OPERATIONS AND SUPPLY CHAIN MANAGEMENT
Vol. 16, No. 3, 2023, pp. 399 - 412
ISSN 1979-3561 | EISSN 2759-9363

Evolution of Performance Measurement Research: An Update on Research Development from 2005 to 2020 and Future Outlook for the Field

Hanan Alhaddi
Hult International Business School, USA
Email: hanan.alhaddi@faculty.hult.edu

ABSTRACT

Research in performance measurement has been growing rapidly over the past seven decades, and it has explored a wide spectrum of issues in multiple industries and sectors, from assessing airports' efficiency to proposing frameworks to improve supply chain management. The field has also attracted the attention of scholars to conduct several literature reviews to understand the evolution of the research in the field, given its multidisciplinary nature. Given the sustained academic interest in this field and lack of consensus on its maturity, this paper uses co-occurrence, citation, and co-citation analyses to examine the field's research development and academic maturity. Findings from this study inform the performance measurement literature in three ways; by highlighting the absence of adequate examination of effectiveness, an essential complement to the extensively studied efficiency in the current literature; by indicating 'practical implementation' as the current stage of research evolution of the field; and by emphasizing the need to adopt novel theoretical perspectives. It also extends the literature review previously conducted by other scholars (from 1950 to 2005) to update the research community on developments in the field from 2005 to 2020. This paper argues that while this research area has shown signs of academic maturity and professionalism, there seems to be continued reliance on a limited number of works despite the field entering a practical implementation phase. Finally, this study proposes a path for empirical verification for some of the key theoretical foundations found in the literature.

Keywords: *Performance measurement, bibliographic data, co-occurrence, co-citation, literature review, research evolution*

1. BACKGROUND

Over the course of roughly seven decades (from 1950s to 2020), the field of performance measurement has been examined by scholars from a wide range of multidisciplinary fields, posing a variety of questions centered around the desire to quantify measurements (Ridgway, 1956), understanding the implications of measurements (Argyris, 1952; Power, 1997), identifying a set of "balanced" measures (Drucker, 1954; Johnson and Kaplan, 1987), and developing robust systems for design and implementation of measurement systems (Bourne *et al*., 2000; Neely *et al*. 1999; Neely *et al*., 2000). To guide research activities, provide empirical evidence to support various propositions, and attempt to address some of the legacy questions in the field, a large number of frameworks have been developed, with some being evidently more predominant than others, particularly revolving around the rediscovery of Drucker's suggestion related to the need for developing a balanced measurement system (Drucker, 1954). For example, among these works were the performance measurement matrix (Keegan *et al*., 1989), the results-determinants framework (Fitzgerald *et al*., 1991), the performance pyramid (Lynch and Cross, 1991), the balanced scorecard (Kaplan and Norton, 1992), auditing and enhancing performance measurement systems framework (Medori and Steeple, 2000), the performance prism (Neely, 2001; Neely *et al*., 2002), forward performance measurement and management integrated framework (Taticchi *et al*., 2008a), a framework to review performance measurement systems (Najmi *et al*., 2005), examining the effects of performance measurement on performance framework (Pavlov and Bourne, 2011), and theoretical framework for the interfaces with the performance measurement and management system (Melnyk *et al*., 2014).

In 2005, Andy Neely published an update on performance measurement research based on a comprehensive literature review (Neely, 2005); the update was to a paper originally published in 1995 where the author, along with Gregory and Platts have previously conducted a literature review and provided thoughtful research agenda considerations relative to performance measurement system design (Neely *et al*., 1995). A decade of research progress in the field has revealed less than exciting results, as Neely noted the overreliance in much of the literature on only a small number of works (predominantly, the balanced scorecard by Kaplan) despite the wide spectrum of disciplines concerned with performance measurement. Additionally, the field did not appear to have professionalized academically as of 2005. The paper also presented several proposed research questions to be considered for future research agenda and concluded, with clear articulation, that there is a practical concern should the overreliance in the literature on a singular framework continues. The challenge namely was running the risk of "becoming trapped by solutions proposed for problems of the past" (Neely, 2005). More specifically, the work carved out two paths of inquiries: one aimed at the academic professionalism and maturity of the field, and one aimed at the research community providing a set of proposed research questions to advance the agenda. Concerned with the evolution and academic professionalism of performance measurement as a

field and realizing the veracity of the challenges articulated, this paper reports and discusses the results of quantitative and comprehensive literature on performance measurement by extending the review to 2020 (from 2015 to 2020). Comparative remarks to Neely's work (2005) are made throughout this paper. This is crucial as it enables the illustration of the field's dynamic evolution and informs the proposed avenues for future research. Additionally, it sets the stage for providing informed responses to the prevalent questions posed by the author in 2005. This study has five main aims: (1) to understand the intellectual structure of current literature, (2) to explain how the literature has evolved, (3) to investigate whether and how the evolution in the research of performance measurement could contribute to the overall development of the field, (4) to provide informed responses to the enduring questions posed by the most recent literature review by Neely (2005), and (5) to propose an informed agenda for future research.

