The Development of a Performance Measurement System for Indirect Procurement: A Delphi Study

Abstract

Purpose – This study aims to address the current gap in knowledge on indirect procurement performance management. It attempts to argue the need for a specific and tailored performance management approach for the indirect procurement function that incorporates a balanced approach, beyond financial measures.

Methodology – The case study approach evaluated Key Performance Indicators (KPIs) from a Balanced Scorecard perspective in the development of a Performance Measurement System (PMS) for a Middle Eastern university's indirect procurement division. It initially reviewed the literature to assess potential indicators for this context. It utilised vision and mission statement analysis alongside expert interviews to augment the literature. The candidate indicators were then evaluated and ranked by an expert panel through applying a four round Delphi technique.

Findings/results – Twenty-nine procurement specific indicators are suggested in a Balanced Scorecard framework. The five highest ranked indicators were not in the Financial Perspective unlike other BSC studies in the broader field of supply chain management.

Research limitation –This is a single case study; its findings cannot be generalised.

Managerial implications – The study suggests a framework and indicators for a procurement performance measurement system for practitioners to consider. It also highlights there is no one-size-fits-all and that organisations need to tailor performance measurement to the organisation and divisional strategy, and operational needs. This study aids the development of guidelines for executives and procurement management that wish to develop indicators and a PMS.

Originality/value – This study contributes to knowledge by partly addressing the under researched field of indirect procurement performance measurement. The literature suggested that various roles in supply chain management require specific performance measurement indicators. This study puts forward a BSC framework with twenty-nine indicators specifically for indirect procurement. Fourteen of these indicators were derived from non-literature sources. This study enhances knowledge and contributes to the limited debate and evidence on indirect procurement performance measurement and the broader performance measurement literature.

Keywords: Performance Measurement, Metrics, Procurement, Balanced Scorecard, Delphi study, Key Performance Indicators

Paper type: Research article

1. Introduction

Performance Measurement (PM) research has been undertaken over the past thirty years (Gopal and Thakkar, 2012, Bititci *et al.*, 2018, Bourne *et al.*, 2018, Bourne, Melnyk and Bititci, 2018). It has gained significant momentum in the last two decades (Taticchi and Balachandran, 2008, Barrows and Neely, 2011).

A singular definition of PM Systems (PMS) is elusive due to the diversity of the field of research and the wide-ranging backgrounds of its seminal authors (Neely *et al*, 1995; Beamon, 1999). Ravelomanantsoa *et al.* (2019) identify a multitude of definitions and designs of PMS over the last one hundred years. Melnyk *et al.* (2014) advise that it is up to each system to adopt a definition that will appropriately and efficiently measure the system.

Generally, a PMS may be considered to be "*the set of metrics used to quantify both the efficiency and effectiveness of actions*" (Neely, Gregory and Platts, 1995). In today's highly competitive market, organisations require systems that measure their performances so that they can be effectively managed (Biazzo and Garengo, 2012). Therefore, connecting PM with strategy is a common problem that needs to be solved (Neely, 2005, Pun and White, 2005, Micheli and Mura, 2017).

There is a large corpus of literature that examines Supply Chain Management (SCM) but the majority of PM studies in that field focus upon the corporate performance of supply chains and direct (revenue generating) procurement of raw materials/goods in the manufacturing sector (Garengo, Biazzo and Bititci, 2005, Neely, 2005, Cox *et al.*, 2005, Maestrini *et al.*, 2017, Guersola, De Lima and Teresinha Arns Steiner, 2018, Maestrini *et al.*, 2018). There is comparatively little research that explores the internal performance of functions that undertake indirect (non-revenue generating) procurement of services (Gunasekaran, Patel and Tirtiroglu, 2001, Gunasekaran, Patel and McGaughey, 2004, Ellram, Tate and Billington, 2004, Cox *et al.*, 2005, Davis and Novack, 2012). In particular, there is a paucity of research that explores PM within a procurement context (Beamon, 1999, Lambert and Pohlen, 2001, Gunasekaran, Patel and Tirtiroglu, 2001, 2004, Wagner and Kaufmann, 2004, Gunasekaran and Kobu, 2007, Gopal and Thakkar, 2012, Mishra *et al.*, 2018) This is of concern since Carter *et al*. (2000) suggests that a purchasing function is unique from other functions and a unique approach might be necessary for this complex activity (Cox *et al.*, 2005). Sangwa and Sangwan (2018) argue the importance of effectively measuring all functions of an organisations in order to address efficiency and support lean initiatives.

A number of studies examine the relationship between strategic procurement and firm performance (Nair, Jayaram and Das, 2015). However, many organisations are measuring suppliers' performance but are not effectively measuring internal processes, such as procurement. Organisations tend to measure procurement's performance based on cost savings. Considering procurement, in particular indirect procurement, services other departments in the organisation, this is a very limited approach (Caniato, Luzzini and Ronchi, 2014). Contemporary performance measurement should include non-financial and financial measures that are linked to the organisation's strategy (Franco-Santos, Lucianetti and Bourne, 2012).

There is an absence of research on performance measurement on indirect procurement functions. This study utilises the Balanced Scorecard (BSC) as the basis for the development

of an indirect procurement PMS. It reports the development of a Performance Measurement System (PMS) that includes ranked and balanced internal KPIs for a Middle Eastern university's in-house indirect service procurement division.

This paper contributes to knowledge in the following ways:

1. Aids in addressing the gap in knowledge within performance measurement for indirect procurement
2. Contributes to the performance measurement debate
3. Augments and integrates Bigliardi and Bottani's (2010) framework into a procurement context
4. Suggests additional procurement specific metrics for practitioners to consider
5. Develops an indirect procurement performance measurement framework by combining BSC, Delphi, and KPIs

2. Performance measurement systems in supply chain management

This literature review examines previous studies on PMS in supply chain management and procurement. It attempts to collate potential indicators for a performance measurement system in procurement. Indicators are adapted from previous BSC systems that are externally focused as opposed to internal procurement. Initial searches for the terms "indirect spend" or "indirect procurement" in the title using the ISI Web of Science and Emerald databases produced few results. No procurement related articles for "indirect spend" and only two articles for "indirect procurement". This demonstrated the lack of empirical research in the indirect procurement field and/or functions. Moreover, the authors could not find any literature in the application of existing PM frameworks in an indirect procurement setting, which focuses on procurement of services, in a non-profit institution such as a university. The search was broadened to Google Scholar (peer reviewed journals and academic books) and included additional search terms: "balanced scorecard", "BSC", "performance measurement", "performance management", and "supply chain". This was supplemented by the "related citation" and "cited by" facilities.

2.1. PMS design

Designing a PMS is an important but complex process (Neely *et al.*, 2000, Taticchi and Balachandran, 2008, Taticchi, 2010, Okwir *et al.*, 2018). It is intellectually challenging but also fulfilling for all parties involved (Neely *et al.*, 2000). There is no right or wrong PMS design since each organisation needs to develop a system that is suitable with its values, operations, requirements, structures and models (Taticchi, 2010). Guidelines have been developed by a number of authors (Neely *et al.*, 2000, Mills, 2017). However, there are limited academic papers on PMSs in the supply chain and procurement domain. The limited studies on PMS in logistics, procurement and supply are focused on the practice–corporate performance link from a financial measures perspective (Komatina, Nestic and Aleksić, 2019, Shao, Moser and Henke, 2012, Zimmermann and Foerstl, 2014, Hofmann, 2014). There is no consensus on a framework for a procurement (direct or indirect) PMS (Belvedere, Grando and Legenvre, 2018).

The adoption of a PMS can be a valuable undertaking for organisations. It may for instance be a useful motivational tool (Verweire and Van Den Berghe, 2004; Smith and Mobley, 2008), as well as being a device for change (Chennell *et al*., 2000). However, the poor choice of

performance measures may have undesirable consequences (Ravelomanantsoa, Ducq and Vallespir, 2019). PMS implementation may also be ineffective due to lack of vision and strategy (Wagner and Kaufmann, 2004; Biazzo and Garengo, 2012), low commitment (Bourne *et al.*, 2000; Wagner and Kaufmann, 2004) and the absence of sensible data (Wagner and Kaufmann, 2004, Melnyk *et al.*, 2014).

2.2. The Balanced Scorecard and Supply Chain Management

Several frameworks exist that provide the basis of a PMS including Balanced Scorecard (BSC (Kaplan and Norton, 1992), Performance Prism (Neely *et al*., 2001), Supportive Performance Measures (Keegan *et al*, 1989), Results and Determinants (Fitzgerald, 1991) and the Integrated PMS (Verweire and Van den Berghe, 2004, Sangwa and Sangwan, 2018). Typically, these identify top-level strategic measures and decompose these into their respective operational measures of performance (Lynch and Cross, 1991, Bititci, 1994, Taticchi, 2010, Biazzo and Garengo, 2012, de Oliveira, Carneiro and Esteves, 2018).

