briefing). Cronbach’s α value of all nine dimensions was recorded within acceptable level and thus reliability of scale was ensured. Part C measured psychological impact through measuring job satisfaction (scale developed by Cammann et al. (1983) and OCB (scale used by Smith et al. (1983). Job satisfaction scale accesses employee’s satisfaction level on a seven-point scale ranging from 1 (strongly disagree) to 7 (strongly agree). OCB scale comprises of 16 items (Like – Helps other who have been absent, punctuality, assist supervisor with their work, etc.) that measure OCB using a five-point Likert scale. Again, the value of Cronbach’s α confirms reliability of both scales. Part D measured organizational performance with the help of scale of Singh (2004). Part E measured job characteristics model with the help of scale developed by Morgeson and Humphrey (2006). The scale comprises of 21 items which assess the model on a five-point scale ranging from 1 (strongly disagree) to 5 (strongly agree). These 21 items were divided into four major constructs namely task characteristics, social characteristics, work context and knowledge characteristics. Present study explores mediation with the help of structure equation modeling using AMOS.

Data analytical strategy
Correlation tables were produced for all variables of the study. Structural equation modeling was used to examine mediation effect of two explanatory models of HPWPs.

Table I.
Demographic profile of sample

<table>
<thead>
<tr>
<th>Category</th>
<th>Gender</th>
<th>Age group</th>
<th>Experience (year)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (54%, 233), F (46%, 199)</td>
<td>&lt; 25 (13%, 56), 26–35 (42%, 181), 36–45 (21%, 91), 46–55 (15%, 65), &gt; 55 (0%, 39)</td>
<td>&lt; 5 (42%, 181), 5–10 (31%, 117), &gt; 10 (27%, 134)</td>
<td>Total (100%, 432)</td>
</tr>
</tbody>
</table>
Psycho-strengthening HPWPs, procedural improvement related HPWPs, employee empowerment related HPWPs, social and safety need driven HPWPs, value creating HPWPs, employee engagement related HPWPs, team-oriented HPWPs and traditional HRM practices and reward-oriented HPWPs as manifest variables. Other latent variables: organizational performance, autonomy, feedback, skill variety, task identity and task significance were modeled with their respective subscales as a manifest variable. Error terms were suitably added with every variable of the study. Proposed model fit did not come out as a good fitted model ($\chi^2$ value ($df = 329, n = 432$) = 723 $p < 0.001$, $\chi^2/df = 2.19$, RMR = 0.23, CFI = 0.77 and RMSEA = 0.24). Different statistical values of the model indicated the need of model revision for a good model fit. Loading of four HPWPs (semi-autonomous team, attitude survey, family insurance and staff suggestion scheme) was not significant, and they were removed from the model. Resultant model showed a moderate fit model ($\chi^2$ value ($df = 238, n = 432$) = 385 $p < 0.001$, $\chi^2/df = 1.61$, RMR = 0.11, CFI = 0.88, RMSEA = 0.09). Revised model looks like the proposed one, but as discussed above, four latent variables have been removed for better model fit.

Mediation analysis of job characteristics theory (see Table II)
To investigate the mediation effect of five variables of job characteristics theory (autonomy, feedback, skill variety, task identity and task significance), regression values of direct, indirect and total effects were calculated. HPWPs reported a significant effect on organizational performance (direct $\beta = 0.60$, $p < 0.05$). Similarly, HPWPs exerted a significant direct effect on autonomy (direct $\beta = 0.66$, $p < 0.05$), feedback (direct $\beta = 0.56$, $p < 0.05$), skill variety (direct $\beta = 0.44$, $p < 0.01$), task identity (direct $\beta = 0.53$, $p < 0.05$) and task significance (direct $\beta = 0.55$, $p < 0.05$). Further, all hypothesized mediating variables exerted a significant direct effect on dependent variable (organizational commitment). Direct relationships of autonomy and organizational performance (direct $\beta = 0.42$, $p < 0.05$), of feedback and organizational performance (direct $\beta = 0.46$, $p < 0.05$), of skill variety and organizational performance (direct $\beta = 0.33$, $p < 0.01$), of task identity and organizational performance (direct $\beta = 0.38$, $p < 0.01$) and of task significance and organizational performance (direct $\beta = 0.33$, $p < 0.01$) were concluded significant.

Further, HPWPs exerted a significant indirect effect on organizational performance (indirect $\beta = 0.30$, $p < 0.01$) after being mediated by autonomy. Indirect effect explained 33.3 percent of total effect of independent variable on dependent variable being mediated by autonomy (total $\beta = 0.90$, $p < 0.001$). The result confirmed a partial mediation effect of autonomy. HPWPs exerted a significant direct (direct $\beta = 0.56$, $p < 0.05$) and indirect (indirect $\beta = 0.31$, $p < 0.01$) effect on feedback. The calculated indirect effect resulted in 31.9 percent of the total effect of HPWPs on performance being mediated by feedback (total $\beta = 0.97$, $p < 0.001$). Additionally, predictor variable exerted significant indirect effects on criteria variable (indirect $\beta = 0.28$, $p < 0.01$), resulting in 38.8 percent of the total effect of HPWPs on organizational performance being mediated by skill variety (total $\beta = 0.72$, $p < 0.05$).
Identically, task significance also reported partial mediation in the relationship of HPWPs (indirect \( \beta = 0.37, p < 0.01 \) and total \( \beta = 0.92, p < 0.001 \)). Indirect effect explained 40 percent of the total effect of HPWPs on organizational performance being mediated by task significance. However, indirect of HPWPs on performance when mediated by task identity is insignificant (indirect \( \beta = 0.26, p > 0.001 \)) and thereby it rules out any mediation effect of task identity. In nutshell, four constructs of job characteristics theory: autonomy, feedback, skill variety and task significance, partially mediate the association of HPWPs and organizational performance. Task identity does not have any mediation effect.

**Theoretical model of psychological impact theory (Figure 2) and revised model (Figure 4)**

To investigate the presence of mediation, a theoretical model comprising all latent (HPWPs, organizational performance, job satisfaction and organization citizenship behavior) and manifest variables was constructed. Again, HPWPs (latent variable) were modeled with psycho-strengthening HPWPs, procedural improvement related HPWPs, employee empowerment related HPWPs, social and safety need driven HPWPs, value creating
HPWPs, employee engagement related HPWPs, team-oriented HPWPs and traditional HRM practices and reward-oriented HPWPs as manifest variables. Other latent variables, such as organizational performance, job satisfaction and organization citizenship behavior, were modeled with their respective subscales as a manifest variable. Error terms were suitably added with every variable of the study. Proposed model reported a good fitted model ($\chi^2$ value ($df = 163$, $n = 432$) = 286, $p < 0.001$, $\chi^2$/$df = 1.75$, RMR = 0.08, CFI = 0.89, RMSEA = 0.07).

Mediation analysis of psychological impact theory (see Table III)

To investigate the mediation effect of two variables of psychological impact theory (job satisfaction and organization citizenship behavior), regression values of direct, indirect and total effects were calculated. HPWPs reported a significant effect on organizational performance (direct $\beta = 0.58$, $p < 0.05$). Similarly, HPWPs exerted a significant direct effect on job satisfaction (direct $\beta = 0.54$, $p < 0.05$) and on organization citizenship behavior (direct $\beta = 0.66$, $p < 0.05$). Further, all hypothesized mediating variables exerted a significant direct effect on dependent variable (organizational commitment). Direct relationships of job satisfaction and organizational performance (direct $\beta = 0.70$, $p < 0.05$) and of organization citizenship behavior and organizational performance (direct $\beta = 0.61$, $p < 0.05$) were concluded significant.

Further, HPWPs exerted a significant indirect effect on organizational performance (indirect $\beta = 0.34$, $p < 0.05$) after being mediated by job satisfaction. Indirect effect explained 36.9 percent of total effect of independent variable on dependent variable being mediated by autonomy (total $\beta = 0.92$, $p < 0.05$). The result confirmed the partial mediation effect of autonomy. HPWPs exerted a significant direct (direct $\beta = 0.54$, $p < 0.05$) and indirect (indirect $\beta = 0.40$, $p < 0.05$) effect on organization citizenship behavior. The calculated indirect effect resulted in 42.5 percent of the total effect of HPWPs on performance being mediated by organization citizenship behavior (total $\beta = 0.94$, $p < 0.05$). In nutshell, two constructs of psychological impact theory, organization citizenship behavior and job satisfaction, partially mediate the association of HPWPs and organizational performance.

Discussion

The present study extends the knowledge and understanding in the domain of HPWPs. First, findings confirm the positive relationship between HPWPs and organizational performance.

<table>
<thead>
<tr>
<th></th>
<th>$\beta$</th>
<th>$B$</th>
<th>SE</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Direct effect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPWPs $\rightarrow$ OP</td>
<td>0.58</td>
<td>0.60</td>
<td>0.78</td>
<td>$&lt;0.05$</td>
</tr>
<tr>
<td>HPWPs $\rightarrow$ JS</td>
<td>0.54</td>
<td>0.63</td>
<td>0.80</td>
<td>$&lt;0.05$</td>
</tr>
<tr>
<td>HPWPs $\rightarrow$ OCB</td>
<td>0.66</td>
<td>0.68</td>
<td>0.77</td>
<td>$&lt;0.05$</td>
</tr>
<tr>
<td>JS $\rightarrow$ OP</td>
<td>0.70</td>
<td>0.77</td>
<td>0.84</td>
<td>$&lt;0.05$</td>
</tr>
<tr>
<td>OCB $\rightarrow$ OP</td>
<td>0.61</td>
<td>0.73</td>
<td>0.69</td>
<td>$&lt;0.05$</td>
</tr>
<tr>
<td><strong>Indirect effect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPWPs $\rightarrow$ JS $\rightarrow$ OP</td>
<td>0.34</td>
<td>0.38</td>
<td>0.72</td>
<td>$&lt;0.05$</td>
</tr>
<tr>
<td>HPWPs $\rightarrow$ OCB $\rightarrow$ OP</td>
<td>0.40</td>
<td>0.43</td>
<td>0.87</td>
<td>$&lt;0.05$</td>
</tr>
<tr>
<td><strong>Total effect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPWPs on OP (JS)</td>
<td>0.92</td>
<td>0.96</td>
<td>0.81</td>
<td>$&lt;0.05$</td>
</tr>
<tr>
<td>HPWPs on OP (OCB)</td>
<td>0.94</td>
<td>0.94</td>
<td>0.76</td>
<td>$&lt;0.05$</td>
</tr>
</tbody>
</table>

**Notes:** $\beta$, Standardized coefficient; $B$, unstandardized coefficient; HPWPs, high-performance work practices; OP, organizational performance; JS, job satisfaction; OCB, organization citizenship behavior.

Table III. Direct, indirect and total effect of HPWPs on organizational performance
performance in a non-western setting, i.e. India. Although previous researchers (e.g. Youndt et al., 1996; Delery and Doty, 1996; Black and Lynch, 2001; Shih et al., 2006; Obeidat et al., 2016) have already applauded the said linkage but similar findings in Indian sample are interesting and demand further exploration and discussion given the vast differences in cultural and organizational settings in India. The paper tried to address the concern of lack of studies in Asian countries (as highlighted by Obeidat et al., 2016) via investigating HPWP's in Indian companies. Although it is extremely premature to suggest, the acceptance of mostly similar set of high-performance practices in eastern world favors universal approach rather than the contingency approach of HPWP's. More structured and focused researches are required to reveal conclusive evidence for the same. Universal perspective of HPWP’s adores the idea of universal validity and utility of HR practices irrespective of organizational and cultural factors.

Second, results are quite encouraging for country like India where the notion of HPWP’s is still in a developing phase. The result certainly addresses the concerns and apprehensions of HR managers and practitioners in the country. A healthy mix of imported HPWP’s (like autonomous teams, management by objective), universally practices HR functions (like scientific recruitment and selection, training and development, regular performance appraisal) and of Indian specific practices (like yoga and meditation, social security measures) highlights the success story of globalization which started three decades back in India. Collective contribution of imported and Indian practices suggests that Indian industry has welcomed government’s efforts (like Foreign Direct Investment, Foreign Institutional Investment, Make-in-India program, Ease of doing easing initiative) in the line of free cross border business transactions.