In a thorough examination of the evolution of performance measurement research, the remaining sections of this paper are organized as follows. Firstly, the research design is presented, describing the bibliometric method used for the analysis. Secondly, findings are discussed, including descriptive evidence of the bibliographic data used to illustrate the evolution of the research field. Thirdly, Findings are synthesized to propose contributions to the performance measurement literature and provide informed responses to the prevalent questions from Neely's most recent and comprehensive literature review (2005). Lastly, insights are drawn to provide a future outlook for the field and propose avenues for future research.

2. PERFORMANCE MEASUREMENT: ENDURING AND EVOLVING RESEARCH QUESTIONS

In 2005, Neely conducted the most comprehensive literature review that conceptualized the research progress of performance measurement. The study carved out two paths of inquiry: one aimed at the academic professionalism and maturity of the field, and one aimed at the research community providing a set of proposed research questions to advance the agenda. Concerning academic professionalism and maturity, performance measurement as a field has matured academically. Despite the interdisciplinary nature of the area, there was an overreliance on a small set of works (e.g., the balanced scorecard). The author questioned the appropriateness of outdated frameworks in solving future problems should the trend continues (in other words, using yesterday's solutions to solve today's problems). Lastly, Neely has posed a conceptual inquiry around whether the field had a future in research. In terms of the research community, the set of research questions proposed revolved around 1) the deployment and implementation of performance measurement systems, 2) the alignment of performance measurement systems with the organization's strategy, and 3) understanding the implications of the extended enterprise (i.e., supply chain network and other key stakeholders). **Table 1** below summarizes the two paths of inquiry.

Table 1 Questions from Neely's 2005 update

Questions related to academic professionalism and maturity	Questions associated with advancing research agenda for the research community
Why hasn't the performance measurement field professionalized and matured academically?	How to design and deploy enterprise performance management rather than measurement systems?
Why does the performance measurement research community depend on a limited number of works from a limited number of contributors?	How to measure performance across supply chains and networks rather than within organizations?
Has the field had its day?	How to measure intangible as well as tangible assets for external disclosure as well as internal management?
Why haven't new dominant ideas or breakthroughs emerged since 1995? Is there a deeper-rooted problem?	How to develop dynamic rather than static measurement systems?
	How can the flexibility of measurement systems be enhanced to cope with organizational changes?

More recently, Cvetkoska and Eftimov (2021) have conducted a study to analyze bibliographic data on performance measurement articles published from 1978 to 2019 in the Scopus database. Despite the similarity in the approach, this paper specifically addressed the enduring research questions posed by Neely (2005). It advanced the Evolutionary Cycle of the Performance Measurement Research framework introduced in Neely's study by adding a sixth phase, Practical Implementation, as the current stage of performance measurement research in the literature.

In a thorough examination of the evolution of performance measurement research, the remaining sections of

this paper 1) provide an update on the citation analysis methodology and data presentation of relevant data, 2) synthesize findings for an update on the research evolution, 3) address the general questions from Neely's 2005 update, and 4) draw insights for an outlook on future performance measurement research. Due to the frequency of referencing Neely's 2005 paper in this study, his work will be referred to as "Neely's 2005 update" for consistency for the remainder of this paper.

3. CITATION ANALYSIS: UPDATE ON THE METHODOLOGY

Following a similar citation/co-citation analysis method used in Neely's 2005 update, this paper leveraged the Web of Science (WoS) database to provide empirical evidence underpinning the development of performance measurement research. The initial search included all publications from 2005 to 2020 that contained the phrase "performance measurement" in the titles, abstracts, or authors' keywords.

The initial search returned a total of 10,254 results. There were three adjustment rounds to refine the results and to focus on the most relevant research works. The first round was to include only "Articles" and "Proceeding Papers" in the search; all other types of publications were excluded. The second round was to categorize the selected documents and rank the top 25 categories, which covered a broad spectrum of disciplines from management and engineering to medicine and religion. The third round was to exclude any category containing less than 10% of records, resulting in a final count of 1,943 articles in the dataset. The remaining five categories were: management (45%), business finance (19%), business (17%), operations research (11%), and industrial engineering (8%). The management category contained the highest number of publications (874 papers), followed by business finance with 60% fewer papers (369 papers). An important observation was the lessening diversity in the performance measurement space, indicating a slightly less widely distributed field of academic study than Neely's 2005 update.

Regarding publication activity, the papers included in the dataset (up to June 2020) showed a steadily increasing trend overall, with a noticeable jump in 2015 (210 publications) from 2014 (112 publications). This finding suggests an answer to the inquiry by Neely in 2005 around whether the field has had its day: the continued academic momentum argues no; the research area seems to continue to attract the attention of the academic research community.