Applications of BSC in SCM are comparatively scarce. Chia *et al.* (2009) use a BSC, comprising of fifteen measures, to assess the performance of logistics' functions in Singapore. Their study identified the top three performance measurement indicators were gross revenue, profit before tax, and cost reduction. Their study reflects the emphasis that is placed upon financial goals across the field of SCM (Gunasakaran and Kobu, 2007). Balfaqih *et al*. (2016) provide an overview of the supply chain performance literature. A number of studies have utilised the Balanced Scorecard approach, but procurement is not the focus. Bhagwat and Sharma (2007) apply the BSC to a broad study on supply chain management. They found that internal business processes neglected crucial measures for day-to-day operations. Cunha Callado and Jack (2015) apply the BSC to four nodes in the supply chain (suppliers, producers, distributor, and retailers). The measure of minimising cost was addressed in their study, which could be considered to be an indirect link to procurement. However, procurement was not specifically considered. Cunha Callado and Jack (2015) found that BSC measures are different per role in the supply chain. They recommended that each role should develop their own BSC measures as a common supply chain BSC is not feasible.

Bigliardi and Bottani's (2010) study identify relevant KPIs for the food supply chain. They put forward two BSC models which include two sets of KPIs. Table I illustrates the KPIs from Chia *et al.* (2009), and Bigliardi and Bottani (2010).

INSERT TABLE I HERE

2.3. Reviewing relevant indicators

This study reviews and evaluates indicators used in previous SCM PM academic papers to identify candidate KPIs. They are then classified into the four BSC perspectives. Previous studies have relied on adapting factors from external customer service literature (Young and Varble, 1997, Brandon-Jones and Silvestro, 2010, Johnston, 2008, Sharma, Kong and Kingshott, 2016). Internal customer service has been under researched, particularly in procurement (Minjoon and Shaohan, 2010, Brandon-Jones and Silvestro, 2010, Brandon-Jones, 2017). In the review, fitting indicators were incorporated without modification. A number of indicators were modified to better reflect the internal procurement context.

2.3.1. Supply Chain Performance Measures

Designing supply chain measures is a challenging task and limiting the number of the metrics used is difficult (Lapide, 2000, Chae, 2009, Gopal and Thakkar, 2012). Moreover, the majority of measures in supply chains are quantitative rather than qualitative (Gopal and Thakkar, 2012), very cost focused (Beamon, 1999) and therefore not balanced. Previous studies have utilised the Supply Chain Operation Reference (SCOR) model and quantitative analyses to develop supply chain performance measures that are principally externally focused (Akkawuttiwanich and Yenradee, 2017). In the absence of specific KPIs for an indirect procurement division, highly cited studies in supply chain KPIs by Gunasekaran *et al*. (2001) and Bradon-Jones (2017) alongside other studies (Cavinato, 1987, Beamon, 1999, Chia, Goh and Hum, 2009, Bigliardi and Bottani, 2010) were reviewed in order to identify potential indicators that can be used, as-is or with some degree of modification.

It is not the purpose of this study to cover all available SCM measures, since the focus is to identify measures for an indirect procurement division. Nevertheless, some specific SCM metrics that were highlighted by Gunasekaran *et al*. (2001, 2004) have a degree of transferability into an indirect procurement division.

2.3.2. Selecting the appropriate financial perspective indicators

Gunasekaran *et al*. (2001, 2004) has significantly contributed to the supply chain PM field, but the majority of the supply chain performance measures that were identified by them are not suitable for indirect procurement. PMS financial indicators tend to have been designed to measure corporate performance and specifically in Gunasekaran *et al*. (2001, 2004), performance around the supply chain of goods. Some financial indicators that were suggested in the literature can be used for measuring an internal procurement function. However, other indicators would need modification, while keeping the essence of the original KPI.

While moving beyond pure financial perspective in PM is important, for procurement "*cost savings*" (Gebauer and Segev, 2000, Gunasekaran, Patel and Tirtiroglu, 2001, Cunha Callado and Jack, 2015), "*cost avoidance*" (Pohl and Förstl, 2011), and *"variations against budget"* (Bigliardi and Bottani, 2010) are relevant.

"*Cost per operation hour*" (Gunasekaran, Patel and Tirtiroglu, 2001, Bigliardi and Bottani, 2010) is a prominent indicator in SCM. However, an internal procurement function, without inventory or manufacturing may benefit from a broader measure on an annual bases, "*division's operating cost (annual)*".

2.3.3. Internal Business Indicators

Internal business process performance measures have a significant impact on the operational performance in supply chain environments (Gunasekaran and Kobu, 2007). Specifically, "*Order lead time*" or "*Cycle time*" is highlighted in a number of key studies (Gunasekaran, Patel and Tirtiroglu, 2001, Bigliardi and Bottani, 2010). This indicator is a leading indicator, if it is measured frequently (e.g. weekly or monthly). Also, *"reliability of cycle time"* is identified as having a high impact on customer satisfaction levels (Towill, 1997, Cavinato, 1987, Gamini, 2011, Parasuraman, Zeithaml and Berry, 1994).

An efficient and effective customer order process is key in SCM (Gunasekaran, Patel and Tirtiroglu, 2001). "*Lean Six Sigma*" initiatives focus on value and striping out waste. Thereby, contributing towards *"waste reduction"* along the internal processes (Cunha Callado and Jack, 2015).

Also, "*effectiveness of scheduling techniques*" which refers to planning and scheduling a variety of tasks that need to take place along the supply chain (Gunasekaran and Kobu, 2007). This relates to procurement in terms of the *"effectiveness of the procurement plan*". That is, the scheduling of time and date for the needed tasks, determination of resources and their relevant flow in the operating system.

2.3.4. Customer Perspective Indicators

The fulfilment of internal customers' needs is a significant objective of a purchasing function and the customers' perspectives should be a priority for top management (Kumar, Ozdamar and Ng, 2005, Bernardo, 2018).

The relevance of "*range of products and services*" KPI (Gunasekaran, Patel and Tirtiroglu, 2001, Mapes, New and Szwejczewski, 1997, Cunha Callado and Jack, 2015) would need to be modified for applicability to an indirect procurement division. It could take the form of "*number of categories supported*". This would indicate whether the division has internal expertise to support the procurement of IT, HR, facilities contracts etc. The more categories a division supports, the more internal customers indirect procurement can serve.

"*Measuring customer service and satisfaction*" is argued to be highly important (Beamon, 1999, Gunasekaran and Kobu, 2007, Cunha Callado and Jack, 2015) and is commonly measured with the SERVQUAL instrument to assess the quality of service provided to external customers (Parasuraman, Zeithaml and Berry, 1988) as well as to internal customers (Large and König, 2009). "*Flexibility*" is a typical supply chain performance measure (Beamon, 1999), and has been previously used in an internal customer setting (Finn *et al.*, 1996).

Gunasekaran and Kobu (2007) identifies "*customer query time*" and "*service level compared to competitors*" as important KPIs. The latter's relevance to an internal division is limited and can be modified to match the internal company environment and therefore changed to "*service level compared to other divisions*".

2.3.5. Selecting the appropriate Innovation and Learning perspective indicators

A number of studies identify "*Order entry method*" as a key component of the Innovation and Learning perspective (Bhagwat and Sharma, 2007, Gunasekaran and Kobu, 2007, Bigliardi and Bottani, 2010). This KPI refers to how customer requirements are converted into a purchase order. For an indirect procurement division that procures service contracts, this can be referred to as "*effective reflection of customer requirements into contracts*". Chia *et al*. (2009) suggest four KPIs under this perspective that can be used in any setting; "*employee satisfaction*", "*employee turnover per year*", "*number of suggestions implemented per employee yearly*" and "*money invested in employee training yearly*", which are all candidate indicators.

A summary of the candidate indicators is shown in Table II. The literature review has identified twenty candidate indicators principally from SCM PM literature. These indicators were then augmented and evaluated by expert judgment.

INSERT TABLE II HERE

3. Research Methodology

This case study gathered and analysed evidence from multiple sources (Mills, Durepos and Wiebe, 2010). With the review of the literature undertaken to ensure that the candidate indicators are well grounded, we describe and justify the next steps in the research approach to augmenting, analysing, and verifying KPIs.

The BSC and KPIs should be aligned with the vision and mission (Biazzo and Garengo, 2012). The procurement services department vision and mission were content analysed to verify alignment with the candidate indicators.

An additional cross check was implemented in line with the view that BSCs should be driven from top management (Biazzo and Garengo, 2012). An unstructured conversational expert interview approach on procurement performance metrics was undertaken with senior management. This approach was chosen to allow flexibility, the ability to explore opinions, ensure understanding and response validity (Lavrakas, 2008, Zhang and Wildemuth, 2009). Convenience sampling was used to select the interviewees. A university corporate director and senior manager with responsibilities for procurement were interviewed. The interviewees were not informed of the previous indicators identified. It was content analysed and additional indicators were added to the candidate list.

The complete list of candidate indicators was empirically evaluated through a Delphi approach. **This approach (combining KPI selection with the Delphi technique) has been used by other studies** (Yeung, Chan and Chan, 2009, Hübner-Bloder and Ammenwerth, 2009, Bigliardi and Bottani, 2010, Tsai and Cheng, 2012). **The** Delphi technique is a consensus multistage method, invented at Rand Corporation in the fifties (Dalkey and Helmer, 1963). It is commonly used in the health and social sciences field (Hasson, Keeney and McKenna, 2000). Even though the technique has been criticised (Goodman, 1987, Williams and Webb, 1994, Hasson and Keeney, 2011), it is well established (Powell, 2003) and is particularly useful for its ability to measure forecasting (Rowe and Wright, 1999) and enhance decision-making (Hasson, Keeney and McKenna, 2000).