Further, the study makes a sincere attempt to investigate the conceptualization problem of HPWP. It definitely increases the understanding of delicate relationship between HPWP’s and organizational performance. Specifically, the study explores two multidimensional models of HPWP’s and organizational performance. The models suggest bundling of HPWP’s using job characteristics theory and psychological impact theory. The argument is examined through analyzing the mediation effect of these theories on the relationship between HPWP’s and organizational performance. The paper finds empirical evidence for two stated propositions and the mediation effect of job characteristics theory and psychological impact theory is confirmed. It means a model of HPWP’s that ensures a significant job profile, allows the utilization of variety of skills, provides greater autonomy and feedback at workplace leads to a higher level of organizational performance. Similarly, organizational performance enhances with the institutionalization of a model of HPWP’s that encourages organization citizenship behavior and provides a greater level of job satisfaction to employees. An overview of previous studies reveals uncertainty over mechanism through which the effect of HPWP’s transcends to organizational performance. The present study provides empirical evidence of the intervening effect of job characteristics and psychological impact theory. Therefore, present study tries to fill the existing research gap in the domain of performance studies.

Further, researchers do not have a unified view on how to develop a model of high-performance practices. These different perspectives of HRM throw some light on the issue. HPWP’s should be designed in such a way that it results into more responsibility and autonomy (job characteristics perspective) and it should take care of employees’ attitude and well-being (psychological impact perspective). Model of HPWP’s developed with the help of these perspectives contributes in building a workforce, which is more committed, more satisfied with their jobs and which maintains work-life balance. This leads to enhanced performance and involvement of employees, reduced absenteeism and turnover. Higher performance could be elicited by ensuring more challenges, more responsibility, more autonomy and more opportunities for employees at work place.
By adopting an extensive range of innovative HPWPs, organizations are able to exploit existing complementarities among such practices to maximize their beneficial independent properties.

The findings are quite significant and important for countries like India, which has drastically different cultural and industrial relations compared to the Western world. According to Gesteland (2005), business culture could be categorized as: deal focused and relationship focused. Eastern countries like India are categorized in "relationship focused" where employees demand social affiliation, mutual trust, and harmonious long-lasting relationships (Gupta and Bhaskar, 2016). Western countries like USA are "deal focused" where employees are more professional, task-specific, direct, and aggressive (Gupta and Bhaskar, 2016). Indian employees prefer job contentment and healthy association with colleagues, while Western employees prefer task accomplishment and target achievement (Lees and Khatri, 2010). Now, any attempt of the implementation of HPWPs requires a serious understanding of cultural context of the country. Paper appropriately acknowledges job satisfaction (contentment at work) and OCB (friendly, social and proactive social relations).

Given the preliminary nature of this study, further research needs to be conducted on various conceptual models of HPWPs, their successful implementation, and performance relationships. We approached mediation through a cross-sectional approach. No researcher can confidently claim causal relations based on cross-sectional studies. Thus, results of present studies could be confirmed through any longitudinal exploration. In addition to discussed perspectives, future studies could explore the mediation effect of individual differences (attitude, belief system, values). The study is marred with few limitations. First, a larger sample size would have made the study more worthwhile. Second, owing to the lack of previous observations in case of a number of theories, results of the study could not be compared for continuity or deviation. Another limitation of the study is subjective organizational performance measure. Future research works could include both financial and non-financial measures to measure organizational performance. In conclusion, adequate investment in employees propels the laws of reciprocity as establishments who have embraced and embedded this concept in their culture will reap that which they have sown.

References


Bambale, A.J. (2011), Understanding Significant Relationships between Organizational Citizenship Behavior (OCB) and Marketing Function, Faculty of Commerce and Management Studies, Kelaniya.


Further reading


Corresponding author
Naval Garg can be contacted at: naval_garg123@yahoo.co.in

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm Or contact us for further details: permissions@emeraldinsight.com
Cognitive leader–member exchange differences between supervisors and subordinates

Yan Shen
Dhurakij Pundit University, Bangkok, Thailand and Shandong Yingcai University, Jinan, China

Abstract

Purpose – The purpose of this paper is to reveal the influence of organizational identification (OI) on organizational citizenship behavior (OCB) directed toward organization (OCB-O) and OCB directed toward individuals (OCB-I), and explore the moderating effects of turnover intention (TI) and the differences of the moderating effects of supervisors’ ratings of leader–member exchange quality (LMX-L) and subordinates’ ratings of LMX quality (LMX-S) in the process.

Design/methodology/approach – Based on social exchange theory, this study took administrative staff of private small and medium enterprises in China as population. A paired-samples design was used, and 437 valid paired samples were collected finally. Data analyses were conducted by using structural equation modeling.

Findings – OI indicated a positive influence on both OCB-O and OCB-I. TI showed a negative moderating effect on the positive relationship between OI and OCB-I, but no significant moderating effect on the positive relationship between OI and OCB-O. LMX-L and LMX-S were not in the same direction and showed different moderating levels. LMX-L had no significant effects on the relationship between OI and OCB-O, nor did on the relationship between OI and OCB-I. On the other hand, LMX-S strengthened both the positive relationship between OI and OCB-O, and the positive relationship between OI and OCB-I.

Originality/value – This study explored the relationships from the social exchange perspective, found that OCB-O and OCB-I had different formation mechanisms and should be explored as distinct variables, and confirmed that it was necessary to use the paired-samples design when studying dyadic phenomena.

Keywords OCB, LMX, Organizational identification, Social exchange theory

Paper type Research paper

1. Introduction

Past research works had shown that employees’ organizational identification (OI) is a central predictor for employees’ efforts on behalf of an organization (Riketta, 2005). Over the past two decades, research on OI is substantially increasing and investigating its effects on employees’ work behaviors (He and Brown, 2013). However, there are still five limitations as follows.

First, the conclusions about the relationship between OI and organizational citizenship behavior (OCB) are inconsistent in prior studies. Major studies find that OCB is a behavioral result of OI (Meyer et al., 2012) and OI is positively related to OCB (Ashforth et al., 2008). However, through meta-analysis, some scholars find that the correlation between OI and extra-role behaviors has become weaker (Riketta, 2005). Further, Zhang and Chen (2013) considered that there is the lack of relationship between OI and OCB, and the possibility might be that OCB motives can be quite personal and relational; another possibility might lie in the complex nature of OCB measurement, as it does not clearly distinguish between citizenship behavior directed toward organizations (OCB-O) and toward individuals (OCB-I). Thus, the current study hopes to clarify the relationship between OI and OCB-O, and the relationship between OI and OCB-I.

Second, major scholars study OCB as a whole (Callea et al., 2016; Zhang and Chen, 2013). However, from the perspective of social exchange theory (SET), in an organization, employees not only have exchange relations with the organization, but also with individuals, and OCB
should be divided into citizenship behavior that benefits the actual organization (OCB-O) and
citizenship behavior that benefits specific individuals (OCB-I) within an organization (Lee and
Allen, 2002). Therefore, the current study maintains that OCB-O and OCB-I are different
aspects of OCB and should be explored separately at the same time. OCB-O reflects employees’
deliberate attempts to restore balance within an organization (Lee and Allen, 2002). OCB-O is
defined more in terms of the “good soldier” or “good citizen” syndrome, whereby tasks are
undertaken that are “right and proper” for the system than for specific persons (Smith et al.,
1983). Thus, such acts focus on, and benefit, an organization (Williams and Anderson, 1991).
OCB-O encompasses concern for the interests of the organization as a whole (e.g. protecting the
organization, offering ideas for improvement) (Evans and Davis, 2014). In contrast, OCB-I’s
purpose is to maintain interpersonal equilibrium (Coleman and Borman, 2000; Smith et al.,
1983). Such acts concern for the interests of fellow employees (e.g. assisting others, sharing
resources) (Evans and Davis, 2014) and interpersonal facilitation (Podsakoff et al., 2000). In such
a context, OCB-I is clearly directed toward specific individuals in the workplace, excluding help
directed to supervisors. This exclusion indicates concerns about the true beneficiaries of such
help (supervisors or the organization) and the motivational complexity that arises because
supervisors often control rewards such as pay and promotion. Consequently, items that do not

Third, major studies on OI are based on social identity theory (Wu et al., 2016; Zhang and
Chen, 2013). However, social exchange relationship between employees and their
organization plays a fundamental role in determining employees’ attitudes and behaviors
(Konovsky and Pugh, 1994). Moreover, the higher the degree of OI is, the greater the extent
to which attitudes and behaviors are affected by leader–member exchange (LMX) in an
organization (Millward and Haslam, 2013). SET is the theoretical foundation of much
organizational psychology and behaviors (Fassina et al., 2008; Kim et al., 2015; Michel and
Tews, 2016). OI, OCB, LMX, turnover and retention all have the nature of social exchange.
Hence, this study takes SET as the theoretical foundation.

Fourth, major studies disregard the intermediate influential factors regarding OI’s
relationship with OCB (Evans and Davis, 2014; Wu et al., 2016). Although some studies have
paid attention to these intermediate factors (Choi et al., 2014; Lin et al., 2017), the results are
inconsistent. The possibility might be that the relationship between OI and OCB-O/OCB-I may
be not a true causality relationship (Bang et al., 2016) and perhaps be affected by other factors.

Fifth, major studies on LMX have collected data from a single level of subordinates and
have not given much consideration to the perspectives of supervisors and subordinates
(Buch et al., 2016; Coyle and Foti, 2015). However, relationships always occur in pairs. The
perception of relationship quality by two parties may be same or different. Different
perceptions have different effects on behaviors. Collecting data from a single source can cause
higher common method variance. Krasikova and LeBreton (2012) asserted that studying
dyadic phenomena (such as LMX) from the perspective of one dyad member is “theoretically
deficient.” Hence, this study uses a paired-samples design to collect data from two levels of
supervisors and subordinates to explore the differences of the effects of supervisors’ ratings of
LMX quality (LMX-L) and subordinates’ ratings of LMX quality (LMX-S).

Taken together, based on SET, the current study hopes to reveal the relationship
between OI and OCB-O/OCB-I, the moderating effects of turnover intention (TI) and the
differences of the moderating effects of LMX-L and LMX-S in the impact process of OI on
OCB-O and OI on OCB-I.

2. Literature review
2.1 Social exchange theory
SET postulates that certain workplace antecedents lead to interpersonal connections, referred
to as social exchange relationships (Cropanzano and Mitchell, 2005). Social exchange
relationships between two parties are different from those of pure economic exchange. In exchange relationships, they develop through a series of mutual exchanges that include not only material things, but also spiritual things that yield a pattern of reciprocal obligation by each party (Blau, 1964).

SET assumes that individuals return the benefits they receive, which likely engenders from them goodwill and helpfulness toward the party with whom they have a social exchange relationship (e.g. Masterson et al., 2000). Thus, when employees receive something valuable that is in their interests, they are likely to reciprocate for the benefit of an organization (Choi et al., 2014). SET proposes that individuals who receive valuable outcomes from others tend to give something back in turn because mutual reciprocity is important for maintaining good relationships (Organ, 1990). Employees seek to ensure equilibrium between what they receive from an organization and what they give in return (Iverson and Zatzick, 2011). When employees’ needs are satisfied, they have a sense of voluntary return that they are able to control (Eisenberger et al., 2001). In return, they not only reciprocate by performing contractual tasks, but they may also act in ways that are disproportionate to the original investment (Karriker and Williams, 2009), for example, long-term commitment to the organization, working proactively to help the organization achieve its goals and extra-role behaviors that benefit the organization or colleagues (Callea et al., 2016). When employees are unsatisfied with what they have got from the organization, they will reduce returns to the organization, such as reduce the degree of OI, form TI, and reduce their extra-role behaviors, in order to reduce the dissonance of exchange imbalance (Karriker and Williams, 2009).