At a high level, the 1,943 papers included in the dataset provided some 21,740 citations, covering 40,143 works and drawing on 40,143 different authors. These papers were cited 24,778 times (22,081 without self-citations) from 2005 to 2020. Similar to the publication activity, there was a steadily increasing trend in the number of citations. Due to the difference in the durations between Neely's study and this paper, providing a basic comparison of the paper count, the number of authors, etc., would result in a misinterpretation of the data. While a "normalization" of the data was possible, the evident jump in research activity from 2015 to 2020 would have skewed the results. Therefore, commenting on the updates as a percentage (where applicable) was a more robust method to draw insights. In this regard, the significant change in the number of authors per publication is worth noting. The 1,352 papers included in Neely's 2005 update dataset drew on 16,697 authors (roughly 12 authors per paper), whereas the 1,943 papers included in this dataset drew on 40,143 authors (approximately 20 authors per paper). One explanation is the enduring tendency in the academic community to publish, an argument presented by Neely's 2005 update that seems to stand 15 years later. Another note is that authors with only 1 citation continued to make up the vast majority of the contributing authors to the papers included in both datasets: 11,929 of the total 16,697 (71.4 percent) and 26,451 of the total 40,143 (66 percent) in 2005 and 2020 respectively.

In this study, the five top papers in terms of citations were Neely's 2005 update on performance measurement research 2005 (Neely, 2005) at 326 citations; followed by (Zhou *et al.*, 2007) at 276 citations; (Bitici *et al.*, 2012) at 240 citations; (Liang *et al.*, 2006) at 237 citations; then (Franco-Santos *et al.*, 2007) at 215 citations. In comparison, the five top papers in terms of citations from Neely's 2005 update were (Kaplan *et al.*, 1992) at 119 citations; (Kaplan *et al.* 1996a) at 63 citations; (Charnes *et al.*, 1978) at 56 citations; (Dixon *et al.*, 1990) at 49 citations; then (Neely *et al.*, 1995) at 42 citations. Aside from the increase in the number of citations in both datasets, two additional vital insights to draw here are: Firstly, Neely's work on performance measurement research (Neely, 2005; Neely *et al.*, 1995) continues to be among the five top papers as of 2005 and 2020. This observation illustrates the ongoing interest of the research community in examining the field's evolution. Secondly, the observed reduction in the citing of the balanced scorecard demonstrates the lessening reliance on this concept.

Regarding authors' contribution, 4,244 authors contributed to the papers included in the dataset (the highest publication record for an author was 14, and the lowest was one). Out of the total number of contributing authors, 537 authors (12.6%) had two or more publications, 153 authors (3.6%) had three or more publications, 57 authors (1.3%) had four or more publications, and only 26 authors (0.02%) had five or more publications. The majority of the remaining 3,707 authors (87.3%) had only one publication. Neely's 2005 update noted a similar trend. All papers included in the dataset were downloaded from the Web of Science database and then uploaded using VOSviewer (a software tool used for constructing and visualizing bibliometric networks developed by Van Eck and Waltman in 2007). VOSviewer enables the construction of networks based on 1) bibliographic data and 2) text data. Bibliographic data draw on various elements based on co-occurrence, citation, bibliographic coupling, and co-citation, where text data enables the construction of "Term Maps."

The remaining analysis section of this paper includes 1) a bibliographic data-based analysis section where keyword co-occurrence, citation, and co-citation are analyzed and 2) a text

data-based analysis section where a term map is constructed and discussed.

4. BIBLIOGRAPHIC DATA-BASED ANALYSIS

Keyword Co-occurrence Analysis. This analysis looked at the relatedness of keywords (combining Author Keywords and Key Words Plus (terms appearing in the article's references but not the article itself)) based on the number of publications in which they occur together. With a threshold set at a minimum of 5 occurrences, out of the total 6,262 keywords, 562 met the criteria. The most frequently occurring keywords were "performance measurement or performance-measurement" (995 occurrences, combined), "management" (334 occurrences), "balanced scorecard" (268 occurrences), and "framework" (204 occurrences). It was necessary to narrow down the keyword set to further analyze the relatedness of the co-occurred terms. Several iterations were done in which settings were adjusted by increasing the number of occurrences per keyword from 5 to 100 in increments of five (i.e., 5, 10, 15, etc.). The Total Link Strength (TLS) score was selected as an attribute metric. TLS is a standard weight metric that indicates the number of links associating two elements; it is essential because it provides another dimension to evaluate the connections among searched terms. The TLS is calculated by summing the number of links connecting/relating one term to another within the dataset included in the study. **Table 2** shows a summary of the findings.

Table 2 Keyword co-occurrence analysis

Keyword	Occurrences	Total Link Strength
Performance measurement	658	3347
Performance-measurement	337	2146
Management	334	2120
Balanced scorecard	268	1633
Framework	204	1425
Performance	199	1061
Impact	188	1433
Model	162	1145
Strategy	147	1072
Design	140	982

The search included two variations of the phrase performance measurement, an unhyphenated and a hyphenated version (i.e., performance-measurement). While it was an option to exclude the hyphenated version (337 occurrences) and only keep the unhyphenated version (658 occurrences), doing so would have caused an overlook of roughly 50% more linkages. **Figure 1** shows the revised network, highlighting three nodes around performance measurement, balanced scorecard, and management. In general, there were strong connections among these keywords.