The Delphi technique was chosen because it was able to provide (1) anonymity to the participants, (2) flexibility as to when the panel members could provide responses, and (3) time flexibility for the researchers. Delphi is also considered more accurate than comparative staticized groups and unstructured interactive groups (Rowe and Wright, 1999). Only a few scholars have combined the BSC, Delphi and KPIs together (Bigliardi and Bottani, 2010, Tsai and Cheng, 2012, Melnyk *et al.*, 2014, Balfaqih *et al.*, 2016), out of which only two are supply chain focused (Bigliardi and Bottani, 2010, Kasiri, Sharda and Hardgrave, 2012). Figure 1 illustrates the research approach.

INSERT FIGURE 1 HERE

To safeguard face and construct validity, a respondent-driven pre-test with one procurement professionals within the university's procurement division (non-members of the expert panel) was completed (Ruel, Wagner III and Gillespie, 2015). The pre-test was completed to confirm the comprehensiveness and understanding of the questionnaire.

In order to ensure the quality of the data gathered and the results, a rigorous phased approach was used by implementing Okoli and Pawlowski's (2004) guidelines. In this study and in line with the classic Delphi technique, a four round Delphi was chosen (Young and Hogben, 1978). The first round was a mixture of open-ended and close-ended questions. Close-ended questions were previously used by Chia *et al*. (2009). The open-ended questions allowed the respondents to suggest additional indicators, if they desired. This was followed by three rounds of closed questions. Test-retest reliability is not required with Delphi, since the experts are expected to revise their opinion in view of the group's opinion (Okoli and Pawlowski, 2004). Moreover, a Delphi study captures a snapshot of group opinion at that moment in time. Therefore, the application of test-retest is invalid (Hasson and Keeney, 2011). There are four elements that are embedded in a Delphi that enhance the validity of the technique (Hasson, Keeney and McKenna, 2000); (1) it is assumed that group opinion is stronger than individual opinion, (2) using experts increase content validity (Goodman, 1987), (3) iteration increase the concurrent validity and (4) high response rates increase the validity of results.

Okoli and Pawlowski (2004) recommend the panel to consist of 10-18 experts. In order to ensure informed individuals, pre-selection criteria was defined and applied (McKenna, 1994):

- panel member is working for the Procurement Department with four years' experience; and
- has attended PM workshops; and
- has attended top level strategy meetings run by the organisation

As the procurement division has a small team, a convenience sampling approach was used. The first 12 professionals that agreed to participate and fulfilled the criteria were accepted. The balanced panel comprised of five managers, six senior professionals, and one PM professional.

4. Analysis and findings

In this study, information and data on KPIs were collected from four sources, (1) the literature, (2) senior management, (3) an existing departmental vision and mission statement, and finally (4) from an expert panel. The analyses of sources 2, 3 and 4 are presented next.

4.1. Vision, mission and senior management analysis

The content analysis of the vision and mission identified two key elements of the mission statement that were not measured by indicators derived from the literature. Two additional indicators were added. They were taken directly from the mission statement of the procurement division. The two additional KPIs, "*clarity of processes and procedures*" and "*usability of procurement tools*", were linked to the BSC's internal perspective.

The senior management unstructured interviews content analysis revealed eighteen indicators. It corroborated nine candidate indicators that were already identified in the literature review. Table V shows these nine indicators. An additional nine indicators were identified. They

tended to be more sophisticated measures of financial performance and internal processes. The additional indicators are:

1. *Percentage of competitive bidding versus sole source bidding ($)*
2. *Number of transactions over $1M during the fiscal year*
3. *Average value of commitment per contracting representative per FY*
4. *Managed spend per fiscal year ($)*
5. *Percentage of spend on long form contracts (LFC/managed route) versus percentage of spend on short form contracts (SFC) and special consultant contracts (SCC) (unmanaged)*
6. *Benefit versus cost of division*
7. *Number of claims and contractual disputes per year*
8. *Ratio of completed contracts during the fiscal year to number of staff*
9. *Percentage of qualified personnel within the team.*

The candidate indicators identified from the literature review (20), vision and mission analysis (2), and senior management interviews' analysis (9 additional) were inputted into the pre-test. Following the pre-test two minor clarifications were made to the questionnaire's instructions and one amendment to a question. After the pre-test, the researcher applied the same process with all individuals of the expert panel separately. i.e. a communication email, followed by a face-to-face meeting, during which a consent to participate to the study was also provided, and a final email which included the questionnaire (first round). Table III outlines an overview of the Delphi process.

4.2. 1st round of Delphi study

Section A asked the expert panel to indicate (by selecting yes/no) whether the proposed thirty-one candidate indicators, in principle, should be measured for monitoring the performance of the Contracting Division. Section B asked the respondents to indicate one additional KPI for each BSC perspective (optional). Section B was optional because the preselected KPIs might have been considered adequate for some panellists (Keil, Tiwana and Bush, 2002).

INSERT TABLE III HERE

All the responses were analysed, and the results fed into the questionnaire for round two. Table IV presents the Delphi round one results using number of yes's (N) and percentage of respondent agreement (%) (Hübner-Bloder and Ammenwerth, 2009). A small volume of KPIs were proposed by the respondents. Content analysis and grouping similar responses together was not required (Yeung, Chan and Chan, 2009). The data was already grouped by the expert participants. There were only eight proposed indicators that were very unique in nature, which did not require further grouping. Thus, the suggested eight additional KPIs remained as proposed for the subsequent round. The suggested indicators by the panel are listed and were classified into; Financial (1,2,3), Internal Business (4,5), Customer (6,7), Innovation and Learning (8):

1. *Percentage of annual contract commitment growth compared to previous financial year*
2. *Percentage of long form contract actions that were presented to the Tender Committee*
3. *Percentage of long form contracts waived to short form contracts*
4. *Percentage of contractors that are evaluated at least twice a year*

5. *Number of signed contracts using the contractor's template or in-house template with significant adjustments*
6. *Number of rejected contract requisitions per month*
7. *Number of internal divisions that contracting has service level agreements with*
8. *Average number of training days per employee per year.*

INSERT TABLE IV HERE

4.3. 2nd round of Delphi study

In the second round of the Delphi, a second questionnaire was developed and the collective expert panel opinion was shared with all individuals of the panel that responded to the first questionnaire. Two individuals did not participate in the first round and were not invited in subsequent rounds. The respondents were asked whether they would like to revisit any part of their initial response that was provided in the first round, but also to provide feedback (yes/no) on the additional KPIs that were proposed by the panellists.

4.4. 3rd round of Delphi study

The questionnaire that was developed for round three contained the finalised opinion of the expert panel illustrated in percentage points, for the originally proposed KPIs. This part was presented to the panel for information only, and no further opinion was solicited for these indicators. The collective expert panel opinion on the additional KPIs was presented in percentage points. The respondents had again the chance to revisit their opinion in view of the expert panel feedback on these additional KPIs.

There is no common panellist consensus threshold in the literature (Hasson, Keeney and McKenna, 2000, Powell, 2003). In order to be somewhat conservative with eliminating KPIs, the research adopted a 51% consensus threshold as proposed by Loughlin and Moore (1979). KPIs below the threshold were removed for the following round.

4.5. 4th and final round of Delphi study

The remaining twenty-nine KPIs were utilised in round four's questionnaire. This round collected importance values, based on a 1-5 Likert scale (1 = not important, 5 = extremely important) (Chia, Goh and Hum, 2009, Bigliardi and Bottani, 2010). Table V shows the results. The panellists were consistent respondents throughout rounds two to four and the researcher requested a resubmittal of any questionnaire that contained missing values.

INSERT TABLE V HERE

4.6. Consensus and consistency analysis

Panellist consensus, and consistency is an importance element of analysis for Delphi studies (Brancheau and Wetherbe, 1987, Nevo and Chan, 2007, Yeung, Chan and Chan, 2009). Kendall's Coefficient of Concordance (W) was applied to assess the level of panellist agreement (Kendall and Gibbons, 1990).

The Kendall's W value for the twenty-nine indicators was 0.3207, suggesting that there is a weak to moderate agreement between the panellists across the final indicators (Schmidt, 1997).

Considering a moderate consensus threshold was set, and lower scores are more dispersed, this is considered acceptable. In other Delphi studies that aimed to identify KPIs, only Yeung *et al*. (2009) provide the Kendall's Coefficient of Concordance for their top seven indicators (0.249). Bigliardi and Bottani (2010), Tsai and Cheng (2012), and Hübner-Bloder and Ammenwerth (2009) do not report Kendall's W. However, Hübner-Bloder and Ammenwerth (2009) epigrammatically provide the inter-rater reliability without any further analysis.

The Kendall's W value per BSC perspective identified that the highest agreement was in Innovation and Learning (W=0.446), followed by Internal Business (W=0.4026), and then Financial (W=0.2165). Finally the Customer perspective (W=0.1141) had very little agreement between the panellists. The Innovation and Learning perspective indicators were common and generic in nature with the exception of "*effective reflection of customer requirements into contracts"*. Therefore, the panellists relatively agreed with each other since they are familiar KPIs. The Kendall's W value for Internal Business indicators can be explained by the high level of agreement on the top three indicators in this perspective. The high scores are probably due to the respondents' assessments of the potential impact (disruption and cost) and delays in the procurement process. The Kendall's W values for the Financial, and Customer perspectives were low. Table V shows the large dispersion of respondent ratings across the majority of indicators in those perspectives.