2.2 Organizational identification and OCB

OI is defined as the perceived oneness with an organization and the experience of the organization’s success or failures as one’s own (Mael and Ashforth, 1992). In other words, OI reflects employees’ feelings of psychological inclusion within their organizations (Walumbwa and Hartnell, 2009). Employees with high level of OI tend to feel psychologically intertwined with their organizations (Mueller and Straatmann, 2014). SET suggests that employees are likely to act in ways that benefit organizations when they have positive relationships in return (Cropanzano and Mitchell, 2005). It is expected that employees with high OI are willing to engage in a positive social exchange with the organization in return for the organization’s payment, for example, honors, promotion opportunities and satisfactory remuneration from the organization. OCB-O reflects employees’ deliberate attempts to restore balance within an organization (Lee and Allen, 2002). So, in addition to performing their duties in accordance with the labor contract, they also volunteer to do extra-role behaviors that are beneficial to the organization, for example, offer ideas for improvement, take the initiative to solve hidden problems to protect the company and defend the organization when other employees criticize it. On the contrary, if employees perceive that their treatment by the organization is unreasonable, they may have a low level of OI (Callea et al., 2016). Such employees believe that their relationship with the organization is distant, and the sense of reciprocal obligation is reduced (Dutton et al., 1994). Therefore, they only do the work stipulated in the labor contract, because it is enough to reach the balance of the exchange relationship with the organization, and need not to do extra-role behaviors. Accordingly, the following hypothesis is proposed:

\[ H1a. \] OI is positively related to OCB-O.

Based on SET, according to the reciprocity of exchange, employees who identify with the organization also hope to be identified in return, and identification from colleagues is an important component. High-quality social exchange relationship can be expected to lead to OI (Choi et al., 2014). OCB-I concerns for the interests of fellow employees (e.g. assisting others,
sharing resources) (Evans and Davis, 2014), the positive cooperative behavior (e.g. altruism and courtesy) (Coleman and Borman, 2000) and interpersonal facilitation (Podsakoff et al., 2000). So we may predict that employees with high OI will show more OCB-I. On the contrary, the lower the degree of employees’ OI is, the less OCB-I among individuals. This leads to:

\[ H1b. \text{ OI is positively related to OCB-I.} \]

### 2.3 The moderating effects of TI

Among employees with TI, some identify with the organization and some do not identify with the organization because of some factors (Mishra and Bhatnagar, 2010; Moura et al., 2009). Employees with TI are concerned about the situation before leaving (Hancock et al., 2013) and have started planning for their departure. Those who have a high level of OI will continue to do something to help the organizational image (e.g. to defend the organization when other employees criticize it) because of the spiritual exchange with the organization. However, they will try to avoid material exchange with the organization other than the labor contract (e.g. to do extra-role works to keep up with developments in the organization), because direct payment cannot be made for engaging in the labors before they leave their organizations (Kim et al., 2015).

So in aggregate, they perform OCB-O less than before. Those who have low-level OI only care about what they deserve under the labor contract before they actually leave, and unlikely to exhibit OCB-O without a reward (Kim et al., 2015). In order to get the treatment stipulated under the labor contract, they only do a good job in role. As long as do a good job in role, they feel that the exchange relationship with the organization is balanced. Moreover, OCB-O is essentially unenforceable with the usual incentives or sanctions (Smith et al., 1983), and a failure to engage in OCB-O is not punished. So, they will not perform OCB-O. Therefore, it can be suggested:

\[ H2a. \text{TI moderates the relationship between OI and OCB-O, such that the positive relationship is weaker when subordinates have higher TI.} \]

Strong identification with an organization makes cooperative behavior toward other organizational members likely because of reciprocity (Kramer, 1991). Employees who plan to stay in office expect a long-term relationship with their colleagues, they value their exchange with colleagues. Employees who intend to leave their jobs have foreseen that the exchange relationship with the colleagues will end with the actual turnover, and the reciprocity will not exist. It is not necessary to spend time or money to help colleagues, because individuals will be unlikely to exhibit OCB-I without a reward (Kim et al., 2015). At the same time, reducing OCB-I will also help to balance the relationship with colleagues before they actually leave and eliminate the guilt associated with being in debt each other.

And those employees who have low-level OI are reluctant to cooperate with colleagues. When they plan to leave, they will further reduce OCB-I, because those who are less oriented toward future outcomes are less likely to engage in interpersonal prosocial behavior (Balliet and Ferris, 2013) and unlikely to exhibit OCB-I without a reward (Kim et al., 2015). However, those who have low-level OI but not plan to turnover know it clearly that the exchange relationships with colleagues will continue for a long time; in order to maintain a basic balance of interpersonal relationships, they perform OCB-I sometimes but will reduce in aggregate. Therefore, the hypothesis is proposed:

\[ H2b. \text{TI moderates the relationship between OI and OCB-I, such that the positive relationship is weaker when subordinates have higher TI.} \]

### 2.4 The moderating effects of LMX

On the basis of \( H1a \) and \( H1b \), the higher the degree of employees’ OI is, the greater the performance of OCB-I and OCB-O. However, this situation is not always the case in practice,
and the possibility might be that OCB motives can be quite personal and relational (Zhang and Chen, 2013). LMX indicates a subordinate’s social exchange relationship with his or her supervisor and reflects the quality of the supervisor–subordinate relationship (Koivisto et al., 2013). Therefore, LMX may affect the relationship between OI and OCB.

According to SET, either member of the dyad invests in the other and in their relationships in an effort to reap tangible and social rewards. LMX is characterized by high expectations regarding employees’ performance in return for the investments made by the leader (Breevaart et al., 2015). When receive additional resources such as autonomy, information, and the opportunity to participate in the decision-making process, subordinates rate the quality of LMX highly (Ilies et al., 2007). Because the additional resources are ultimately provided by the organization, the previous balance of the exchange relationship between employees and the organization is broken. To restore balance, they have to do extra-role behaviors to return the additional resources although in a low level of OI. And for those who are in a high level of OI, the benefits of surpassing the terms of the current labor contracts are exactly what they expected; in order to bring the social exchange relationship with the organization to a higher level, they are willing to act as a “better soldier” or a “better citizen,” and then show more OCB-O. It is assumed that subordinates’ rating of LMX quality (LMX-S) has a positive moderating effect on the relationship between OI and OCB-O. Thus, the following hypothesis is proposed:

**H3a.** LMX-S moderates the relationship between OI and OCB-O, such that the positive relationship is stronger when subordinates rate LMX highly.

Subordinates in high-quality relationships are trusted by their supervisors and given greater latitude for decision making (Townsend et al., 2002), greater empowerment (Keller and Dansereau, 1995) and greater influence in their teams (Singh and Srivastava, 2009). This causes the original balance of the relationships with the colleagues to be broken. In order to achieve a new state of equilibrium, subordinates who are in high-level OI engage in behavior that is not defined in their role descriptions, such as helping colleagues who have high workloads or who have been absent. And subordinates who are in low-level OI will also do more OCB-I moderately.

Comparatively, subordinates in lower-quality LMX often report that they are jealous of individuals in higher-quality LMX and tend to perceive the differential treatment given to employees as unfair (Lee, 2001). Such subordinates do not think that they have enough respect and consequently do not have a sufficient sense of value about their teams. Thus, their sense of obligation to help colleagues is reduced. Consequently, it is assumed that subordinates’ rating of LMX quality (LMX-S) moderates the positive relationship between OI and OCB-I. Thus, the following hypothesis is proposed:

**H3b.** LMX-S moderates the relationship between OI and OCB-I, such that the positive relationship is stronger when subordinates rate LMX highly.

LMX is a dyadic interpersonal relationship between a supervisor and a subordinate. Theoretically, LMX is an objective existence; consequently, the quality of LMX from the supervisor’s and subordinate’s perspectives should be two measurements of the same concept and should at least have a certain level of correlation. Graen and Cashman (1975) found a correlation of 0.50 between leader LMX and member LMX. Studies have mostly used subordinates’ rating of LMX quality as the single source of data (Dulac et al., 2008; Henderson et al., 2008; Rockstuhl et al., 2012). However, others have reported much lower correlations, for example, Gerstner and Day (1997) reported that the correlation between perspectives was 0.29 in a meta-analytic study, and Matta et al. (2015) reported a correlation of only 0.25 (p < 0.01). Thus, it is necessary to examine member LMX and leader LMX at the same time so as to reveal the influence of LMX on subordinates.
Supervisors change the attitudes and behavior of subordinates through the exchange relationship (Erdogan et al., 2006). When supervisors believe that the quality of LMX is high, they express their trust and appreciation to the subordinates, relay greater decision-making powers to them, and other benefits of surpassing the terms of the labor contracts. Based on the norm of reciprocity, successful LMX that develops depends on the result of a series of role-making episodes, in which leaders express their expectations and employees show the degree to which they are able and willing to live up to these expectations (Breevaart et al., 2015). In order to repay benefits provided by supervisors, subordinates accept more tasks for their organizations and show more helpings to the colleagues, which are outside their normal roles.

Low-quality LMX is based on economic exchanges, namely, exchanges based on the formal requirements of employment contracts (Liden et al., 1997). When supervisors believe that LMX quality is poor, they demonstrate their official relationships with subordinates and are stricter with their subordinates and apply higher standards (Tse et al., 2013). There are no additional requirements for the subordinates; moreover, supervisors do not require their subordinates to exhibit more OCB-O and OCB-I. In accordance with the reciprocity of exchange, the subordinates, even if they identify with their organizations, are less likely to exhibit OCB-O and OCB-I.

Thus, this study assumes that supervisors’ rating of LMX quality (LMX-L) moderates the positive relationship between OI and OCB. Consequently, the following hypotheses are proposed:

- \( H4a \): LMX-L moderates the relationship between OI and OCB-O, such that the positive relationship is stronger when supervisors rate LMX highly.
- \( H4b \): LMX-L moderates the relationship between OI and OCB-I, such that the positive relationship is stronger when supervisors rate LMX highly.

In sum, the research framework is shown in Figure 1.

3. Methodology

3.1 Sample and data collection

This study used administrative staff of the private service-oriented small- and medium-sized enterprises (SMEs) as the population in Shizhong District, Ji’nan, Shandong, China. Employees join SMEs for different purposes: some for long-term career development, some just for making a living temporarily and waiting for the opportunity to move into a more ideal organization. The management decisions of the private service-oriented SMEs have strong independence and can truly determine the remuneration based on the contribution of employees to the enterprise (Cheng and Waldenberger, 2013). In addition to the specified material benefits, employees are often given non-material benefits, such as promotion and honor. At the same time, private service-oriented SMEs are under great pressure. Especially, the employees of some key administrative positions often carry out intensive work. If remuneration or personal development opportunities are not as good as expected, they are likely to leave the company. That is to say, in private SMEs, employees’ turnover rate is high (Parkvithee, 2017), employees’ identification to the organization is uneven (Gu et al., 2010), the role of social exchange is more immediate and timely, so the relationship among OI, OCB-O, OCB-I, TI and LMX is more easily observed.

The current study chose 59 enterprises from the population. Administrative teams were then chosen from each enterprise after communicating with the human resource directors. The supervisor of each administrative team was required to choose several subordinates who had been working in the team for more than six months. Among these subordinates, half were keen to have additional jobs assigned to them by the supervisor, while the other half only undertook their jobs and were not keen on assigned extra work.
In order to ensure the authenticity of data, the supervisor and subordinates of one team were asked to come to the same room at the same time and independently complete the questionnaires. In order to protect the participants’ privacy, supervisors named each subordinate A, B, C, D, E, F, etc., respectively, and subordinates used the codes to replace their real names. The supervisor evaluated the quality of the LMX with each subordinate. Each subordinate conducted self-evaluation on OI, TI, OCB-O, OCB-I and the quality of LMX with the supervisor. After completion, each supervisor’s and his/her subordinates’ questionnaires were bound in a book and given a number. In the process of collecting questionnaires, it was must to ensure that the supervisor’s rating of LMX quality with Subordinate A corresponded to Subordinate A’s questionnaire, and the supervisor’s rating of LMX quality with Subordinate B corresponded to Subordinate B’s questionnaire, and so on. A complete paired questionnaire contained one supervisor’s questionnaire and his/her one subordinate’s questionnaire. Finally, 437 valid paired questionnaires were collected.