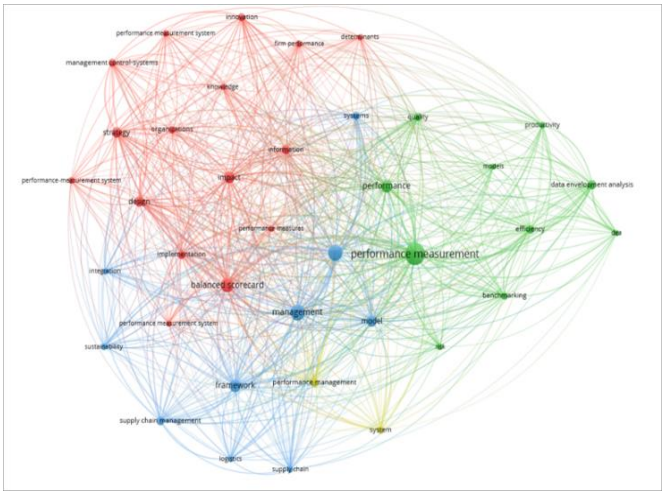

Figure 1 Keyword co-occurrence network

Worthy of note, there was a strong link between performance measurement and keywords like efficiency, models, and productivity. At the same time, the balanced scorecard showed a strong link to keywords like determinants and management control systems. Based on the network analysis, there are two main takeaways: 1) the enduring keyword co-occurrence among performance measurement, management, and balanced scorecard, and 2) the unclear delineation among the second tier of keywords (smaller nodes). In other words, multiple duplications and overlaps of keywords were observed in each node. For example, the "performance measurement system" (or "performance-measurement system") seems to appear in the literature on the balanced scorecard, management, and performance measurement (the three primary nodes with the most linkages).

Citation Analysis. The citation analysis was conducted three times using the author, source, and publication (document) as the unit of analysis each time. The most frequently cited authors were: Garengo (512 citations), Bititci (435 citations), Widener (409 citations), and Neely (359 citations). Interestingly, there was not correlation between the number of publications and the number of citations per author. The author with the most publications (Saunila, 15) was cited only 102 times. The top ten most frequently cited authors were from less diverse disciplinary backgrounds (compared to Neely's 2005 work). A total of 20,039 sources were identified. The top 5 most cited sources were *International Journal of Operations and Production Management* (2,323 citations), *Management Accounting Research* (1,891 citations), *Supply Chain Management - An International Journal* 1,543 citations), *Accounting, Organizations and Society* (1,307 citations), and *Omega - International Journal of Management Science* (1,208 citations). Combined, they make up roughly 41% of the total citations. **Table 3** shows the top 10 most frequently cited sources.

Table 3 Citation analysis by source

Source	Citations
International Journal of Operations and Production Management	2323
Management Accounting Research	1891
Supply Chain Management - An International Journal	1534
Accounting, Organizations and Society	1307
Omega - International Journal of Management Science	1208
European Journal of Operational Research	961
Journal of the Operational Research Society	735
International Journal of Project Management	679
Total Quality Management and Business Excellence	654
Accounting Review	627

Of the 1,943 publications, Neely's 2005 update on performance measurement research was the most cited (326 citations). A dominant presence of operations management, management, and supply chain management fields was observed, academic disciplines that are naturally closely related. This observation supports the argument made earlier in this paper regarding the slightly less diverse field of performance measurement over the past 15 years. **Table 4** below lists the top 10 most frequently cited papers.

A few critical remarks are worth noting, particularly when compared to Neely's 2005 update. Firstly, when looking annually, the rate of publication over the past 1.5 decades has been relatively stable (roughly 135 and 134 papers published per year as of 2005 and 134, respectively). One note is the adjustment made for 2020 (removed 66 papers from the complete set) to account for only full years of publication from 2005 to 2014. Secondly, there has been less overreliance in the literature on the work of any singular author (compared to the observed dominance of Kaplan and Norton in Neely's 2005 update). Thirdly, the significant increase in the number of citations per paper indicates a field that has matured academically and professionally over time. For example, Neely's 2005 update showed that less than 1 percent of the papers included in the 2005 dataset were cited more than 30 times (10 out of 1,352 papers). Academic maturity is evident as this statistic grew to roughly 10 percent in 2020 (189 out of 1,943 papers).

Table 4 Citation analysis by document

Publication	Publication Year	Total Citations	Field/Discipline
Neely (2005)	2005	326	Operations Management
Zhou *et al.* (2007)	2007	276	Environmental Performance
Bititci *et al.* (2011)	2012	240	Operations Management
Liang *et al.,* (2006)	2006	237	Supply Chain
Franco-Santos *et al.,* (2007)	2007	215	Management
Toor *et al.,* (2010)	2010	194	Management
Barros *et al.,* (2012)	2012	171	Management
Yang *et al.* (2009)	2009	171	Operations Management
Meng (2012)	2012	166	Management
Leuschner *et al.,* (2013)	2013	165	Supply Chain Management

Co-citation Analysis. This analysis was conducted using cited authors, cited sources, and cited references as the unit of analysis each time. Regarding the author as the unit of study, the TLS score has also been selected as a weight metric. In this context, TLS indicates the number of links associating an author with other authors. There were 40,143 authors included in the dataset.