Comparing the third and the fourth round, response consistencies were high. KPIs that received only 60% of yes' in the third round of the Delphi study also scored low importance values (3 or below) in the fourth round. Figure 2 displays this relationship.

INSERT FIGURE 2 HERE

In contrast with other Delphi studies (Yeung, Chan and Chan, 2009), there was very minor opinion change by the panellists in this study. The positive relationship between the third and fourth round of the Delphi, as per Figure 2, demonstrates the consistency and dependency of data between these rounds (Von der Gracht, 2012). Therefore, the KPIs that were very close to the cut off threshold following the 3rd round of Delphi did not achieve high importance scores in the fourth round. This shows that the panellists voted and completed the questionnaire based on their strong and consistent opinions.

5. Discussion, the Developed BSC Model

The literature review highlighted the limitation of early PMSs being overly financially focused. The BSC was considered to be one of the tools that had the capability of addressing this issue, providing focus and balance. Figure 3 displays the KPIs in the developed BSC model, post Delphi analysis.

It is evident that the developed framework and indicators provides balanced measures. This is demonstrated by each perspective having one or more of the top six KPIs that are within the very important to extremely importance range (score above 4).

In the final round the top three ranked indicators in each of the Financial, and Internal Business perspectives were drawn from the literature and senior management. Fifteen indicators came

from the literature. Eighteen indicators were derived from management. However, nine overlapped with the literature indicators, leaving an additional nine that were empirically derived from senior management. Three indicators were from the Delhi panellist and finally two from the vision and mission statement analysis. Table V illustrates the final indicators' ratings and sources. The research suggests an additional fourteen indicators tailored for procurement beyond the literature review.

Chia *et al*'s. (2009) study of performance in logistics functions in Singapore highlights the importance of the indicators; gross revenue, profit before tax, and cost reduction, which reflects the emphasis that is placed upon financial goals across the field of SCM (Gunasakaran and Kobu, 2007). In contrast, this study finds priorities in the Internal Business, Innovation and Learning, and the Customer perspectives over the Financial Perspective.

Bhagwat and Sharma (2007) find that internal business processes neglected crucial measures for day-to-day operations in SCM. This research addresses this gap by suggesting and developing measures on the Internal Business perspective for procurement's day-to-day operations.

Cunha Callado and Jack (2015) recommended that each role in SCM should develop their own BSC measures as a common supply chain BSC is not feasible. This study attempted to address this and suggests twenty-nine indicators specifically for procurement.

INSERT FIGURE 3 HERE

6. Conclusion, implications and future research

This study developed a novel framework utilising the BSC and Delphi techniques. The approach ranked KPIs for measuring the internal performance of an indirect procurement division.

Traditionally cost savings are identified as a priority measurement for procurement. Cost savings were highly ranked in this study; however, they were not the highest rated indicator. The high ranking of non-financial measures support the view of a contemporary approach to performance measurement, and procurement as an internal service to other functions.

6.1. Contribution to theory

This study partly addresses the gap in knowledge on procurement performance management systems. It contributes to theory building by addressing the lack of specific performance measures for indirect procurement. It assessed twenty-nine indicators through the Delphi technique, which were embedded in a BSC framework. Fourteen of the final indicators originated from the empirical evidence. The top five highest ranked indicators were in non-financial perspectives. This adds support to the current debate for a balanced performance measurement system. It also contributes to the debate that procurement should not be measured on cost savings alone. The phased approach in this study also contributes to PMS implementation guidelines.

6.2. Implications for practice

Procurement has traditionally been a cost minimisation function. It plays a fundamental role in organisations and within their supply chains. As strategies develop that incorporate factors beyond cost minimisation, to support resilient and robust supply chains, procurement must incorporate these factors in order to support organisational strategies. To achieve this performance measurement and KPIs are required.

This study's approach and framework can be tailored to individual organisations allowing them to drive improvements, and support a strategic approach to procurement. It suggests twenty-nine procurement indicators and a BSC framework for practitioners to consider. While there is no one-size-fits-all, the framework and indicators are a starting point for consideration and can be further tailored to support the particularly organisation's procurement strategy and goals.

This studies development of a Balanced Scorecard, supported by a Delphi approach, is a potential method to refocus the procurement function to incorporate key principles of strategy and the dynamic environment that organisations and supply chains operate in. This approach can aid senior management, in a variety of sectors and roles, with the difficult task of implementing and operationalisation of strategies. Executives and directors could use this approach to aid change management in their organisations and develop procurement to support the organisations' strategies. Also, the use of the BSC and Delphi approach is more manageable for organisations with limited resources, such as SMEs, to apply in comparison to other complex quantitative approaches.

The results highlight the importance of non-financial measures. Viewing procurement as an internal service would aid integration and support a process-based approach rather than functional.

6.3. Future research

The framework provides a foundation for further research and suggests a starting point for the applicability of these indicators for organisations. However, this research was limited to one organisation. Further empirical work could assess key performance indicators from a quantitative approach. As procurement can be viewed as servicing internal customer, internal customers' views and indicators would potentially add value to the framework.

REFERENCES

Akkawuttiwanich, P. and Yenradee, P. (2017) 'Evaluation of SCOR KPIs using a Predictive MILP Model under Fuzzy Parameters', *Int. J Sup. Chain. Mgt Vol,* 6(1), p. 172.

Antonio André, C.C. and Jack, L. (2015) 'Balanced scorecard metrics and specific supply chain roles', *International Journal of Productivity and Performance Management,* 64(2), pp. 288-300.

Balfaqih, H., Nopiah, Z.M., Saibani, N. and Al-Nory, M.T. (2016) 'Review of supply chain performance measurement systems: 1998–2015', *Computers in Industry,* 82, pp. 135-150.

Barrows, E. and Neely, A. (2011) *Managing performance in turbulent times: analytics and insight.* John Wiley & Sons.

Beamon, B.M. (1999) 'Measuring supply chain performance', *International Journal of Operations & Production Management,* 19(3), pp. 275-292.

Belvedere, V., Grando, A. and Legenvre, H. (2018) 'Testing the EFQM model as a framework to measure a company's procurement performance', *Total Quality Management & Business Excellence,* 29(5-6), pp. 633-651.

Bernardo, N. (2018) *Agile Procurement : Volume I: Adding Value with Lean Processes.* Cham: Cham : Springer International Publishing : Imprint: Palgrave Macmillan.

Bhagwat, R. and Sharma, M.K. (2007) 'Performance measurement of supply chain management: A balanced scorecard approach', *Computers & Industrial Engineering,* 53(1), pp. 43-62.

Biazzo, S. and Garengo, P. (2012) *Performance Measurement with the Balanced Scorecard: A Practical Approach to Implementation Within SMEs.* Springer Verlag.

Bigliardi, B. and Bottani, E. (2010) 'Performance measurement in the food supply chain: a balanced scorecard approach', *Facilities,* 28(5), pp. 249-260.

Bititci, U.S. (1994) 'Measuring Your Way to Profit', *Management Decision,* 32(6), pp. 16-24.

Bititci, U.S., Bourne, M., Cross, J.A., Nudurupati, S.S. and Sang, K. (2018) 'Editorial: Towards a Theoretical Foundation for Performance Measurement and Management', *International Journal of Management Reviews,* 20(3), pp. 653-660.

Bourne, M., Franco-Santos, M., Micheli, P. and Pavlov, A. (2018) 'Performance measurement and management: a system of systems perspective', *International Journal of Production Research,* 56(8), p. 2788.

Bourne, M., Melnyk, S. and Bititci, U. (2018) 'Performance measurement and management: theory and practice', *International Journal of Operations & Production Management,* 38(11), pp. 2010-2021.

Brancheau, J.C. and Wetherbe, J.C. (1987) 'Key Issues in Information Systems Management', *MIS Quarterly,* 11(1), pp. 23-45.

Brandon-Jones, A. (2017) 'E-procurement quality from an internal customer perspective: Construct development, refinement, and replication using a mixed-methods approach.(Report)', *International Journal of Operations & Production Management,* 37(12), pp. 1741-1772.

Brandon-Jones, A. and Silvestro, R. (2010) 'Measuring internal service quality: comparing the gap-based and perceptions-only approaches', *International Journal of Operations & Production Management,* 30(12), pp. 1291-1318.

Caniato, F., Luzzini, D. and Ronchi, S. (2014) 'Purchasing performance management systems: an empirical investigation', *Production Planning & Control,* 25(7), pp. 616-635.

Carter, J.R., Smeltzer, L.R. and Narasimhan, R. (2000) 'Human Resource Management within Purchasing Management: Its Relationship to Total Quality Management Success', *Journal of Supply Chain Management,* 36(1), pp. 52-62.

Cavinato, J.L. (1987) 'Purchasing performance: what makes the magic?', *Journal of Purchasing and Materials Management,* 23(3), pp. 10-16.

Chae, B. (2009) 'Developing key performance indicators for supply chain: an industry perspective', *Supply Chain Management: An International Journal,* 14(6), pp. 422-428.

Chia, A., Goh, M. and Hum, S.-H. (2009) 'Performance measurement in supply chain entities: balanced scorecard perspective', *Benchmarking: An International Journal,* 16(5), pp. 605-620.