3.2 Measures
A set of standardized measures was used for data collection. Each question was answered on a five-point Likert-type rating scale comprising: strongly disagree (1), disagree (2), neither agree nor disagree (3), agree (4) and strongly agree (5).

3.2.1 Organizational identification (OI). A six-item scale developed by Mael and Ashforth (1992) was used to measure OI. According to Lin et al. (2017), Cronbach’s $\alpha$ is 0.91.

3.2.2 Turnover intention (TI). A three-item scale developed by Khatri et al. (2001) was used to measure TI. According to Dong et al. (2014), Cronbach’s $\alpha$ is 0.84.

3.2.3 Leader–member exchange (LMX). A seven-item scale developed by Graen and Uhl-Bien (1995) was used to measure LMX. According to Graen and Uhl-Bien (1995),
Cronbach’s α for the measure is in the 80–90 percent range. This study adjusted the wording of the supervisors’ and subordinates’ questionnaires to guide the participants to provide their evaluations of the relationships between the two parties.

3.2.4 OCB-O. An eight-item scale developed by Lee and Allen (2002) was used to measure OCB-O. According to Evans and Davis (2014), Cronbach’s α is 0.84.

3.2.5 OCB-I. An eight-item scale developed by Lee and Allen (2002) was used to measure OCB-I. According to Evans and Davis (2014), Cronbach’s α is 0.85.

3.2.6 Control variables. Considering the possible effects of other variables, the following six variables were controlled: gender, age, education, position level, working years and department scale.

In order to ensure the reliability and validity of the Chinese versions and the consistency with the originals, four times translation between English and Chinese had been conducted before measured.

4. Results and discussions
4.1 Reliability analysis
Reliability analyses were conducted by using SPSS 22.0, which indicated that Cronbach’s α coefficients for all variables/dimensions in this study were satisfactory (see Table I). The findings indicated that each variable/dimension had the right internal consistency and can be further analyzed.

4.2 Determination of variables’ independence
OCB was considered as a whole in many previous studies; however, in this study, OCB-O and OCB-I were regarded as separate variables, and further, OCB-O and OCB-I were highly related (r = 0.45, p < 0.00) (see Table I). Although they were not alarmingly high but it was necessary to carry out an alternate model fit test by combining OCB-O and OCB-I and then compare it with the measurement model where they were considered as a separate construct. This paper tested two measurement models to verify the structure of our data by using Amos 22.0. In Ma, which represented the hypothesized model, the items of the two structural variables (OCB-O and OCB-I) loaded on two distinct but correlated latent factors. In the alternative model – Mb, the items of OCB-O and OCB-I were forced into the same latent factor – OCB. Confirmatory factor analyses (CFA) were conducted to test which model’s fit indexes were better. According to the criteria of fit indexes, results indicated that Ma was better than Mb (see Table II). These findings supported that OCB-O and OCB-I were best viewed as distinct variables. This confirmed to Zhang and Chen’s (2013) conjecture, which was OCB-O and OCB-I should be clearly distinguished.

4.3 Item parceling
Reducing the number of items, in theory, increases model fit (Bandalos, 2002). According to Bandalos’ (2002) method of item parceling, the items of OI scale were parcelled from 6 into 3 new indicators; the items of OCB-O, OCB-I scales were parcelled from 8 into 4 new indicators and the items of LMX-S and LMX-L scales were parcelled from 7 into 4 new indicators, respectively. On the other hand, the items of TI scale were still 3, remained unchanged.

4.4 Validity analysis
4.4.1 Confirmatory factor analysis (CFA). CFA was conducted to examine whether the hypothesized measurement model fit the data with six factors – OI, TI, OCB-O, OCB-I, LMX-S and LMX-L. The index proved that the overall fitness of scale was good (see Table III).

4.4.2 Convergent validity. Convergent validity was assessed by composite reliability (CR) and average variance extracted (AVE). It is deemed acceptable if CR values are greater than
<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>Cronbach's α</th>
<th>CR</th>
<th>AVE</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Organizational identification (OI)</td>
<td>3.82</td>
<td>0.73</td>
<td>0.89</td>
<td>0.89</td>
<td>0.74</td>
<td>0.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Turnover intention (TI)</td>
<td>2.10</td>
<td>0.97</td>
<td>0.89</td>
<td>0.89</td>
<td>0.73</td>
<td>−0.16**</td>
<td>0.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. OCB directed toward individuals (OCB-I)</td>
<td>3.59</td>
<td>0.73</td>
<td>0.91</td>
<td>0.91</td>
<td>0.71</td>
<td>0.37**</td>
<td>−0.27**</td>
<td>0.84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. OCB directed toward organization (OCB-O)</td>
<td>3.91</td>
<td>0.61</td>
<td>0.89</td>
<td>0.90</td>
<td>0.68</td>
<td>0.46**</td>
<td>−0.30**</td>
<td>0.45**</td>
<td>0.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Subordinates' rating of LMX quality (LMX-S)</td>
<td>3.60</td>
<td>0.61</td>
<td>0.89</td>
<td>0.89</td>
<td>0.66</td>
<td>0.32**</td>
<td>−0.23**</td>
<td>0.27**</td>
<td>0.30**</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>6. Supervisors' rating of LMX quality (LMX-L)</td>
<td>3.88</td>
<td>0.60</td>
<td>0.84</td>
<td>0.82</td>
<td>0.64</td>
<td>0.08*</td>
<td>−0.04</td>
<td>0.04</td>
<td>0.03</td>
<td>0.23**</td>
<td>0.73</td>
</tr>
</tbody>
</table>

**Notes:** $n = 437$. Italic statistics on the diagonal represent the square root of AVE. *$p < 0.05$; **$p < 0.01$
0.7, AVE values are greater than 0.5 (Fornell and Larcker, 1981). As shown in Table I, CR values and AVE values were all above the recommended benchmark, indicating that convergent validity was sufficient for this study.

4.4.3 Discriminative validity. Discriminative validity was assessed by the square root of AVE (\(\sqrt{AVE}\)), which should be greater than the correlations involving the constructs (Fornell and Larcker, 1981). As shown in Table I, \(\sqrt{AVE}\) values were all above the recommended benchmark, indicating that discriminative validity was sufficient for this study.

4.4.4 Correlation analysis. Pearson’s correlation analysis was executed to examine correlations among continuous variables satisfactory, which showed that OI had a significant correlation with OCB-O \((r = 0.46, p < 0.00)\), and had a significant correlation with OCB-I \((r = 0.37, p < 0.00)\) (see Table I). Based on this, \(H1a\) and \(H1b\) were preliminarily verified.

4.5 Structural equation modeling (SEM)
This study examined the hypothesized measurement model with three factors – OI, OCB-O and OCB-I. The test result of SEM showed a good fit (see Table IV). The standardized path coefficients of the paths from OI to OCB-O \((\beta = 0.51, p < 0.001)\) and from OI to OCB-I \((\beta = 0.42, p < 0.001)\) were all significant and in the expected directions (see Figure 5). So, \(H1a\) and \(H1b\) were, respectively, supported.

4.6 Hierarchical regression analysis
We considered that the relationship between OI and OCB could be an artifact of a cross-sectional design and not a true causality relationship (Bang et al., 2016). In order to guard against this potential criticism, hierarchical regression analyses were carried out to explore the proposed moderating effects of TI, LMX-S and LMX-L.

### Table II.
Multiple fit indexes of models

<table>
<thead>
<tr>
<th>Models</th>
<th>(\chi^2)</th>
<th>(\chi^2/df)</th>
<th>GFI</th>
<th>AGFI</th>
<th>NFI</th>
<th>CFI</th>
<th>RMSEA</th>
<th>SRMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ma</td>
<td>387.10</td>
<td>3.76</td>
<td>0.89</td>
<td>0.85</td>
<td>0.90</td>
<td>0.93</td>
<td>0.08</td>
<td>0.05</td>
</tr>
<tr>
<td>Mb</td>
<td>1,397.48</td>
<td>13.44</td>
<td>0.57</td>
<td>0.43</td>
<td>0.64</td>
<td>0.66</td>
<td>0.17</td>
<td>0.15</td>
</tr>
<tr>
<td>Criteria</td>
<td>The less the better</td>
<td>(\leq 5.0)</td>
<td>(\geq 0.80)</td>
<td>(\geq 0.80)</td>
<td>(\geq 0.90)</td>
<td>(\geq 0.90)</td>
<td>(\leq 0.08)</td>
<td>(\leq 0.08)</td>
</tr>
</tbody>
</table>

**References:** Bentler and Bonett (1980), Schumacker and Lomax (2004), Doll et al. (1994), MacCallum and Hong (1997), Bentler and Bonett (1980), Hu and Bentler (1999), McDonald and Ho (2002), Hu and Bentler (1999)

**Source:** This study’s research

### Table III.
Fit indexes of CFA

<table>
<thead>
<tr>
<th>Fit index</th>
<th>(\chi^2)</th>
<th>(\chi^2/df)</th>
<th>GFI</th>
<th>AGFI</th>
<th>NFI</th>
<th>CFI</th>
<th>RMSEA</th>
<th>SRMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>484.97</td>
<td>2.29</td>
<td>0.91</td>
<td>0.88</td>
<td>0.93</td>
<td>0.96</td>
<td>0.06</td>
<td>0.04</td>
</tr>
</tbody>
</table>

**Source:** This study’s research

### Table IV.
Fit indexes of SEM

<table>
<thead>
<tr>
<th>Fit index</th>
<th>(\chi^2)</th>
<th>(\chi^2/df)</th>
<th>GFI</th>
<th>AGFI</th>
<th>NFI</th>
<th>CFI</th>
<th>RMSEA</th>
<th>SRMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>155.51</td>
<td>3.79</td>
<td>0.94</td>
<td>0.90</td>
<td>0.95</td>
<td>0.97</td>
<td>0.08</td>
<td>0.04</td>
</tr>
</tbody>
</table>

**Source:** This study’s research
In the regression, the standardized scores of the variables were entered. As an example, in order to assess whether the relationship between OI and OCB-O was moderated by TI, the interaction term (OI × TI) was built. In Step 1, gender, age, education, position level, working years and department scale were entered. In Step 2, the main effects of OI and TI were added. In Step 3, the interaction term (OI × TI) was put in. The process was repeated to conduct all analyses by using SPSS 22.0. The results are reported in Tables V–X. In order to interpret significant interactions, simple slope analyses and the two-way standardized worksheet by Dawson (2014) were performed (see Figures 2–4).

As shown in Table V, TI was not found to moderate the relationship between OI and OCB-O ($\beta = -0.04, p = 0.36$). Thus, $H2a$ was not supported.

This is perhaps related to the subordinate’s assessment of the exchange relationship with the organization before and after departure. Such considerations include an assessment of their current career (Hausknecht and Trevor, 2011). Among the employees with high TI, some identify with the organization and some do not identify with the organization.