Looking at **Table 5** below (showing the top 10 authors based on TLS score), Kaplan had the highest TLS score (33,134), followed by Chenhall (21,727), Neely (21,049), Ittner (17,243), then Simons (13,393). One vital insight to draw here is the endurance of the balanced scorecard as the basis for the relatedness to other authors' works. This insight presents an interesting quandary as the overreliance on the works by Kaplan and Norton, in particular, has decreased. One possible explanation is that despite the absence of significant new publications by Kaplan and Norton (i.e., balanced scorecard), the work of these authors continues to be influential and relevant to other streams of research in performance measurement. This is expected as the purpose of the balanced scorecard is to serve as *a* performance measurement tool. Another remark is the direct relationship between the TLS score and the Citations. In other words, except for a couple of deviations, the higher the TLS score, the higher the citations.

Constructing the network analysis required two iterations, each at a different author citation threshold setting. The first was setting the minimum threshold at 20 citations per author, which yielded a significantly dense network. The

second iteration was done by reducing the minimum threshold to 50 citations per author, which generated the network shown in **Figure 2**. As a general practice, the default settings were used first and then adjusted based on the density of the network and the difficulty (or easiness) of extracting valuable insights from it. Looking at the initial network, an essential understanding to draw is the distance between the nodes. For example, while Kaplan and Neely seem to have strong linkages to other authors (based on the number of links extruded from each node), the distance between the two appeared to be the closest since both works focus on performance measurement.

Table 5 Co-citation analysis by author

Author	Citations	Total Link Strength
Kaplan	1367	33134
Chenhall	504	21727
Neely	826	21049
Ittner	498	17243
Simons	341	13393
Abernethy	222	11515
Bourne	283	8749
Merchant	159	7245
Henri	153	6871
Gunasekaran	312	6784

With a less dense network, a closeup of the links was possible. **Figure 2** shows that while Kaplan (accounting) was more connected to Ittner (accounting) and Chenhall (accounting), Neely (operations management) was more related to Gunasekaran (operations management). Further, while performance measurement was a central theme, it was connected to similar disciplines from 2005. For example, Ittner (2015) discussed the primary organizational and technical issues organizations face as they implement performance measurement systems. Chenhall *et al.*, (2005) focused on the strategic alignment between performance measurement systems and manufacturing.

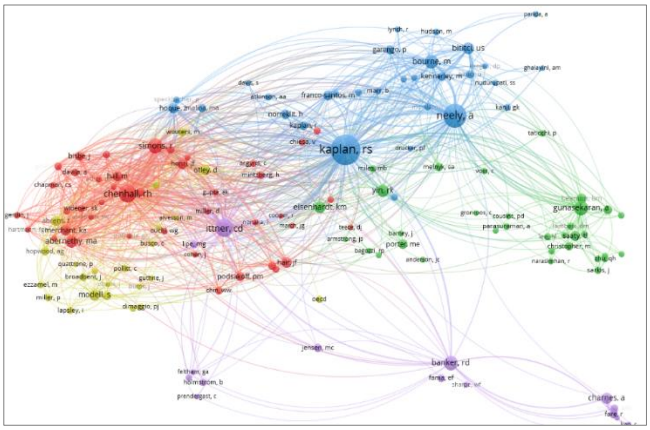

Figure 2 Co-citation by author network

Regarding cited sources as the unit of analysis, two iterations were done to refine the results. The first was to set the minimum threshold for the number of citations per source to 20. Of the 21741, 564 sources met were cited more than 20 times. There was also a direct relationship between the TLS score and the number of citations per source. A few key observations are worth noting. Firstly, *Accounting, Organizations, and Society* and *The International Journal of Operations and Production Management* show strong relatedness (indicated by the high TLS score) to other sources in the performance measurement field. Secondly, both journals are linked heavily to *Harvard Business Review.* One possible explanation is sharing some common grounds relative to the balanced scorecard (heavily published by *Harvard Business* Review). Thirdly, in general, *Harvard Business Review* seems to have less influence on the performance measurement literature in 2020 compared to data from Neely's 2005 update, which is logical as the same has been observed for the balanced scorecard. **Table 6** below shows the top 10 sources by co-citation.

Table 6 Co-citation analysis by source

Source	Citations	Total Link Strength
Accounting, Organizations, and Society	4064	202265
International Journal of Operations and Production Management	2798	134749
International Journal of Production Economics	1506	79158
Management Accounting Research	1159	73742
Management Accounting	1336	72511
Journal of Operations Management	1142	67516
Harvard Business Review	1414	56727
Academy of Management Journal	897	56471
Strategic Management Journal	996	55936
International Journal of Production Research	829	53562

As alluded to earlier, the initial network included 564 cited sources; hence, it was difficult to draw any meaningful insights due to its density. This observation informed the second iteration to increase the threshold to 100, resulting in 148 sources cited (see **Figure 3** below).

The predominant nodes showed the most co-cited sources: the Accounting, Organizations, and Society, International Journal of Operations and Production Management, Journal of Operations Management, Harvard Business Review, and European Journal of Operations Research. Worthy of mention that while Harvard Business Review appeared as one of the highly co-cited sources, it ranked 35th (138 citations) on the Citation Analysis. Again, Accounting, Organizations, and Society and The International Journal of Operations and Production Management share many linkages through Harvard Business Review, as the latter has been a primary source of publication for the balanced scorecard.