Cox, A., Chicksand, D., Ireland, P. and Davies, T. (2005) 'Sourcing Indirect Spend: A Survey of Current Internal and External Strategies for Non-Revenue-Generating Goods and Services', *Journal of Supply Chain Management,* 41(2), pp. 39-51.

Cunha Callado, A.A. and Jack, L. (2015) 'Balanced scorecard metrics and specific supply chain roles', *International Journal of Productivity and Performance Management,* 64(2), pp. 288-300.

Dalkey, N. and Helmer, O. (1963) 'An experimental application of the Delphi method to the use of experts', *Management Science,* 9(3), pp. 458-467.

Davis, T.S. and Novack, R.A. (2012) 'Why metrics matter', *Supply Chain Management Review,* 16(4), pp. 10-17.

de Oliveira, C.A., Carneiro, J. and Esteves, F. (2018) 'Conceptualizing and measuring the "strategy execution" construct', *Journal of Business Research*.

Ellram, L.M., Tate, W.L. and Billington, C. (2004) 'Understanding and Managing the Services Supply Chain', *Journal of Supply Chain Management,* 40(4), pp. 17-32.

Finn, D.W., Baker, J., Marshall, G.W. and Anderson, R. (1996) 'Total quality management and internal customers: measuring internal service quality', *Journal of Marketing Theory & Practice,* 4(3), p. 36.

Franco-Santos, M., Lucianetti, L. and Bourne, M. (2012) 'Contemporary performance measurement systems: A review of their consequences and a framework for research', *Management accounting research,* 23(2), pp. 79-119.

Gamini, G. (2011) 'Reliability of the internal service encounter', *International Journal of Quality & Reliability Management,* 28(9), pp. 1003-1018.

Garengo, P., Biazzo, S. and Bititci, U.S. (2005) 'Performance measurement systems in SMEs: A review for a research agenda', *International Journal of Management Reviews,* 7(1), pp. 25-47.

Gebauer, J. and Segev, A. (2000) 'Emerging technologies to support indirect procurement: two case studies from the petroleum industry', *Information Technology and Management,* 1(1-2), pp. 107-128.

Goodman, C.M. (1987) 'The Delphi technique: a critique', *Journal of Advanced Nursing,* 12(6), pp. 729-734.

Gopal, P.R.C. and Thakkar, J. (2012) 'A review on supply chain performance measures and metrics: 2000-2011', *International Journal of Productivity and Performance Management,* 61(5), pp. 518-547.

Guersola, M., De Lima, E.P. and Teresinha Arns Steiner, M. (2018) 'Supply chain performance measurement: a systematic literature review', *International Journal of Logistics Systems and Management,* 31(1), p. 22.

Gunasekaran, A. and Kobu, B. (2007) 'Performance measures and metrics in logistics and supply chain management: a review of recent literature (1995–2004) for research and applications', *International Journal of Production Research,* 45(12), pp. 2819-2840.

Gunasekaran, A., Patel, C. and McGaughey, R.E. (2004) 'A framework for supply chain performance measurement', *International Journal of Production Economics,* 87(3), pp. 333-347.

Gunasekaran, A., Patel, C. and Tirtiroglu, E. (2001) 'Performance measures and metrics in a supply chain environment', *International Journal of Operations & Production Management,* 21(1), pp. 71-87.

Hasson, F. and Keeney, S. (2011) 'Enhancing rigour in the Delphi technique research', *Technological Forecasting and Social Change,* 78(9), pp. 1695-1704.

Hasson, F., Keeney, S. and McKenna, H. (2000) 'Research guidelines for the Delphi survey technique', *Journal of Advanced Nursing,* 32(4), pp. 1008-1015.

Hofmann, E. (2014) *Performance Measurement and Incentive Systems in Purchasing : More Than Just Savings.* Berlin, Heidelberg: Berlin, Heidelberg : Springer Berlin Heidelberg : Imprint: Springer.

Hübner-Bloder, G. and Ammenwerth, E. (2009) 'Key Performance Indicators to Benchmark Hospital Information Systems—A Delphi Study', *Methods of Information in Medicine,* 48(6), p. 508.

Johnston, R. (2008) 'Internal service – barriers, flows and assessment', *International Journal of Service Industry Management,* 19(2), pp. 210-231.

Kasiri, N., Sharda, R. and Hardgrave, B. (2012) 'A balanced scorecard for item-level RFID in the retail sector: a Delphi study', *European Journal of Information Systems,* 21(3), pp. 255-267.

Keil, M., Tiwana, A. and Bush, A. (2002) 'Reconciling user and project manager perceptions of IT project risk: a Delphi study1', *Information Systems Journal,* 12(2), pp. 103-119.

Kendall, M. and Gibbons, J.D. (1990) 'Rank Correlation Methods Edward Arnold', *A division of Hodder & Stoughton, A Charles Griffin title, London*, pp. 29-50.

Komatina, N., Nestic, S. and Aleksić, A. (2019) 'Analysis of the performance measurement models according to the requirements of the procurement business process', *International Journal of Industrial Engineering and Management*.

Kumar, A., Ozdamar, L. and Ng, C.P. (2005) 'Procurement performance measurement system in the health care industry', *International Journal of Health Care Quality Assurance,* 18(2), pp. 152-166.

Lambert, D.M. and Pohlen, T.L. (2001) 'Supply Chain Metrics', *International Journal of Logistics Management, The,* 12(1), pp. 1-19.

Lapide, L. (2000) 'True measures of supply chain performance', *Supply Chain Management Review,* 4(3), pp. 25-28.

Large, R.O. and König, T. (2009) 'A gap model of purchasing's internal service quality: Concept, case study and internal survey', *Journal of Purchasing and Supply Management,* 15(1), pp. 24-32.

Lavrakas, P. (2008) 'Encyclopedia of Survey Research Methods'.

Loughlin, K.G. and Moore, L.F. (1979) 'Using Delphi to achieve congruent objectives and activities in a pediatrics department', *Academic Medicine,* 54(2), pp. 101-106.

Lynch, R.E. and Cross, K.F. (1991) *Measure up! Yardsticks for continuous improvement.* Blackwell, Cambridge.

Maestrini, V., Luzzini, D., Caniato, F., Maccarrone, P. and Ronchi, S. (2018) 'The impact of supplier performance measurement systems on supplier performance', *International Journal of Operations & Production Management,* 38(11), pp. 2040-2061.

Maestrini, V., Luzzini, D., Maccarrone, P. and Caniato, F. (2017) 'Supply chain performance measurement systems: A systematic review and research agenda', *International Journal of Production Economics,* 183(PA), pp. 299-315.

Mapes, J., New, C. and Szwejczewski, M. (1997) 'Performance trade-offs in manufacturing plants', *International Journal of Operations & Production Management,* 17(10), pp. 1020-1033.

McKenna, H.P. (1994) 'The Delphi technique: a worthwhile research approach for nursing?', *Journal of Advanced Nursing,* 19(6), pp. 1221-1225.

Melnyk, S.A., Bititci, U., Platts, K., Tobias, J. and Andersen, B. (2014) 'Is performance measurement and management fit for the future?', *Management Accounting Research,* 25(2), pp. 173-186.

Micheli, P. and Mura, M. (2017) 'Executing strategy through comprehensive performance measurement systems', *International journal of operations,* 37(4), pp. 423-443.

Mills, A.J., Durepos, G. and Wiebe, E. 2010. Sage encyclopedia of case study research. Thousand Oaks, CA: Sage.

Mills, C. (2017) *Performance Management : A Practical Guide.* Author House UK. p. 274.

Minjoon, J. and Shaohan, C. (2010) 'Examining the relationships between internal service quality and its dimensions, and internal customer satisfaction', *Total Quality Management & Business Excellence,* 21(2), pp. 205-223.

Mishra, D., Gunasekaran, A., Papadopoulos, T. and Dubey, R. (2018) 'Supply chain performance measures and metrics: a bibliometric study', *Benchmarking: An International Journal,* 25(3), pp. 932-967.

Nair, A., Jayaram, J. and Das, A. (2015) 'Strategic purchasing participation, supplier selection, supplier evaluation and purchasing performance', *International Journal of Production Research,* 53(20), pp. 6263-6278.

Neely, A. (2005) 'The evolution of performance measurement research: Developments in the last decade and a research agenda for the next', *International Journal of Operations & Production Management,* 25(12), pp. 1264-1277.

Neely, A., Gregory, M. and Platts, K. (1995) 'Performance measurement system design: A literature review and research agenda', *International Journal of Operations & Production Management,* 15(4), pp. 80-116.

Neely, A., Mills, J., Platts, K., Richards, H., Gregory, M., Bourne, M. and Kennerley, M. (2000) 'Performance measurement system design: developing and testing a process-based approach', *International Journal of Operations & Production Management,* 20(10), pp. 1119-1145.

Nevo, D. and Chan, Y.E. (2007) 'A Delphi study of knowledge management systems: Scope and requirements', *Information & Management,* 44(6), pp. 583-597.

Okoli, C. and Pawlowski, S.D. (2004) 'The Delphi method as a research tool: an example, design considerations and applications', *Information & Management,* 42(1), pp. 15-29.