<table>
<thead>
<tr>
<th>Variables</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent variable: OCB-O</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>0.06</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>Age</td>
<td>0.02</td>
<td>-0.02</td>
<td>-0.03</td>
</tr>
<tr>
<td>Education</td>
<td>-0.05</td>
<td>-0.02</td>
<td>-0.02</td>
</tr>
<tr>
<td>Position level</td>
<td>0.22***</td>
<td>0.11*</td>
<td>0.10*</td>
</tr>
<tr>
<td>Working years</td>
<td>-0.09</td>
<td>-0.07</td>
<td>-0.06</td>
</tr>
<tr>
<td>Department scale</td>
<td>0.02</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>Organizational identification (OI)</td>
<td>0.41***</td>
<td>0.41***</td>
<td>0.41***</td>
</tr>
<tr>
<td>Turnover intention (TI)</td>
<td>-0.24***</td>
<td>-0.24***</td>
<td>-0.24***</td>
</tr>
<tr>
<td>OI × TI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.04</td>
<td>0.27</td>
<td>0.27</td>
</tr>
<tr>
<td>$\Delta R^2$</td>
<td>0.05</td>
<td>0.24</td>
<td>0.00</td>
</tr>
<tr>
<td>$F$</td>
<td>3.62***</td>
<td>70.62***</td>
<td>0.85</td>
</tr>
</tbody>
</table>

Notes: $n = 437$. Values in the table are standardized regression coefficients. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$

Source: This study’s research

<table>
<thead>
<tr>
<th>Variables</th>
<th>M4</th>
<th>M5</th>
<th>M6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent variable: OCB-I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>0.04</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>Age</td>
<td>0.07</td>
<td>0.05</td>
<td>0.03</td>
</tr>
<tr>
<td>Education</td>
<td>-0.01</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>Position level</td>
<td>0.25***</td>
<td>0.16***</td>
<td>0.15**</td>
</tr>
<tr>
<td>Working years</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Department scale</td>
<td>0.10*</td>
<td>0.15***</td>
<td>0.15***</td>
</tr>
<tr>
<td>Organizational identification (OI)</td>
<td>0.28***</td>
<td>0.29***</td>
<td>0.29***</td>
</tr>
<tr>
<td>Turnover intention (TI)</td>
<td>-0.24***</td>
<td>-0.25***</td>
<td>-0.25***</td>
</tr>
<tr>
<td>OI × TI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.08</td>
<td>0.22</td>
<td>0.24</td>
</tr>
<tr>
<td>$\Delta R^2$</td>
<td>0.09</td>
<td>0.15</td>
<td>0.02</td>
</tr>
<tr>
<td>$F$</td>
<td>6.94***</td>
<td>42.15***</td>
<td>8.53***</td>
</tr>
</tbody>
</table>

Notes: $n = 437$. Values in the table are standardized regression coefficients. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$

Source: This study’s research

Table V. The moderating effect of TI on the relationship between OI and OCB-O

Table VI. The moderating effect of TI on the relationship between OI and OCB-I.
Those who identify with their organizations are willing to repay the treatment, welfare and growth opportunities they have received. They may also hope to continue receiving such support from their former organizations in their new careers. Thus, such employees are willing to maintain good exchange relationships with their former organizations (Tavares et al., 2016); in addition to doing a good job in the role before actually leaving the organization, they will continue to exhibit extra-role behavior that is beneficial to the organization. That is to say, employees who identify with the organization will continue to perform OCB-O even if they have TI. On the contrary, those who do not identify with the organization do not want to exchange with the organization whether or not they have TI (Kumar and Singh, 2012).

As shown in Table VI and Figure 2, TI was found to moderate the relationship between OI and OCB-I ($\beta = -0.13; p < 0.01$). Thus, $H2b$ was supported.

As shown in Table VII and Figure 3, LMX-S was found to moderate the relationship between OI and OCB-O ($\beta = 0.12; p < 0.01$). Thus, $H3a$ was supported.
As shown in Table VIII and Figure 4, LMX-S was found to moderate the relationship between OI and OCB-I ($\beta = 0.31$, $p < 0.001$). Thus, $H3b$ was supported.

As shown in Table IX, LMX-L was not found to moderate the relationship between OI and OCB-O ($\beta = 0.05$, $p = 0.23$). Thus, $H4a$ was not supported.

As shown in Table X, LMX-L was not found to moderate the relationship between OI and OCB-I ($\beta = 0.04$, $p = 0.40$). Thus, $H4b$ was not supported.

From the results, subordinates' rating of LMX quality (LMX-S) and supervisors' rating of LMX quality (LMX-L) had different influences on subordinates. LMX-S had a significant positively moderating effect on the relationship between subordinates' OI and OCB-O/OCB-I. However, LMX-L had no significant moderating effect on them. These show that subordinates' attitude and behaviors are mainly affected by their own cognition and internal evaluation. In other words, subordinates' cognitive systems are the main factor for their attitudes and behaviors, and supervisors' cognition cannot affect directly subordinates. The results are consistent with the cognitive behavioral theory, which is one's behaviors are not
directly determined by events, but rather by the way these events are cognitively processed and evaluated (Ellis, 1991).

Further, one item of note from Table I was that the correlation between LMX-L and LMX-S was only 0.23 ($p < 0.01$), namely, the perception of LMX between supervisors and subordinates was unequal. Theoretically, LMX is an objective existence; consequently, the quality of LMX from the supervisor’s and subordinate’s perspectives should be two measurements of the same
concept and should at least have a certain level of correlation. Based on this idea, studies have mostly used subordinates’ perceptions of LMX as the single source of data (Buch et al., 2016; Coyle and Foti, 2015; Dulac et al., 2008). By focusing on only one perspective, these studies have implicitly assumed that LMX quality assessed from the perspective of one member of a dyad is sufficient to describe the nature of that relationship (Matta et al., 2015). Yet, supervisors’ and subordinates’ views of the relationships often do not converge (Sin et al., 2009); a few studies reported a low leader–member agreement, for example, Gerstner and Day (1997) reported the correlation between perspectives was 0.29, and Matta et al. (2015) reported that correlation was only 0.25 ($p < 0.01$). In dyadic relationships, there are two cognitive subjects, each with his or her own perception. The possibility of cognitive differences between supervisors and their subordinates regarding LMX quality might be that one’s perception may be affected by several factors such as personality, emotional displays (Medler-Liraz and Seger-Guttmann, 2018), political leadership skills (Buch et al., 2016) and leadership style (Michel and Tews, 2016; Saboe et al., 2015). Another possibility might be that the relationships occur in an organizational context, and environmental variables such as the characteristics of organizational culture and organizational politics (Johnson et al., 2017) may have different effects on supervisors’ and subordinates’ perceptions of LMX quality. Therefore, a paired-samples design was confirmed to be necessary in this study, as Krasikova and LeBreton (2012) pointed out that studying dyadic phenomena (such as LMX) from the perspective of one dyad member is “theoretically deficient.”

Based on the above empirical analyses, the test results of all the hypotheses are shown in Figure 5.

Figure 5.
Standardized path coefficients and moderating data

Notes: S, responded from subordinates; L, responded from supervisors. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$
5. Conclusions

5.1 Main findings
This study found that OI indicated a positive influence on both OCB-O and OCB-I. However, the relationships between OI and OCB-O/OCB-I are not a true causality relationship (Bang et al., 2016), and affected by TI and LMX. TI showed a negative moderating effect on the positive relationship between OI and OCB-I, but had no significant moderating effect on the positive relationship between OI and OCB-O. Supervisors' and subordinates' ratings of LMX quality were not in one same direction and shred different moderating levels. Supervisors' rating of LMX quality had no significant effects on the relationship between OI and OCB-O, nor did significant effects on the relationship between OI and OCB-I. On the other hand, subordinates' rate of LMX quality strengthened both the positive relationship between OI and OCB-O, and the positive relationship between OI and OCB-I.

5.2 Theoretical contributions

(1) The theoretical basis was extended. Previous research on OI was mostly based on social identity theory. And this study explored the relationship among the variables from social exchange perspective and got some new findings.

(2) The research content of OCB is complemented appropriately. Major scholars study OCB as a whole. However, this study found that OCB-O and OCB-I had different forming mechanisms and should be explored as distinct variables.

(3) In terms of research methods, this study confirmed that it was necessary to use a paired-samples design and collect data from both levels when studying dyadic phenomena (such as LMX). Based on this, differences between the effects of supervisors' and subordinates' rating of LMX quality were found during the exchange process in the organization.

5.3 Practical implications
From the perspective of social exchange, the current study provides some important implications to organizational practices:

(1) SMEs should strive to improve subordinates' OI so as to motivate them to perform more extra-role behaviors toward the organization and fellow employees.

(2) Although the subordinates' turnover rate is high in private service-oriented SMEs, supervisors should not be afraid of the rational turnover of subordinates and should improve the quality of the exchange relationship between organizations and employees to motivate subordinates with TI to show more extra-role behaviors to the organization before they actually quit.

(3) Subordinates' rating of LMX rather than supervisors' personal perspective deserved more considerations, because supervisors and subordinates often have different views on the same thing. Supervisors consider that the relationships with subordinates are good, while the subordinates may not consider so. And subordinates' attitude and behaviors are mainly affected by their own cognition, and supervisors' cognition cannot affect directly on subordinates. Only when subordinates consider that the quality of the relationships with their supervisor is high, they will show more extra-role behaviors to help the organization and colleagues regardless of the level of OI.
5.4 Limitations and future research

(1) This paper did not find the direct effect of supervisors’ perceptions of LMX on subordinates; the possibility might be that its effects may be curvilinear, and another possibility might be that there are other moderating variables influencing the degree of LMX agreement (Sin et al., 2009). How supervisors’ cognition affects subordinates? These issues are not discussed in the current study and should be explored in future studies.

(2) The current study tried to ensure the homogeneity of enterprises and the representativeness of samples. Because of time limitations, the current study did not compare and analyze the influence of corporate backgrounds and culture. Further research could address this omission.

(3) The measures translation in the current study is relatively simple. Because the scales in English are used in Chinese samples, in next studies, the items should be translated from English into Chinese by a first translator and then independently back-translated into English by a second translator following the translation-back-translation method (Brislin, 1970). Discrepancies between the original and the back-translated versions must be discussed between the translators and some final adjustments could be made.

References


Further reading

About the author
Yan Shen is Doctoral Candidate in Business Administration at the China-ASEAN International College of Dhurakij Pundit University, Bangkok, Thailand. She is Professor of Organizational Psychology at Shandong Yingcai University, China. Her research interests include organizational behavior, leadership and human resource management. Yan Shen can be contacted at: sheny628@126.com

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Abstract

Purpose – The importance of foresight is discussed in relation to why traditional scenario planning methodology is problematic at achieving it. The "survivor syndrome" is borrowed from the human resources literature and presented as a metaphor for foresight to illustrate how better "scenarios" can be achieved by understanding the syndrome better. A practice perspective is given on the use of a seven-theme framework as a method of interviewing survivors. The paper aims to discuss this issue.

Design/methodology/approach – The paper draws from an empirical research that took place during the 2008 global financial crisis to illustrate the richness of the insights that would otherwise not be obtainable through scenario planning methods that do not involve "survivors." In that research, semi-structured interviews were employed with key personnel at multiple levels of one private and one public organization that had undergone a redundancy process at the time of the crisis to explore its effect on the remaining workforce.

Findings – The "survivor syndrome" itself would be minimized if managers consider the feelings of survivors with more open communication. Survivors in private firms were found generally to experience anxiety, but are more likely to remain more motivated, than their counterparts in the public sector. These detailed insights create more accurate "scenarios" in scenario planning exercises.

Originality/value – Organizational performance can be better enhanced if the survivor syndrome can be better managed. In turn, scenario planning, as a form of organizational foresight, is better practiced through managing the survivor syndrome. Scenario planning methodology has proliferated well in the human resource management literature.

Keywords Performance management, Scenario planning

Paper type Research paper

Introduction

Foresight has become one of the latest management buzzwords, rising above many other tools of strategic competitive advantage, particularly in times of uncertainty and environmental turbulence. Principally, the ability to incorporate into the present decisions of organizations (organizational foresight) or specifically into the strategic decisions (strategic foresight), the expectations of future conditions, foresight is the capacity to think systematically and develop as individuals and as an organization to prepare for those future eventualities. More elegantly put, “foresight is a unique and highly-valued human capacity that is widely recognized as a major source of competitive advantage and cultural renewal within nations and corporations” (Chia, 2002, p. 5). However, how exactly foresight can be practiced remains at the forefront of researchers’ agenda, and has offered a variety of answers, not the least one involving the use of scenario planning methodology.