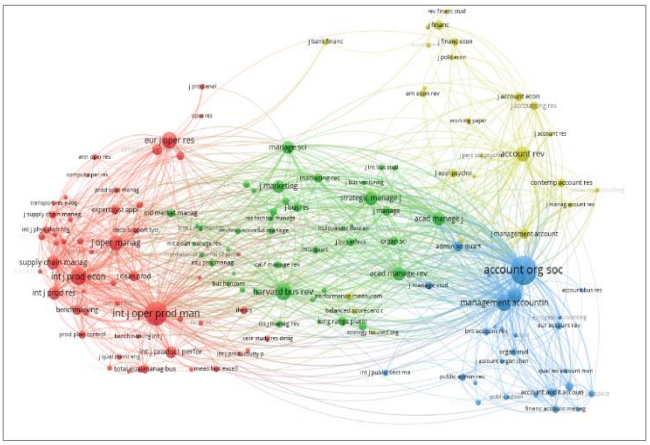

Figure 3 Co-citation by source network

As it relates to co-citation by cited references, the analysis was conducted to understand the level of relatedness in the references being co-cited across the publications in the dataset. The dataset included 68,035 cited references in the publications. **Table 7** shows the top 10 co-cited references by the number of citations. Kaplan *et al.*, (1992), the preliminary work on using the balanced scorecard as a tool for performance measurement, was the most cited reference with 296 citations. This endurance is insightful to further inform the academic maturity assessment of performance measurement as a field. It is also clear that the context of performance measurement has evolved into the implementation/organizational integration of the concept. This noted evolution was incorporated into the update to Neely's 2005 proposed Evolutionary Cycle of Performance Measurement Research, which will be discussed more in-depth later in this paper.

Table 7 Co-citation by reference

Cited Reference	Citations	Total Link Strength
Kaplan, RS. and Norton, D.P. (1992), "The balanced scorecard: measures that drive performance," *Harvard Business Review,* January-February, pp. 71-9	296	3146
Chenhall, R. H. (2003). Management control systems design within its organizational context: findings from contingency-based research and directions for the future. Accounting, Organizations and Society, 28(2–3), 127–168	127	2112
Kaplan, RS. and Norton, D.P. (1996a), The balanced scorecard: Translating strategy into Action, Harvard Business School Press, Boston, Boston, MA	189	2044
Ittner, C. D., Larcker, D. F., & Randall, T. (2003). Performance implications of strategic performance measurement in financial services firms. Accounting, Organizations and Society, 28(7–8), 715–741	106	2007
Chenhall, R. H. (2005). Integrative strategic performance measurement systems, strategic alignment of manufacturing, learning and strategic outcomes: an exploratory study. Accounting, Organizations and Society, 30(5), 395–422	101	1934
Neely, A., Gregory, M., & Platts, K. (1995). Performance measurement system design. International Journal of Operations & Production Management, 15(4), 80–116	140	1630
Bourne, M., Mills, J., Wilcox, M., Neely, A., & Platts, K. (2000). Designing, implementing and updating performance measurement systems. International Journal of Operations & Production Management, 20(7), 754–771	95	1456
Henri, J.-F. (2006). Management control systems and strategy: A resource-based perspective. Accounting, Organizations and Society, 31(6), 529–558	73	1445
Otley, D. (1999). Performance management: a framework for management control systems research. Management Accounting Research, 10(4), 363–382	97	1439
Kaplan, R. S., & Norton, D. P. (2001). *The strategy-focused organization: how balanced scorecard companies thrive in the new business environment*. Boston, Mass.: Harvard Business School Press.	89	1375

The predominant nodes showed Chenhall (2003), Kaplan *et al.*, (1992), Kaplan *et al.*, (1996a/b), and Neely *et al.*, (1995) as having the highest TLS score indicated by the links between the nodes. As alluded to earlier, relevant literature in performance measurement continued to reference the balanced scorecard as the basis for many of the works. It was also interesting to observe the consistent relatedness among the areas of performance measurement, balanced scorecard, accounting, and operations management (evidenced by Neely's work on performance measurement systems design, Neely *et al.*, (1995). Lastly, aside from work by Charnes *et al.* (1978), the remaining works seem to be heavily co-referenced.

5. TEST DATA-BASED ANALYSIS

Term Map. Within the context of this paper, a Term Map differed from the Keyword Co-occurrence Network in two aspects: firstly, a "keyword" analysis based on words that have

been identified explicitly by the author (i.e., Author's Keywords) or terms appearing in the references of the article but not the article itself (i.e., Key Word Plus). The Term Map, however, is an analysis based on the frequency of words appearing in the title and the abstract of the work without being labeled explicitly as a "keyword" (Van Eck *et al.,* 2007). Secondly, the premise used in both analyses is different because a keyword analysis has been designed (or somewhat directed/shaped) by the author and is limited to including the keyword list and reference list (in the case of Keyword Plus). In contrast, a term analysis would reveal any additional and unintended themes or patterns that the author did not influence (or are limited to being only captured from the publication's keyword list or reference list).