Okwir, S., Nudurupati, S.S., Ginieis, M. and Angelis, J. (2018) 'Performance Measurement and Management Systems: A Perspective from Complexity Theory', *International Journal of Management Reviews,* 20(3), pp. 731-754.

Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1988) 'SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality', *Journal of Retailing,* 64(1), pp. 12-40.

Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1994) 'Reassessment of Expectations as a Comparison Standard in Measuring Service Quality: Implications for Further Research', *Journal of Marketing,* 58(1), pp. 111-124.

Pohl, M. and Förstl, K. (2011) 'Achieving purchasing competence through purchasing performance measurement system design—A multiple-case study analysis', *Journal of Purchasing and Supply Management,* 17(4), pp. 231-245.

Powell, C. (2003) 'The Delphi technique: myths and realities', *Journal of Advanced Nursing,* 41(4), pp. 376-382.

Pun, K.F. and White, A.S. (2005) 'A performance measurement paradigm for integrating strategy formulation: A review of systems and frameworks', *International journal of management reviews : IJMR,* 7(1), pp. 49-71.

Ravelomanantsoa, M.S., Ducq, Y. and Vallespir, B. (2019) 'A state of the art and comparison of approaches for performance measurement systems definition and design', *International Journal of Production Research,* 57(15-16), pp. 5026-5046.

Rowe, G. and Wright, G. (1999) 'The Delphi technique as a forecasting tool: issues and analysis', *International Journal of Forecasting,* 15(4), pp. 353-375.

Ruel, E., Wagner III, W.E. and Gillespie, B.J. (2015) *The practice of survey research: Theory and applications.* Sage Publications.

Sangwa, N.R. and Sangwan, K.S. (2018) 'Development of an integrated performance measurement framework for lean organizations', *Journal of Manufacturing Technology Management,* 29(1), pp. 41-84.

Schmidt, R.C. (1997) 'Managing Delphi Surveys Using Nonparametric Statistical Techniques*', *Decision Sciences,* 28(3), pp. 763-774.

Shao, J., Moser, R. and Henke, M. (2012) 'Multidimensional supply performance framework: A conceptual development and empirical analysis', *International Journal of Production Economics,* 138(1), pp. 26-34.

Sharma, P., Kong, T.T.C. and Kingshott, R.P.J. (2016) 'Internal service quality as a driver of employee satisfaction, commitment and performance: Exploring the focal role of employee well-being', *Journal of Service Management,* 27(5), pp. 773-797.

Taticchi, P. (2010) *Business Performance Measurement and Management.* Springer.

Taticchi, P. and Balachandran, K.R. (2008) 'Forward performance measurement and management integrated frameworks', *International Journal of Accounting and Information Management,* 16(2), pp. 140-154.

Towill, D.R. (1997) 'The seamless supply chain - the predator's strategic advantage', *International Journal of Technology Management,* 13(1), pp. 37-56.

Tsai, Y.-C. and Cheng, Y.-T. (2012) 'Analyzing key performance indicators (KPIs) for E-commerce and Internet marketing of elderly products: A review', *Archives of Gerontology and Geriatrics,* 55(1), pp. 126-132.

Verweire, K. and Van den Berghe, L. (2004) *Integrated Performance Management: New Hype or New Paradigm?* London: SAGE Publications Ltd.

Von der Gracht, H.A. (2012) 'Consensus measurement in Delphi studies: Review and implications for future quality assurance', *Technological Forecasting and Social Change,* 79(8), pp. 1525-1536.

Wagner, S.M. and Kaufmann, L. (2004) 'Overcoming the main barriers in initiating and using purchasing-BSCs', *Journal of Purchasing and Supply Management,* 10(6), pp. 269-281.

Williams, P.L. and Webb, C. (1994) 'The Delphi technique: a methodological discussion', *Journal of Advanced Nursing,* 19(1), pp. 180-186.

Yeung, J.F.Y., Chan, A.P.C. and Chan, D.W.M. (2009) 'A computerized model for measuring and benchmarking the partnering performance of construction projects', *Automation in Construction,* 18(8), pp. 1099-1113.

Young, J.A. and Varble, D.L. (1997) 'Purchasing's Performance as Seen By Its Internal Customers: A Study in a Service Organization', *Journal of Supply Chain Management,* 33(3), pp. 36-41.

Young, W.H. and Hogben, D. (1978) 'An experimental study of the Delphi technique', *Education Research Perspective,* 5, pp. 57-62.

Zhang, Y. and Wildemuth, B.M. (2009) 'Unstructured interviews', *Applications of social research methods to questions in information and library science*, pp. 222-231.

Zimmermann, F. and Foerstl, K. (2014) 'A Meta-Analysis of the “Purchasing and Supply Management Practice–Performance Link”', *Journal of Supply Chain Management,* 50(3), pp. 37-54.

Table I – List of indicators used in empirical studies on BSC applications in the supply chain field.

BSC perspective	Summary of performance indicators measured (Chia et al., 2009)	Summary of indicators in BSC model (Company 1) (Bigliardi and Bottani, 2010)	Summary of indicators in BSC model (Company 2) (Bigliardi and Bottani, 2010)
Financial	1. Return on investment	1. Information carrying cost	1. Supplier cost saving activities
	2. Gross revenue	2. Supplier cost saving activities	2. Variations against budget
	3. Profit before tax	3. Variations against budget	3. Return on investment
	4. Cost reduction	4. Cost per operation hour	4. Cost per operation hour
	-	5. Return on investment	-
Customer	5. Market share	6. Customer query time	5. Customer query time
	6. Number of customers retained	7. Order lead time	6. Order lead time
	7. Customer satisfaction	8. Distribution lead time	7. Distribution lead time
	-	9. Distribution performance	8. Distribution performance
	-	10. Delivery reliability	9. Delivery reliability
	-	11. Effectiveness of distribution planning schedule	10. Effectiveness of distribution planning schedule
	-	12. Quality of delivery goods	11. Quality of delivery goods
	-	13. Customer perceived value of product	12. Customer perceived value of product
	-	14. Flexibility of service system to meet particular customer needs	13. Flexibility of service system to meet particular customer needs
	-	15. Responsiveness to urgent delivery	14. Responsiveness to urgent delivery
Business	8. Quality of services	16. Purchase order cycle time	15. Accuracy of forecasting techniques
	9. New services implemented per year	17. Effectiveness of master production schedule	16. Purchase order cycle time
	10. On time delivery	18. Supplier rejection rate	17. Planned process cycle time
	11. Waste reduction	19. Total inventory cost	18. Total inventory cost
	-	20. Frequency of delivery	-
Learning and Growth	12. Employee satisfaction	21. Supplier assistance in solving technical problems	19. Order entry method
	13. Employee turnover per year	22. Supplier ability to respond to quality problems	20. Level of information sharing
	14. Number of suggestions implemented per employee yearly	23. Buyer-supplier collaboration in problem solving	-
	15. Money invested in employee training yearly	-	-

Table II - Summary of candidate indicators adapted from the literature.

BSC perspective	(Beamon, 1999, Bigliardi and Bottani, 2010, Brandon-Jones, 2017, Cavinato, 1987, Chia et al., 2009, Cunha Callado and Jack, 2015, Gunasekaran et al., 2001, Gunasekaran and Kobu, 2007, Large and König, 2009, Pohl and Förstl, 2011)
Financial	Variations against budget
	Division's operating cost (annual)
	Cost saving, cost avoidance
Internal Business	Cycle time
	Reliability of cycle time
	Waste reduction
	Lean six sigma
	Effectiveness of procurement plan
Customer	The customer query time
	Number of categories supported
	Customer service and satisfaction. Service quality (Measuring customer perception of service - SERVQUAL)
	Flexibility
	The customer query time
	Service level compared to other divisions
Innovation and Learning	Effective reflection of customer requirements into contracts
	Employee satisfaction
	Employee turnover per year
	Number of suggestions implemented per employee yearly
	Money invested in employee training yearly

Table III – Summary of key findings across all Delphi rounds.

Round 1
Original KPIs: 31 proposed KPIs were included in the questionnaire, soliciting a yes/no response. Additional KPIs: 8 new KPIs were generated in this round. Total number of items going to the next round: 39.
Round 2
Original KPIs: Group opinion shared with the expert panel and revised opinion was solicited on 31 KPIs. Only one individual changed his original response on 3 KPIs. Additional KPIs: A yes/no response was solicited from the expert panel on the 8 additional KPIs. Total number of items going to the next round: 39 (31 for information only).
Round 3
Original KPIs: Finalised group opinion from the previous round was presented here for information only. Additional KPIs: Three out of ten respondents revised their opinions on 4 of the additional KPIs. Eliminated KPIs: 10 KPIs were eliminated from the list based on a 51% threshold. - out of 31 original predefined KPIs, 5 of them were eliminated, i.e. 16%, - out of 8 additionally proposed KPIs, 5 of them were also eliminated, i.e. 62% - no KPI that was proposed by management was eliminated Total number of items going to the next round: 29.
Round 4
Total KPIs: 29 KPIs were presented to the expert panel and importance scoring, based on a 1-5 Likert Scale (1 = not important, 5 = extremely important), was solicited from the expert panel. - 6 KPIs (20.68% of all KPIs), which are distributed across all four BSC perspectives, achieved mean values equal to 4 (= very important) or above, on the Likert scale. o "Cycle time" and "Effectiveness of procurement plans" achieved a mean value score of 4.6, "Effective reflection of customer requirements into contracts" a mean value of 4.4, "Reliability of cycle time" a mean value of 4.3, "Internal customer satisfaction % – SERVQUAL" a mean value of 4.1 and finally "Cost savings" a mean value of 4 o Five out of these six KPIs were linked to literature as well as proposed by management, and half of them fall under the Internal Business perspective - 17 KPIs (58.62% of all KPIs) are between a score of 3 (=important) and below 4 (=very important) and distributed across all the BSC perspectives - 6 KPIs (20.68% of all KPIs) scored less than a mean importance value of 3 (= important) - the most popular perspective (from an importance point of view) for the expert panel, is the Internal Business one followed by the Customer perspective, then the Innovation and Learning perspective and finally the Financial one - all KPIs that fall between an importance level of "very important" to "extremely important", and the majority of KPIs above an importance level of "important", have a link to the literature - Kendall's coefficient of concordance for ranks (W) - Kendall W = 0.3207 (aggregate value) ($p=<0.0001$). - KPIs with 60% yes' in the 3rd round, collected an average importance value ≤ 3 in the 4th round of Delphi.