The much documented global financial crisis of 2008 has provoked an outburst of research to understand it better, such as the causes and lessons learned so that firms may revert smoothly back to their routine activities (see Chau et al., 2012). While luck may have had a significant role to play (Parnell et al., 2012) in volatile environments, at times when luck is not
on a firm’s side, it is resorted to utilizing foresight as a valuable vehicle with which to forecast, scan and scenario plan for the future. In this paper, the importance of scenario planning (as a methodology of foresight) is presented as a metaphor for how employees felt after coming close to redundancy during a major restructure of the organization. Understanding and stabilizing these feelings are useful in engaging in foresight/scenario planning exercises as these staff are responsible for restoring the organization. The term “survivor syndrome” is used in this paper to refer to this set of feelings, following Baruch and Hind (2000), specifically in the context of downsizing or redundancy, but can concern any other reason for reducing a workforce, such as downsizing, implementation of artificial intelligence systems or simply improving management processes, which itself has been a growing phenomenon and is seen as a part of work life (Datta et al., 2010). As a mental condition originally understood from the discipline of psychology that stems off from post-traumatic stress disorder, it relates to any “tragedy” felt by the sufferer and the overall guilt about surviving, what should have been done and what the survivor actually did (Hendriksen, 2018). A major reason for the failure of many firms is not just the original causes of the downsizing but also their inability to manage the survivor syndrome afterwards, as survivors are unlikely to behave in the same way after the downsizing, despite the small majority who suffer “learned helplessness” – a condition that explains why sufferers (such as victims of domestic abuse) remain in the same situation, believing little can change and accepting the continued suffering (Appelbaum et al., 1997).

van Dick et al. (2016) found that identification (the ability of the survivor to identify themselves within the organization immediately after the downsizing) is a mediator of individual performance, so it is imperative to understand the self-categorization process of individuals that can either plunge the organization into further difficulty or pull it out of existing trouble, which in turn will assist an economy to recover.

Hence, the reason for using foresight as a metaphor to understand the survivor syndrome better – to “problematize” the body of knowledge, to borrow a term from the management research literature (Alvesson and Sandberg, 2011) – is because these “survivors” are those an economy must rely on to pull it out of an existing recession, and so such insights have valuable, practical implications and offer transferrable lessons learned for managers for future recessions of a variety of capitalist types (for a review of such varieties of capitalism, see Witcher and Chau, 2012). Sahdev (2004) argues, because the survivor syndrome can be likened to a breach of the psychological contract and a violation of organizational justice, the reconciliation process (and reorientation of survivors after the redundancy) would involve top leaders (including politicians) in mediating between internal and external institutional forces, thus linking the behaviors of the survivors directly to macro-economic conditions, such as economic growth.

The referent of a major economic recession is also supportive of Taleb’s (2007) argument to establish “convexity” in achieving a positive outcome (Derbyshire and Wright, 2014), when the 2008 global financial crisis was possibly a “black swan event” due to its high unpredictability and high impact, making the practice of scenario planning particularly difficult. Foresight is therefore facilitated by the micro-practice perspective of the firm (e.g. Sarpong et al., 2013) to smooth out the harder strategic options during a recession. The attention of this paper is therefore on understanding the detailed management issues, such as individual staff motivation levels from the broad management literatures (e.g. Brockner et al., 1986; de Vries and Balazs, 1997; Paulsen et al., 2005; Bean and Hamilton, 2006), to advise how to manage better in major economic recessions.

To do so, the present paper draws from a firsthand research project to offer insights obtained through the use of conducting interviews with survivors of two large organizations (of around 500 employees in each) that underwent a major redundancy program. The purpose of the interviews was to obtain deeper insights on the themes relating to the survivor syndrome, so that it could be better managed by the senior managers to re-orientate them back into the routine work of the organization in the post-redundancy phase. The organizations were
selected based on them being a major employer in the region, which meant that jobs were scarce and thus represented an extreme survivor syndrome that the interviewees faced. The fact that the organizations represented both private and public sectors meant the insights obtained would represent good generalizability; while there were some differences between the two organizations that required noting, the purpose of the research was not to conduct a direct comparison between them. Non-probability sampling within the organizations for identifying the interviewees was used because the topic was perceived sensitive to many survivors after the downsizing so it was best not to have pre-discussed matters with them, and most of the survivors were overloaded with additional work as the structure had just been reduced so it was too much to pressurize an already busy workforce for additional volunteers. Over 30 people volunteered to take part in the study, although the final sample consisted of 18 individuals because the data had already become saturated and sufficient insights had been obtained. The sample included 6 females (12 males), 9 from the private firm (9 from the public firm) and 6 in managerial positions (12 at operational levels). The arithmetic mean average age was about 50 years. The interviews took place between six and eight months after the redundancy rounds of both companies; this enabled the interviewees to have calmed down, carried on and thought forward with respect to their organizational performance and their own careers, enabling a more accurate reflection of their experiences.

In this illustrative example, semi-structured interviews were used: no specific set of questions was asked to the interviewees, as the nature of the research was inductive with the intention of obtaining as many new insights as possible, although the questions fell within the topic of the survivor syndrome. The conversations were mainly one-sided, allowing the interviewees to discuss their experiences and only a few prompts were made by the interviewer to ensure the conversations did not go off-topic. Each interview lasted around an hour in length, was digitally recorded and transcribed verbatim afterwards. The interview transcriptions were analyzed manually as the researchers knew the contexts extremely well, using traditional theme tree and content analysis (thematic analysis) techniques to group and present the conversations within their popular researched themes (e.g. Miles and Huberman, 1994). A reflexive account of the researchers’ observations and personal judgments was kept to control for bias, in line with recommendations made a decade ago in the present journal (see Chau and Witcher, 2009).

This paper contributes empirically in two ways: first, rich and in-depth insights from survivors are sought that have the benefit of understanding social practice that connects the future and past (as argued for by Sarpong and Maclean, 2014), which complement extant research findings that are too specific and isolated; and second, these views pivot around how defining characteristics of private and public firms shape the emotions and motivation of employees during economically critical times. In so doing, it also contributes theoretically by arguing the appropriateness of interviewing on the themes of the survivor syndrome as an augmented scenario planning methodology for the specific situation of managing in conditions of severe economic crises (but does not necessarily replace other well-established approaches of general planning already in use for other general conditions); this is purported as an improved mechanism as it overcomes the doubts over the suitability of research participants within the technique and strikes an appropriate balance between plausibility and probability, as the participants are exclusively all those at stake and have come close to the most plausible and highly probable conditions concerned.

It is now a decade since the highly impactful global financial crisis of 2008. This paper contributes to the special issue by offering a practice reflection on the use of the once considered “breakthrough” methodology of scenario planning. Thus, it argues that interviewing survivors on the themes identified in this paper relating to the survivor syndrome is an improved method of conducting scenario planning because of the near-redundancy experiences that would have led them to a different behavior within the organization, that could not otherwise be obtained.
through laboratory-based methods of strategizing. The harmony between this interviewing technique and the aim of scenario planning suggests a natural proliferation of the use of scenario planning for the broad purposes of managing organizational performance.

What are foresight and scenario planning, and why they matter?
They matter, at least for foresight, as in the words of Chia (2002: p. 5) “[…] the ability to read, interpret, foresee and redefine emergent global socio-political and economic trends […] are all important assets that no forward-looking nation or organization can afford to ignore.” While it is generally agreed that foresight is a capacity of individuals and firms with which to develop a competitive advantage, the exact way this is achieved has been heavily disputed. In this section, these key propositions are reviewed and discussed in the light of our argument for how it matters for firms considering downsizing strategies as a strategic option during an economic crisis, as a way to remain forward-looking and competitive. Foresight is a metaphor in this sense because of its importance in considering future eventualities, but is only possible in knowing the counterfactual of a strategic downsizing decision, which can only be obtained from those who were placed in a vulnerable position of that decision and “survived” it and can subsequently tell their stories and comment on that apprehension.

Foresight though is understood under a number of guises. Neugarten (2006) reviews the usable value of foresight as competitive intelligence by likening its limitations to those of a biological eye; looking too directly into the future is problematic (e.g. tunnel vision ignores the surroundings and blind-spots), and looking forward should not neglect the importance of looking sideways. For the organization, additional vision should involve those players who are not ordinarily deemed core to decision making. For example, Chau and Quire (2018) identify the most common — women — and their particular value in foresight exercises in the technology sector. Research by Sarpong and Maclean (2014) and Sarpong et al. (2013) also emphasizes the importance of human participants — that is, the need to examine ordinary organizational members — in order for human capacity to connect with the past, present and future in a social practice. In this instance, survivors of redundancy rounds are those key human participants of the organizational process whose “feelings” are core to the future outcome of a present decision, and can be understood as an opportunity to preview a particular strategic consequence, thereby linking together all states of past (the economic condition that resulted in the strategic consideration), present (strategic choice of downsizing) and future (the resultant labor force). It is therefore necessary to examine the themes of the downsizing/redundancy literature (e.g. Pina e Cunha et al., 2006) in order to understand the minutiae of an organization’s working, to enable the conduct of effective foresight exercises.

Scenario planning is probably the broadest form of foresight in that it does not predict the specifics of the future per se, but attempts to understand the critical uncertainties that organizations face in their strategic context and to improve the quality of strategic decisions (Meissner and Wulf, 2013). Argued in Schoemaker’s (1995) seminal article as a valuable tool for strategic thinking and based on first use of the methodology over half a century ago at Royal Dutch/Shell Company in the 1960s (see Wack, 1985), scenario planning constructs scenarios to overcome highly uncertain situations for managers to predict; over the long term, the usual problems of overconfidence and tunnel vision can be compensated for. Its application has been varied and continues to be of considerable research interest in new and different contexts — for example, for energy and online platforms in Alizadeh et al.’s (2016) and Raford’s (2015) respective recent reviews.

At the core of scenario planning lies the main question, “if then?” to set out such scenarios, and numerous attempts have been made to refine the most appropriate technique for carrying out this task (see Wulf et al., 2010). These have typically involved the regular two-by-two matrixes as well as the use of repeated strategic workshops to go through intensive steps of procedure and discussion (see Franco et al., 2013). Cairns et al. (2016) argue
the use of such workshops helps make the unfamiliar more familiar, but also suggest the inclusion of scenario refinement and improvisation, such that “the improved scenario(s) must be designed to make the familiar unfamiliar [emphasis added], to provoke challenge, but aim to be credible and relevant” (p. 101). Perhaps one of the key problems is that involving the need to understand better the significance of historical events (Bradfield et al., 2016) that average out in future decisions.

Termed the “scenario planning paradox,” it is argued there is insufficient theory to support scenario planning methodology, therefore rendering it chaotic, so empiricism is crucial to make it useful (Spaniol and Rowland, 2018). In other words, more data are needed to support procedure. Specific steps for conducting scenario planning, like those of Burt et al. (2006) or Konno et al. (2014), offer valuable “how-to” guidance on the methodology, and are heavily premised on environmental analyses. However, following Taleb (2007) on recognizing “black swan” events that subvert from industry trends for which organizations struggle to plan due to unpredictable conditions, there is the need to build on more organizationally resilient approaches that establish convexity to redistribute events for a more positive outcome (Derbyshire and Wright, 2014). “Surprise” (borrowing from “potential surprise theory”) suggests there is some theoretical grounding surrounding the uncertainty around decision making during turbulent environments (Derbyshire, 2017). Much questioned in the extant literature is the extent to which the 2008 global financial crisis constitutes a recent black swan event (Witcher and Chau, 2014) – i.e., one which came as a surprise. As history has indicated the likelihood of its occurrence, some might argue that some black swans turn white, and some events are in fact “gray swans” (as events average out over time).