The 1,943 papers included in the dataset identified a total of 32,749 terms. A count of 1,220 had a minimum of 10 occurrences when extracted from the title and abstract fields. To refine the network and to draw insight from the most prominent terms, the occurrence threshold was increased to 50. A total of 243 terms occurred more than 50 times in the title and abstract of each publication. The most occurred terms were "model" (1,527 occurrences), followed by "practice" (985 occurrences), then "measure" (773 occurrences). **Figure 4** shows a closeup of the three emerged nodes (model, practice, and measure), providing further support to one crucial position proposed in this paper: the transitioning of the field into an implementation/practice phase and the evident interest of academics and researchers in the modeling aspect within performance measurement.

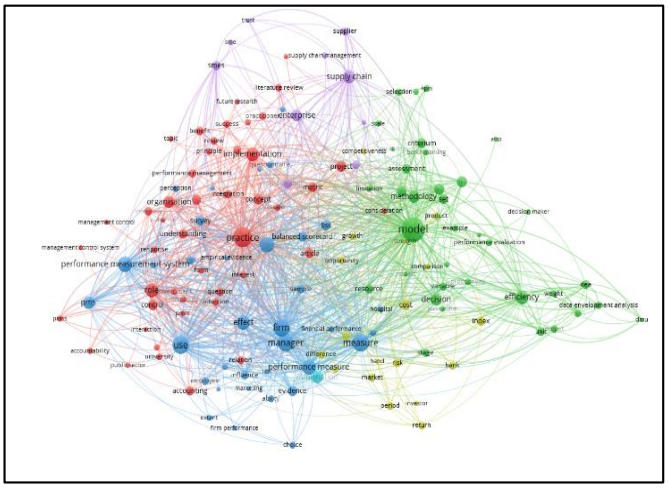

Figure 4 Term Map

The remainder of this paper synthesizes the findings further. It presents discoveries in the form of 1) an informed, inclusive update to the research evolution in performance measurement (including an update to the evolutionary research cycle, 2) a set of responses to the general questions around academic maturity and professionalism of the field, 3) insights into Neely's previously proposed agenda research questions, and 4) insights for future outlook for performance measurement research field.

6. PERFORMANCE MEASUREMENT RESEARCH EVOLUTION – AN INCLUSIVE UPDATE

Before the 1990s, the field of performance measurement was best described as widely diverse, with little consensus in the literature on aspects related to the theoretical foundation (Neely, 2005; Neely *et al.*, 1995). Early on, significant studies focused on quantifying business performance and understanding its organizational implications (Argyris, 1952; Hays and Abernathy, 1980; Ridgway, 1956; Power, 1997). It was a period of problem identification among scholars and academics, as noted in Neely's 2005 update. A richness followed that period in academic research around proposing various frameworks to bring structure and scoping to the field; one prominent phenomenon was the overreliance in the literature on mostly a singular concept (namely, the balanced scorecard). By the late 1990s, the research community was interested in performance measurement methods of application of the proposed frameworks, followed by a period of empirical investigation of those frameworks. In early 2000, evidence suggested that academic interest shifted to theoretical verification of the proposed methodologies, a shift conceptualized in Neely's proposed Evolutionary Cycle of Performance Measurement. He noted that there was no clear delineation among these phases in this framework, "but the cycle appears to be a reasonable proxy for understanding the development of the performance measurement field" (Neely, 2005).

The section to follow is structured to synthesize the findings from this study and provide an update on the following points as noted in Neely's 2005 update: wide disciplinary diversity, lack of academic professionalism, a plateau of field research, and the overreliance on a limited number of frameworks.

Wide disciplinary diversity. Findings from this study suggest that while performance measurement can still be considered a diverse discipline, this diversity has been lessening over the past 15 years. Predominantly, the more significant portion of academic activities in performance measurement has been consolidated into the management/business, operations management, and supply chain areas. This position can be argued based on the fact that of the 141 field categories initially containing performance measurement-related work (in the Web of Science database), the top five categories (3.5 percent) included roughly 30 percent of the performance measurement work.

Lack of academic professionalism. Partially supported by the above point, performance measurement has professionalized academically to some extent. The research community's interest has grown to be more sophisticated and selective in focusing on fewer disciplines. This observation has been noted repeatedly in the network linkages and TLS scores connecting the most co-cited authors, co-cited sources, and references. For example, the overall close relatedness among key authors like Kaplan, Chenhall, and Neely in their work on the balanced scorecard, strategic alignment, and operations management within the performance measurement area.

Plateau of field research. This is essentially the question of whether or not the field has "had its day," which was posed in Neely's 2005 update. The findings from this study suggest that the academic interest and momentum in the field are enduring. This position can be argued based on the steadily increasing trend in publications and the annual publication rate, from 34 publications per year in 2005 to 208 publications per year in 2019 (removing the partial year of 2020 for consistency). The field continues to attract the attention of academics and scholars.