Table IV– First round Delphi study – Agreement with thirty-one suggested KPIs that should be measured.

BSC Perspective	Internal KPIs for measuring the performance of Contracting	n	%
Financial *How do we look to senior management?*	% of competitive bidding versus sole source bidding ($)	10	100%
	Managed spend per fiscal year ($)	9	90%
	Cost benefit (cost savings + cost avoidance)	8	80%
	Cost savings	8	80%
	Average $ value of commitment per contracting representative / FY	8	80%
	Cost avoidance	7	70%
	Division's operating cost (annual)	6	60%
	Cost benefit / division's operating cost	6	60%
	% of spend on LFC vs. % of spend on SFC/SCC	6	60%
	Number of transactions over $1M during the fiscal year	6	60%
	Variation of cumulative signed contract values against budget / FY	5	50%
Internal Business *What must we excel at?*	Effectiveness of procurement plans	10	100%
	Ratio of completed contracts during the fiscal year to the number of staff	10	100%
	Number of claims/contractual disputes per FY	10	100%
	Cycle time	9	90%
	Reliability of cycle time	8	80%
	Clarity of processes and procedures	8	80%
	Usability of procurement tools	7	70%
	Waste reduction	4	40%
	Lean six sigma	2	20%
Customer *How do internal customers see us?*	Internal customer satisfaction % - SERVQUAL	9	90%
	The customer query time	8	80%
	Flexibility	7	70%
	Service level compared to other divisions	7	70%
	Number of categories supported	3	30%
Innovation and Learning *Can we continue to improve and create value?*	Effective reflection of customer requirements into contracts	9	90%
	Employee turnover per year	8	80%
	Employee satisfaction	7	70%
	% of qualified personnel within the team	7	70%
	Number of suggestions implemented per employee yearly	6	60%
	Money invested in employee training yearly	6	60%
	Total number of yes':	224	

Table V – Fourth round Delphi study – Likert Scale scores for 29 indicators for measuring performance.

BSC Perspective	KPI no.	Internal KPIs for measuring the performance of Contracting	n (%)					Importance values 1 to 5 (1 = not important, 2 = slightly important, 3 = important, 4 = very important and 5 = extremely important)				Linked to literature	Proposed by
			Not important (%)	Slightly important (%)	Important (%)	Very important (%)	Extremely important (%)	Mean	Median	Mode (Frequency)	Range		
Financial How do we look to senior management?	1	**Cost savings**	**1 (10%)**	-	**2 (20%)**	**2 (20%)**	**5 (50%)**	**4**	**4.5**	**5 (5)**	**1 to 5**	x	x
	2	Cost benefit (cost savings + cost avoidance)	1 (10%)	1 (10%)	-	5 (50%)	3 (30%)	3.8	4	4 (5)	1 to 5	x	x
	3	Cost avoidance	1 (10%)	1 (10%)	2 (20%)	3 (30%)	3 (30%)	3.6	4	4 (3)	1 to 5	x	x
	4	% of competitive bidding versus sole source bidding ($)	1 (10%)	-	4 (40%)	4 (40%)	1 (10%)	3.4	3.5	4 (4)	1 to 5		x
	5	Average $ value of commitment per contracting representative / FY	1 (10%)	1 (10%)	3 (30%)	5 (50%)	-	3.2	3.5	4 (5)	1 to 4		x
	6	Managed spend per fiscal year ($)	2 (20%)	1 (10%)	4 (40%)	-	3 (30%)	3.1	3	3 (4)	1 to 5		x
	7	Division's operating cost (annual)	2 (20%)	1 (10%)	4 (40%)	1 (10%)	2 (20%)	3	3	3 (4)	1 to 5	x	x
	8	% of annual (FY) contractual commitment growth compared to the previous FY	2 (20%)	1 (10%)	4 (40%)	3 (30%)	-	2.8	3	3 (4)	1 to 4		
	9	% of spend on LFC vs. % of spend on SFC/SCC	2 (20%)	2 (20%)	3 (30%)	3 (30%)	-	2.7	3	3 (3)	1 to 4		x
	10	Cost benefit / division's operating cost	3 (30%)	2 (20%)	3 (30%)	-	2 (20%)	2.6	2.5	3 (3)	1 to 5		x
	11	Number of transactions over $1M during the fiscal year	1 (10%)	3 (30%)	5 (50%)	1 (10%)	-	2.6	3	3 (5)	1 to 4		x
Internal Business What must we excel at?	12	**Cycle time**	-	-	-	**4 (40%)**	**6 (60%)**	**4.6**	**5**	**5 (6)**	**4 to 5**	x	x
	13	**Effectiveness of procurement plans**	-	-	-	**4 (40%)**	**6 (60%)**	**4.6**	**5**	**5 (6)**	**4 to 5**	x	x
	14	**Reliability of cycle time**	**1 (10%)**	-	-	**3 (30%)**	**6 (60%)**	**4.3**	**5**	**5 (6)**	**1 to 5**	x	x
	15	Usability of procurement tools	-	2 (20%)	1 (10%)	4 (40%)	3 (30%)	3.8	4	4 (4)	2 to 5		
	16	% of contractors that are evaluated at least twice a year	1 (10%)	1 (10%)	2 (20%)	2 (20%)	4 (40%)	3.7	4	5 (4)	1 to 5		
	17	Clarity of processes and procedures	-	1 (10%)	5 (50%)	1 (10%)	3 (30%)	3.6	3	3 (5)	2 to 5		
	18	Number of claims/contractual disputes per FY	-	1 (10%)	3 (30%)	5 (50%)	1 (10%)	3.6	4	4 (5)	2 to 5		x
	19	Ratio of completed contracts during the fiscal year to the number of staff	-	2 (20%)	3 (30%)	3 (30%)	2 (20%)	3.5	3.5	4 (3)	2 to 5		x
	20	Number of signed contracts using the contractor's template or using a KAUST template with significant changes	2 (20%)	4 (40%)	3 (30%)	1 (10%)	-	2.3	2	2 (4)	1 to 4		
Customer How do internal customers see us?	21	**Internal customer satisfaction % - SERVQUAL**	-	-	**2 (20%)**	**5 (50%)**	**3 (30%)**	**4.1**	**4**	**4 (5)**	**3 to 5**	x	x
	22	The customer query time	-	-	4 (40%)	3 (30%)	3 (30%)	3.9	4	3 (4)	3 to 5	x	
	23	Service level compared to other divisions	-	2 (20%)	3 (30%)	3 (30%)	2 (20%)	3.5	3.5	4 (3)	2 to 5	x	
	24	Flexibility	-	3 (30%)	2 (20%)	3 (30%)	2 (20%)	3.4	3.5	2 (3)	2 to 5	x	x
Innovation and Learning Can we continue to improve and create value?	25	**Effective reflection of customer requirements into contracts**	-	-	**2 (20%)**	**2 (20%)**	**6 (60%)**	**4.4**	**5**	**5 (6)**	**3 to 5**	x	
	26	Employee satisfaction	-	-	4 (40%)	3 (30%)	3 (30%)	3.9	4	3 (4)	3 to 5	x	
	27	Employee turnover per year	-	1 (10%)	4 (40%)	-	5 (50%)	3.9	4	5 (5)	2 to 5	x	
	28	% of qualified personnel within the team	-	1 (10%)	6 (60%)	1 (10%)	2 (20%)	3.4	3	3 (6)	2 to 5		x
	29	Number of suggestions implemented per employee yearly	3 (30%)	3 (30%)	1 (10%)	3 (30%)	-	2.4	2	1 (3)	1 to 4	x	