Averaging out events realistically, or striking the balance between those conditions that are plausible and probable (Ramirez and Selin, 2014), is difficult when environmental conditions are turbulent and highly unstable, so some form of go-between to get closer to the most appropriate scenario is required. The need to scenario plan is heightened when conditions are uncertain and when competitive environments are dynamic (Oliver and Parrett, 2018). It is found that the wider the pool of participants involved in generating the scenarios, the more realistic the scenario might be derived (Zapata and Kaza, 2015), but reliance is placed significantly on additional scenario developers who might already be few in existence. Other common “pitfalls” or established problems associated with scenario planning in general include bias (availability and reporting, group-think, culture), relevance (appropriateness and availability of participants) and longevity and resilience (short-termism involved in some work and availability of tools and techniques) (KPMG, 2011). Scenario planning can also “muddy communications” if presentation of too many scenarios complicates current decision making and lengthens the tails of the distribution (Roxburgh, 2009).

Nonetheless, we argue the use of a methodology that captures a specific event of volatility is core in improving scenario planning, and can address the above problems, such as by relating to the survivor syndrome. This view is consistent with that of storytelling and the use of an inductive practice approach is a good conceptual lens for successfully analyzing scenario planning and process data (Bowman et al., 2013). The assumption of a linear relationship between scenario planning and strategy development is wrongly conceived by many pre-existing models as human beings are a rather complex species (Rowland and Spaniol, 2017). Some form of extensive scenario orientation (O’Brien and Meadows, 2013), involving the extensive use of appropriate informants, is the bridge to effective strategy creation. The need to draw directly from the intricacies of human beings was recognized as building new social capital in order to access new information, novel strategic options and collaborative opportunities (Lang and Ramirez, 2017). The common pitfalls mentioned earlier are thus minimized: for example, the problem of bias in opinions can be overcome by the personal jobs at stake of the participants and relevance to context would be high as the participants of the exercise are those who have come close with the
situation the organization is scenario planning for. Reducing “muddy communications” and curtailing the distribution of possible scenarios to become more probable/plausible ones will be more achievable as the participants are specifically selected and scenarios are all clustered around the employment conditions in question.

Hence, our suggestion is to use a seven-theme framework relating to the survivor syndrome which offers up closer insights (or “scenarios”) for understanding situations relating to the better management of the syndrome for the benefit of the survivors. These insights might otherwise not have been achievable through the use of other traditional scenario planning techniques, so this approach helps close the gap (Chermack, 2005) between the theory of scenario planning methodology and practice. The themes framework, explained in the next section, is also consistent with thinking on the meaning of theory within foresight (see Piirainen and Gonzalez, 2015) – that knowledge is created (the seven themes), a process of usage (incorporation into interviewing) and which helps to theorize about the future (by raising specific scenarios).

**Themes relating to the survivor syndrome**

The following themes were identified from an extensive review of the extant literature relating to feelings of employees who survived a redundancy round at their place of work, although “survivor syndrome” was not the only phrase used in the paper search (other keywords included “feelings,” “aftermath,” “redundancy,” etc., among others). Other related issues that emerged from the literature review were broadly fitted into the themes given below to minimize the number in total, so that the themes could be used collectively as a simple framework with which to ask questions during interviews with other survivors as part of an organization’s strategic foresight (particularly, scenario planning) exercise.

**Fairness**

Brockner et al. (1987) carried out a field and laboratory study utilizing the justice theory which illustrated that survivors’ work behaviors and attitudes are influenced by the selection process undertaken by the organization to lay off their co-workers. The survivors’ reactions to the victims of perceived injustice resulted in survivors distancing themselves from the organization and displaying dysfunctional behavior and attitudes. A selection process that is perceived by survivors to be unjust can therefore have negative consequences and reduce commitment and individual work performance. Travaglione and Cross (2006) however found the converse situation in which any injustice survivors witnessed was overpowered by the feeling they had survived the redundancy process and were treated well in comparison with their colleagues. The manner in which downsizing is executed and the perceived fairness of the downsizing process also influence the behavioral and attitudinal consequences of survivors (Brockner, 1992). A series of studies (see work by Brockner et al., 1987, 1990, 1993) examined the effect of perceived fairness in a controlled environment. This involved survivors witnessing that the organization had dealt with the redundancy process fairly using open communication, an explanation of the redundancy along with a fair selection process and consultation. They felt this would leave survivors feeling committed to the organization, thus increasing their motivation, and ultimately overall company performance. Furthermore, Davy et al. (1991) recommended that requesting input from employees into any organizational decision process further enhances fairness, while Trevor and Nyberg (2008) discussed the impact procedural justice has on the perceived fairness of survivors. Therefore, influencing survivors’ perceived fairness prior to downsizing has been shown to manipulate survivors’ reactions (Travaglione and Cross, 2006) significantly. Similarly, if a survivor has witnessed an unfair redundancy process she/he tends to have a broken psychological contract (Legge, 1995; Sahdev, 2004) which can lead to disruptive behavior, with fear displacing loyalty and trust.
Fineman (2003) went further to describe that watching colleagues during a redundancy has been depicted as similar to that of experiencing grief.

**Organizational assistance**

The survivor syndrome is more likely to surface if the survivors perceive that organizational support for those who were laid off is low (Travaglione and Cross, 2006). Survivors must not feel that they have been neglected and that assistance had been provided throughout, during and after the process of organizational transformation during the recession. Organizational assistance is also linked to the perceived fairness of survivors. If an organization is seen to assist survivors and victims, then perceived fairness is raised, although the opposite effect occurs if assistance is not provided.

**Feelings**

Managers need to realize that the way in which victims of layoffs are treated will be observed by those who remain (Travaglione and Cross, 2006) and they should therefore consider survivors’ feelings throughout the redundancy process. Most research examines the survivor’s affective reactions to redundancy such as anxiety, stress or emotional well-being (Appelbaum and Donia, 2001a), whereas redundancy can also lead to more specific emotional states, such as guilt or shame (Fineman, 2003). Survivor guilt is defined as including symptoms such as depression, fear and anger, along with feelings of envy toward those who have left the organization (Sahdev, 2003). Brockner et al. (1986) tested a range of emotions that survivors displayed and questioned whether subjects that felt guilt worked harder to eliminate this feeling and the impact this had on the individual’s motivation and attitude. However, as this test was undertaken within a laboratory environment, there was no real threat of job loss to the individual, so it lacked credibility. It could also be plausible to argue that survivors should not feel guilt, as they are typically not responsible for making the decision to dismiss their colleagues. Thinking, known as Adams’ (1965) Equity Theory, demonstrates how redundancy can create an assortment of psychological states in survivors such as job insecurity, positive inequity, anger and relief. These emotional states will likely have an impact on the survivor’s motivation, job satisfaction and commitment (Brockner et al., 1987).

**Motivation**

Chipunza and Berry (2010) examined the relationship between a survivor’s attitude, commitment and motivation, and recommended that an organization’s human resources process should involve survivors within the consultation process, in order to improve motivation and commitment. This research was carried out in Zimbabwe, and may have been subjected to specific cultural and austere economic constraints. Nonetheless, similarly, Weiner’s Equity Theory suggests that motivation is influenced by procedural justice (Trevor and Nyberg, 2008), with job insecurity playing an important role in determining the motivation of survivors. Brockner (1992) shows that motivation is negatively influenced by job insecurity. In addition, motivation may be further affected due to the change in survivor’s roles as they adopt increased workloads along with a fear of further layoffs can lead to a reduction in motivation and an increase in anxiety level (Appelbaum et al., 2003).

**Communication**

Communication is often viewed as one of the most important aspects of the downsizing process (Brockner, 1992; Marks, 2006). It has also been identified as one of the reasons why the redundancy process fails, as a lack of information is likely to cause more damage. Survivors will desperately seek as much information as possible (Brockner et al., 1990) to allay their fears about the future (Paulsen et al., 2005). It is therefore important for employees to feel that they
are being listened to and that their feelings, views and suggestions are taken into account, which will also increase the level of employees’ trust in the management (Tzafrir et al., 2004). Survivors need to know as soon as possible that they will not be losing their jobs and they will further benefit if information concerning the reason for the redundancy is relayed to them effectively and efficiently. While it is vital that positive information be relayed throughout the process, it is also important that bad news is openly communicated (Noer, 1993). Furthermore, survivors will want to receive news about victims and how they have been treated along with information relating to any assistance offered to the victims as this will provide the survivors with an indication of how they may be treated in the future (Thornhill and Gibbons, 1995). The effective management along with extensive communication (Marks, 2006) and the development of a “best practice” HR strategy (Noer, 1993), and not just organizational learning (Bui et al., 2016), can assist in diminishing the syndrome.

Workload
Survivors may be expected to take on increased workload in the case of organizational restructuring, but not many organizations have plans in place for them to adjust to this new workload (Appelbaum et al., 2003), thereby resulting in survivors feeling a loss of control and uncertainty, which can cause stress (Chipunza and Berry, 2010). As workloads increase due to the re-distribution of the remaining work among the survivors, the survivors are left facing job insecurity. Work boundaries are often not clear (e.g. Chau et al., 2017). However, Brockner et al. (1993) find that sometimes survivors who had perceived an increase in their workload may feel positive and have enriched their skills.

The future
It is argued that a broader approach for downsizing can have a positive long-term impact in the longer term (Appelbaum and Donia, 2001b). This should comprise a complete strategic transformation and become part of a continuous improvement plan to improve productivity, cut costs and increase turnover (de Vries and Balazs, 1997). The future will also see relationships between individuals and organizations changing with employees becoming more focused on career self-management (Appelbaum and Donia, 2001b). Cameron (1994) proposes a prescription for best organizational practices which includes an approach for the use of long-term strategies, good preparation and employee involvement, for example.

These seven themes, in no particular order of preference or importance, are therefore recommended for use as a broad list of categories to cover when interviewing survivors when conducting scenario planning exercises. Any insights obtained within these themes form the new “scenarios” that advise how to manage redundancies in the future, as well as offer up suggestions for rebuilding the organization.

Survivor interview framework in use
The following practice insights were obtained on each of the themes derived from the above section relating to the survivor syndrome, after interviewing survivors. The insights offer a practice perspective in relation to the private and public organizations, in which only brief differences are presented. They are of a nature that is rich and otherwise difficult to obtain had the principal research subjects not been these survivors, which might be the case in other scenario planning methodologies.

Perception of the selection process and fairness
About double the number of respondents in the private firm felt the process of redundancies was carried out unfairly, compared to that in the public organization. Survivors tended to be somewhat suspicious and fearful of the future, and this observation is in line with Fineman’s
(2003) finding that trust and loyalty are diminished by survivors who perceive a process being unfair. A typical comment was on the lines of:

[...] errh, it probably wasn't fair, in terms of a process I think it is still difficult to quantify coz I'm not sure I fully understood what their process was in the first place. (Male, 8 years service)

Only one respondent (in the private firm) had no view on fairness, but perhaps this was biased in that she worked closely with the chief executive during the redundancy process and therefore had increased her feeling of loyalty toward him and the organization, having been so engrossed in the rationale behind the downsizing. Those who felt that the selection process was fair worked in various positions, from managerial to manual work. These insights conform with Robbins' (1999) earlier work that identified the symptoms of the survivor syndrome as fear of change, loss of confidence in management and loss of loyalty, all of which influence an individual's perception of fairness. It seems that if survivors feel they have been ignored at any part of the downsizing process, their levels of anxiety, anger and mistrust have increased.

**Perception of organizational assistance**

The survivor syndrome is most likely to surface if employees have perceived that organizational support for those that are laid off is low. In the case of the private firm, an employee helpline was in operation throughout the redundancy process from the moment it was announced. Only three of the respondents actually mentioned this during their interview. When questioned further, it was found that those who were aware of the helpline had not made use of it and were not aware of anyone else who had utilized this facility. A respondent who was a trade union representative further explained that if any of the employees had issues, they would find it hard to approach the management and would therefore ask for assistance directly through the trade union. The union offered re-training which consisted of various courses at subsidized rates. In contrast, no such helpline existed in the public organization. Two factors seemed to explain this. First, it did not appear to be within the culture of the workforce to ask for assistance, plus the lack of communication from the organization on the assistance available. Second, the influence of the trade union was extremely prominent and was therefore easily accessible. An employee commented interestingly:

[...] you have to come back to the culture. I thought that [the union] helped, and I think this is one of the better parts of the union – they got help. I am pleased with that side of things, but it is very difficult at times. How far can you go with people? You can’t force them to take help. (Female, 12 years service)

This had been a long tradition with the type of work undertaken within the organization, and employees were probably more inclined to seek assistance via this route than make use of assistance offered by the organization.