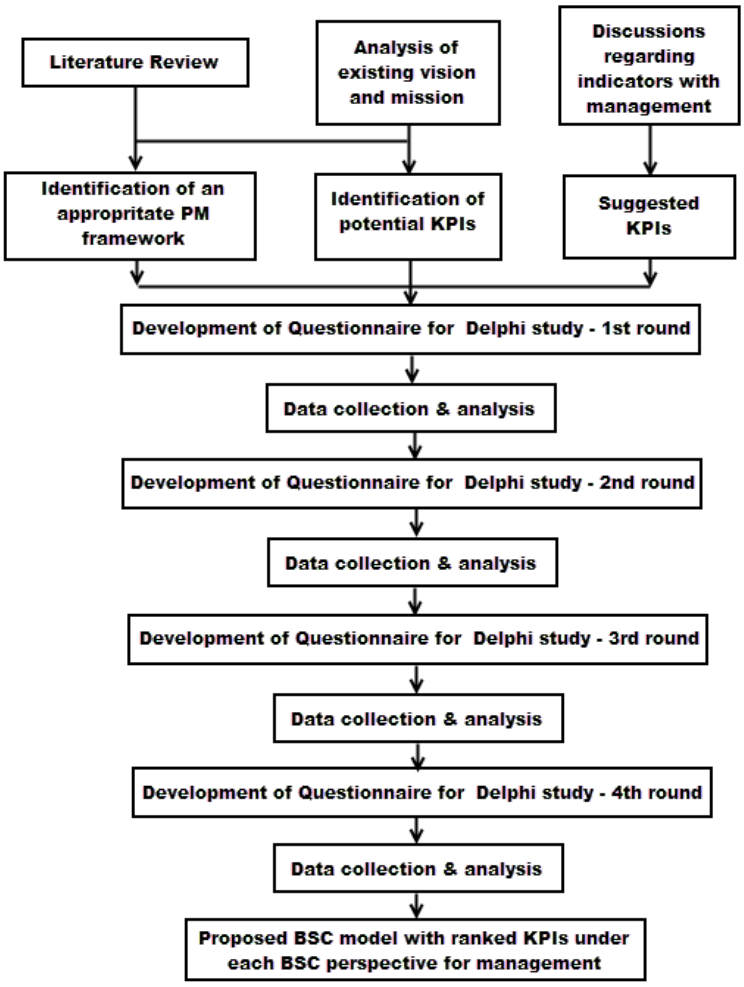

Figure 1 – Research Approach.

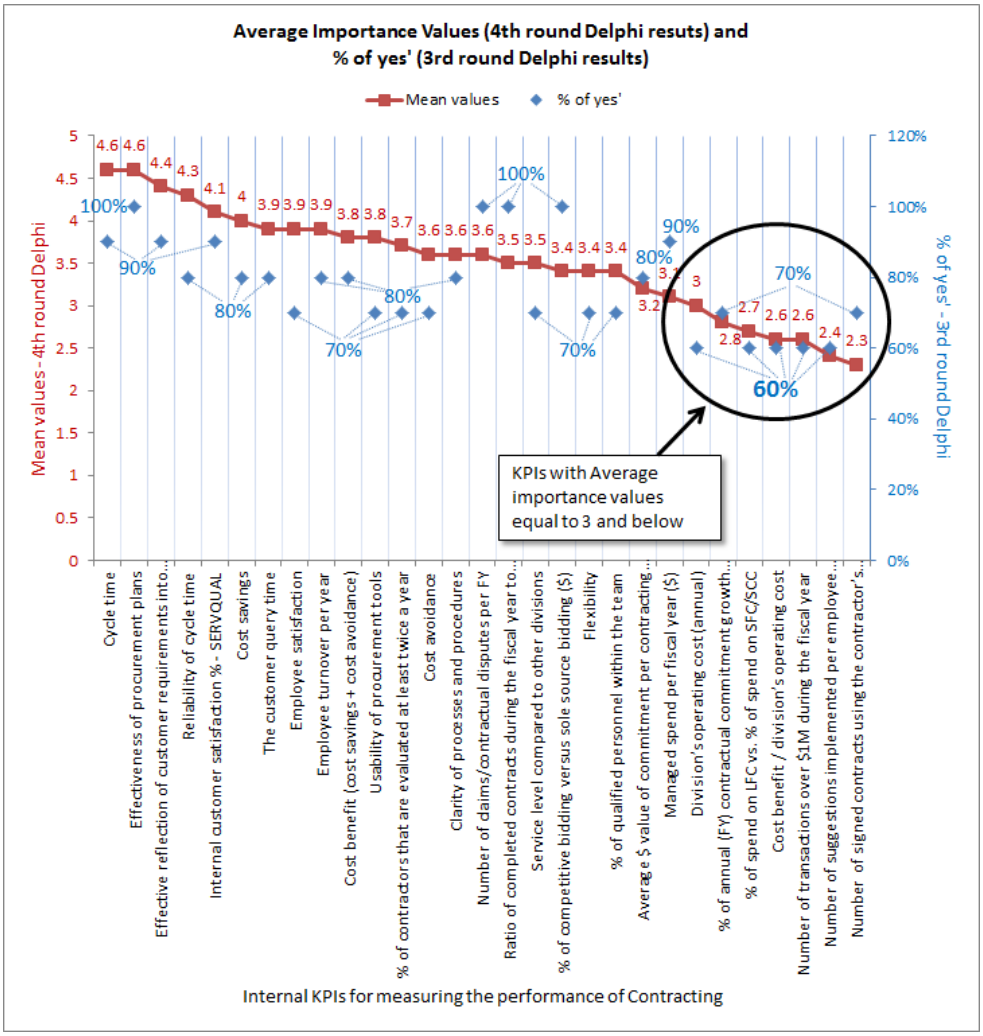

Figure 2 – Comparison of Third and Fourth round Delphi study

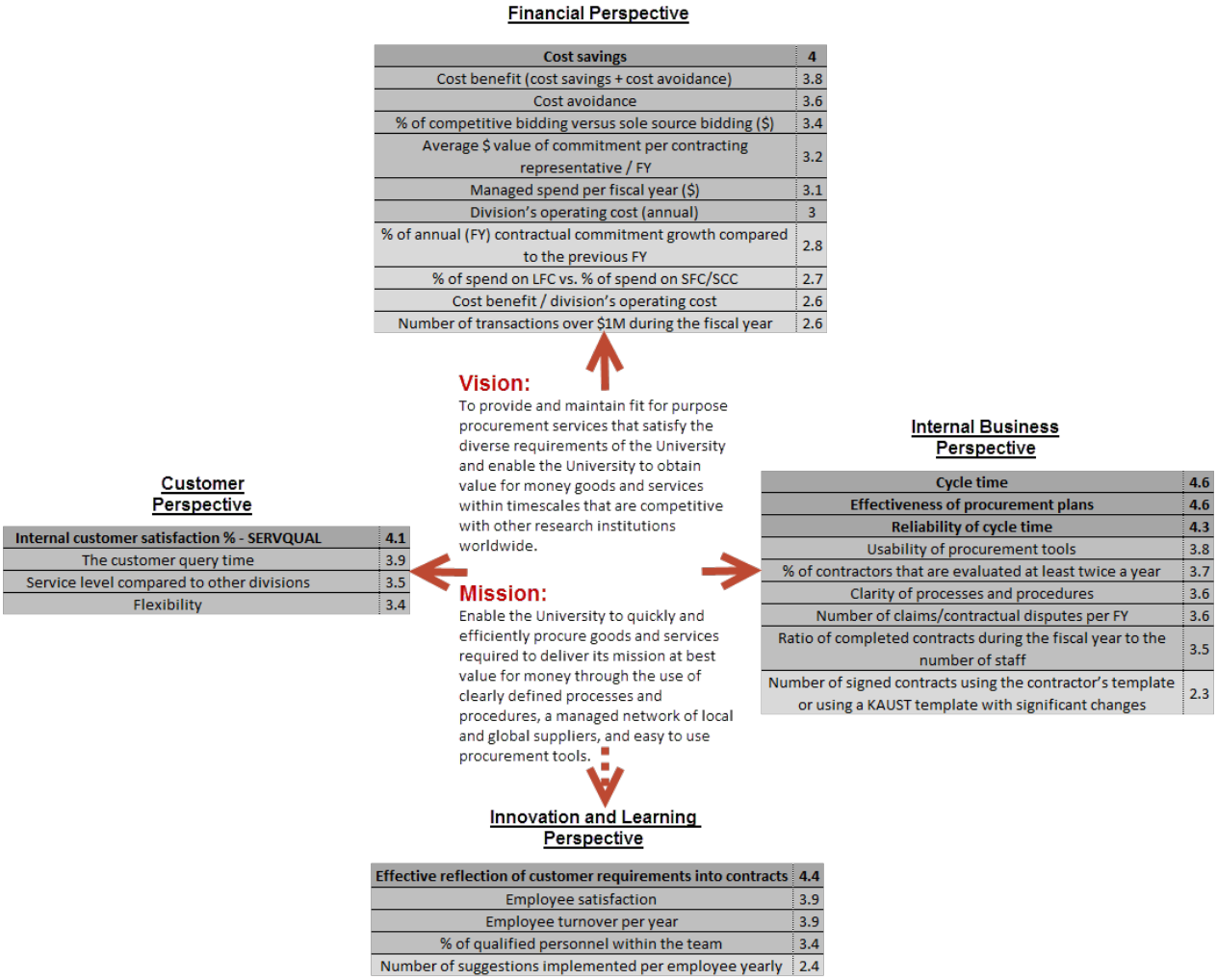

Figure Error! No text of specified style in document. **– The developed BSC model[1].**

[1] (Dark grey: KPIs with importance values ≥ 4; Lighter grey: KPIs with importance values ≥ 3 but < 4; Lightest grey: KPIs with importance values < 3. Red dotted arrow shows an implicit link.)