Four respondents from the private firm and two from the public organization felt that no assistance had been offered at all. Again, this may have been due to the lack of communication from the organization to the workforce, or the individuals were of the mind-set that they did not require assistance and did not look for it. This is worrying, given Armstrong-Stassen’s (1994) suggestion that management need to ensure that perceived organizational support is sustained at all times to maintain motivation within the remaining workforce, and the survivors in particular need to feel that they have not been neglected and assistance has been provided, whether they utilize it or not. Again, organizational assistance is linked with perceived fairness which is increased if an organization is seen to assist survivors and victims, whereas the opposite effect will occur if assistance is not provided (see Baruch and Hind, 2000).
Perception of communication

The perception of communication within the two contexts was different. In the public organization, all respondents perceived that the top-down communication was sufficient. This information was about criteria based upon which an employee could stay or had to leave. In contrast, the bottom-up communication seemed intentionally ignored, partially due to the pre-selected preference of managers. Thus, it affected the perception of fairness. In the private firm, the initial communication of a loss of key clients/customers, leading to some redundancies, was viewed as positive at first. Following that initial meeting, contact was reduced and the majority of the survivors commented on a major lack of communication. For example, a survivor commented:

[...] very uncertain times for everybody – some people not receiving some of the information they need and a lot of people just in limbo really, not sure if it’s going to affect them, or not affect them, because some people were told their jobs were in danger, while others didn’t have a clue. (Male, 27 years’ service)

This lack of communication could have been due to a number of reasons. For example, the organization was probably not eager to relay bad news to its workforce. While Noer (1993) argues that bad news should be openly communicated as much as good news, this was not the case for either organization, and additionally there did not appear to be any plans put in place to help the survivors adjust to their new workloads. Furthermore, survivors wanting to seek information to allay their fears about the future (Datta et al., 2010) may, therefore, piece together information from different sources which could result in feeding the rumor mill which further damages the employer/employee relationship with consequent feelings of mistrust and resentment.

Feelings during the redundancy process

It emerged that most of the respondents had similar feelings during the redundancy process, and concur partly with research undertaken by Brockner et al. (1986) which identified anxiety and stress as the two main emotions. Some additional observations emerged:

- Feelings of shock: the initial reactions of the respondents appear to be of shock. The redundancy in the private firm was initiated by the loss of the organization’s major customer. In order to survive, the company announced in a memo to all staff that the only way the organization could reduce costs to the required levels would be to downsize and restructure the organization. In the public organization, the selection criteria which were sent out to employees at 5 p.m. on a Friday left many employees in shock as well. Shock was not a key element that was documented explicitly in the extant literature. With the process of being made redundant referenced by Fineman (2003) as being similar to that of experiencing grief, it is unsurprising that it has not been pursued further. The respondents’ sense of shock may have been caused by several factors. First, the scale of redundancies would have a severe impact on the remaining workforce – that being, an increase in workload, stress and anxiety for the future of the organization (Robbins, 1999). Second, the length of time taken to relay this information to staff in the private firm and the surprise that key customer was lost in the first place, when it had been stable for quite some time with no apparent hint of the impending loss of business. Third, the timing of the selection criteria being issued in the public organization was seen as devastating as employees did not have a chance to discuss with their colleagues those criteria and their possible impacts on each individual. Open communication would have been vital in this situation in order to keep rumors at bay and survivors informed, which in turn would have raised their perception of fairness (de Vries and Balazs, 1997).
Feelings of anger: only one respondent admitted explicitly to feeling anger. One other referred to his colleagues feeling anger but not himself. This contradicts Noer’s (1993) research that identified anger as being one of his four “feeling clusters” and Adams’ Equity Theory (see Brockner et al., 1986) which advocated how redundancy can create an assortment of emotional states. There may be many reasons why the majority of respondents did not feel anger. For example, it could be they felt secure that their role would remain safe and had high levels of self-esteem (Brockner et al., 1986). Consequently, they did not fear the outcome of the layoffs. Furthermore, due to the period of time that has lapsed since the redundancy and undertaking this research (between six and eight months), their perceptions of events at the time may have changed. Demography could also be another factor which has influenced their feelings. Nearly two-thirds of the sample had experienced redundancies before which could also explain the difference between these insights and those of Noer’s (1993) and Brockner et al.’s (1986).

Feelings of guilt: all respondents except one expressed that they did not have any feeling of guilt, which contradicts Brockner et al.’s (1986) research. The respondent who did express feelings of guilt was female and had worked for the organization for 20 years, while no male admitted to such a feeling, consistent with Fineman’s (2003) view that males tend not to be as subjected to emotions as females within the workplace. Furthermore, the males may have had higher levels of self-esteem due to their positions and security in their current role along with their length of service, and therefore felt justified that they have been able to keep their jobs (cf. Brockner et al., 2004).

Motivation
A redundancy that is not handled effectively will result in survivors demonstrating negative behavior and attitude, such as a lack of motivation (Appelbaum et al., 2003). The period of a survivor’s tenure also has an impact on their motivation levels (Furnham et al., 2009). Furnham et al. (2009) depict motivation as intrinsic to the role and their findings demonstrated that survivors from a management level or higher are less concerned with their working conditions and clarity in their work than those employees of a lower level. This view is consistent with the observations, as the majority of the respondents, mainly from the private firm, expressed that they still felt motivated following the redundancy process. As one respondent commented passionately:

[...] I’ve always liked the challenge – it doesn’t matter how much pressure. I’ve always loved the work before the pressure. I suppose it’s pride – you get something, you’ve achieved it. It isn’t a job, it’s in your blood, there’s nothing you can do. I worry if this place is going to be here in 50 years’ time, but I’m not going to be here, but I would still like to think that what I’ve worked for will still be here for somebody else. (Male, 27 years’ service)

This is also consistent with the view that self-esteem will impact on a survivor’s motivation, and individuals that display high self-esteem are more motivated than their counterparts who show low levels of self-esteem (Brockner et al., 1993, 2004; Marks, 2006). However, it seems that the female respondents who were not in a managerial position also displayed good levels of motivation.

Workload
Once a redundancy has taken place, it is inevitable that survivors’ workloads will increase as they take on their departed colleagues’ responsibilities along with their own. It is therefore necessary for organizations to have plans in place for the survivors to adjust to this increase (Appelbaum et al., 2003), otherwise any uncertainty may cause stress.
One example of how an organization may deal with workload increases is explained by a union representative, as follows:

[...] The organization tried to exploit what was happening – [they] saw us as negative hours the time we were paid for not being at work, and tried to impose annualized hours on us [...]. On the other side of the coin, for every tool that management can use against you, I can always quote the Working Times Directive – I can’t do that, I’ve got X number of hours between work breaks, etc. [...] In any negotiation, you’re gonna lose something, you’re gonna gain something – at the end of the day, it’s a two-way street, and in the end, common sense has got to prevail because you’re all aiming for the same goal. So, yes my workload has increased, marginally. (Male, 5 years’ service)

This view was consistent with all the respondents, suggesting that survivors are willing to accept the increase in their workloads despite facing job insecurity. There was no evidence that anyone felt the increase in workload had a positive effect by the need to enrich their skills (see Brockner et al., 1993). Four of the respondents even mentioned the need to resort to physical ailments following the increase in workload!

**View of the future**

A total of 12 respondents, most of whom were from the public organization, expressed that they were worried about the future and did not display a positive outlook. This manifested in a workforce that lacked confidence in the organization’s management and no eagerness from the survivors to look forward to the future, resulting in reduced performance and adding to the threat of the organization’s future survival. One respondent commented with great passion and concern:

[...] I worry about [the future] a great deal. I worry about people sitting on the Board who know nothing about the industry. I worry about Chief Executives who only rule by committee and no by leadership qualities. I worry about interfering Chairmen who are meant to be non-exec and have more say than they should. I worry about outside people dictating to what happens to this place. We are just custodians of this place and we should, instead of like I said just now, people looking to make their own little empire and their own mark on it, they should look on it like – if you do what I do in the mornings, get down here at 5:30 and look out across, and you see an odd house and think I’m so lucky to be here and you should look after it and respect it. But I see other people – all they see is a fast buck and making their mark and that annoys me and that’s what worries me. (Male, 33 years’ service)

The evidence challenges Thornhill et al.’s (1997) findings that management have a higher perception of the organization’s effectiveness following a redundancy process. It is particularly interesting as many of the respondents held managerial positions and did not feel this way. It appears that some of the survivors perceived a badly handled redundancy process which in turn had an impact on their thoughts for the future. With the six individuals who did not have any worries about the future, four of whom were already shortly due for retirement, which could explain their lack of concern; one respondent had worked closely with the chief executive during the redundancy process, and stated that she “does try very much not to worry about the future.” Even though the respondents were on average middle-aged, the findings have shown that regardless of age or gender, individuals tend to worry about the future following a redundancy.

**Proliferation of scenario planning methodology**

Literature from human resources management was used to understand how the survivor syndrome can be better managed, and if more accurate “scenarios” are better presented, an improved tool is conceived to carry out scenario planning. In this way, foresight is a metaphor for the implications of the survivor syndrome. The nature of the above observations and their deviation from the extant literature meant they would not have been possible to obtain had interviewing not been conducted with “survivors.” For example, rich
characteristics like feelings of fairness, support and shock would have been difficult to hypothesize from scenario planning participants had they not been close to redundancy themselves. Similarly, the lack of communication and lack of confidence in management are post-experiences actually felt and difficult to gauge if scenario planners were to bring them up as possible scenarios. In the same way, the traps and pitfalls explained earlier in this paper have been better controlled for: bias, for example, would not exist as the views are from genuine participants who would otherwise not be affected, which in turn, overcomes the issue of relevance, and of course, establishing greater plausibility.

Given the benefits of the use of survivors as part of scenario planning, we recommend building in the seven themes as part of the interview schedule to ensure their coverage. However, unlike other theorists who specify steps or procedures, we only argue for their inclusion rather than impose a rigid order for the approach. The application of the seven-themed framework and method of interviewing used may be repeated for understanding similar issues where the context might vary, or be augmented. The interview themes of fairness, organizational assistance, feelings, motivation, communication, workload and the future are the core anchors for designing interviews if used as an instrument for conducting scenario planning, particularly in the part where scenarios are raised. Survivors were used in the research as they pose the closest research subject to those who would be victims but escaped redundancy by a narrow margin. In repeated use, alternatives to survivors can be substituted depending on the economic and political contexts and availability of alternatives to survivors, but they must satisfy the criterion of people who have come extremely close to a situation without having fallen so.

In the “tin anniversary” of the 2008 global financial crisis, its context, data from researched subjects and representative phenomena still stand as valuable conditions on which to conduct research on performance management improvement and methodology. Scenario planning methodology was a pioneering approach of the 1950s (Wack, 1985), and has proliferated well in the guise of organizational foresight, via the HR field (such as managing the survivor syndrome), in performance management. Its proliferation is due to the good closeness of fit between the need of scenario planning to raise accurate future scenarios and the feelings felt by survivors of the survivor syndrome who have come close to being a victim and incited a retaliatory response concerning the future of the firm. So, to summarize: improved organizational performance is the ultimate aim, which can be achieved by better strategic decision making, informed by foresight, in the form of scenario planning, which is well achieved through exploring human subjects who meet the criterion of a “survivor.” Hence, as foresight sits in the middle and acts as the medium, it is the metaphor for all the insights raised by the survivors presented in the illustration part of this paper. So, foresight is an important metaphor; it is scenario planning; and for the future of organizations […] it matters!

References


Further reading


Corresponding author

Vinh Sum Chau can be contacted at: v.s.chau@kent.ac.uk

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com