Overreliance on a limited number of frameworks. While the overreliance on the balanced scorecard has decreased (based on the number of cited documents and cited authors), this has produced less exciting news. Firstly, as noted earlier, while there is an apparent reduction in the overreliance on the balanced scorecard, it continues to be among the most referenced concepts in the literature. Secondly, the misalignment between the citation and co-citation analysis results around the balanced scorecard is somewhat puzzling. In other words, while there are fewer publications of balanced scorecard-based papers (i.e., original research), the fundamental publications (particularly early on in the 1990s) of the balanced scorecard (namely, Kaplan *et al.*, 1992; Kaplan *et al.*, 1996a/b; Kaplan *et al.*, 2001) are among the topmost frequently cited references. These four references have been cited 678 times out of 1,856 (37 percent). In the order of ranking, Kaplan *et al.* (1992), Kaplan *et al.* (1996a), Kaplan *et al.* (2001), and Kaplan *et al.* (1996b) have been ranked first, second, eighth, and thirteenth most cited references, respectively. One possible explanation is their usage as one of the preliminary and fundamental works in performance measurement, which is reasonable since the norm in academic publishing is to include fundamental works as part of the background and literature review. Not to mention that the balanced scorecard is a predominant "performance measurement tool"; hence, it was expected to have it continuously referenced in performance measurement papers for the foreseeable future. Furthermore, findings from this study inform an update to the evolutionary research cycle as proposed in Neely's 2005 update. **Figure 5** illustrates the revised proposed cycle (addition of the "Practical Implementation" phase.

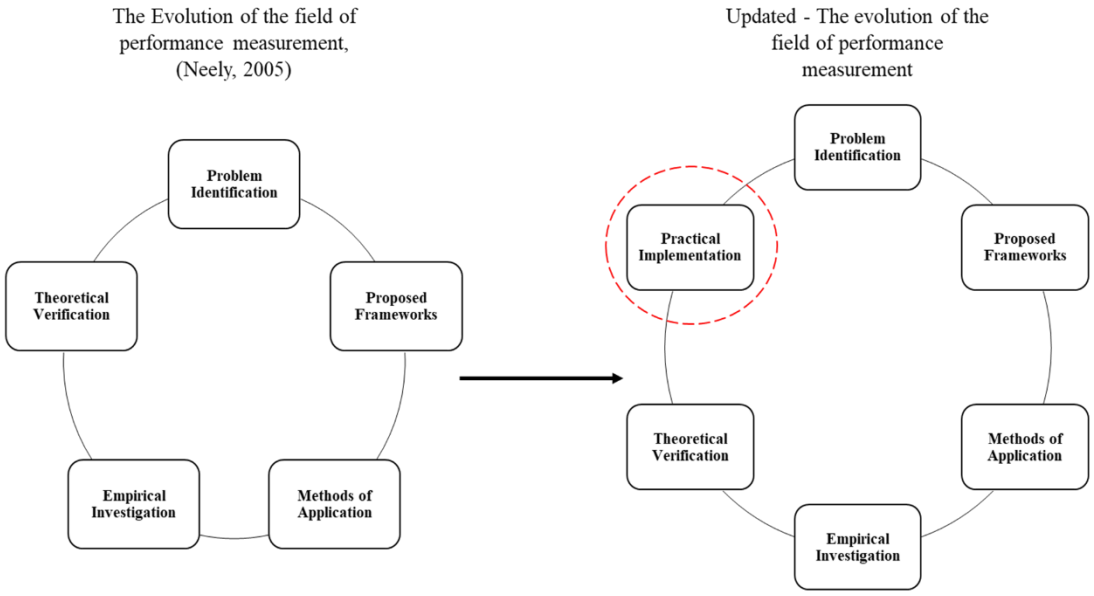

Figure 5 Updated evolutionary cycle of performance measurement research

With an evident academic interest in understanding the instrumentation of performance measurement systems in organizations, it can be argued that the field has evolved into a practical implementation phase. During this phase, the focus shifted to the design, institutionalization, and implementation challenges. While some studies aimed at understanding performance measurement practices (Gomes *et al.*, 2011; Jauhola, 2005; Ukko *et al.*, 2017; Ukko *et al.*, 2020; Rouse *et al.*, 2010; Silvi *et al.*, 2015), others aimed at empirically investigating the implementation of performance measurement systems (Bracci *et al.*, 2017; Dimitrov *et al.*, 2007; Ittner, 2014; Jardioui *et al.*, 2015; Keathley-Herring *et al.*, 2017; Liang *et al.*, 2018; Micheli *et al.* 2011; Pekkola *et al.* 2016; Rosado *et al.*, 2015; Sofyani *et al.*, 2018). Additionally, several studies in the literature conducted practice-based examinations into various aspects of performance measurement (Gomez *et al.*, 2011; Micheli *et al.*, 2014; Saunila *et al.*, 2015).

6.1 Addressing the prevalent questions from Neely's 2005 update

Table 8 and **Table 9** summarize synthesized findings from this bibliographic study to provide 1) answers to questions and 2) insight into research progress around specific aspects from Neely's 2005 update, respectively.

Table 8 Responses to questions on academic maturity and professionalisms

From Neely's 2005 update: Questions related to academic professionalism and maturity	From this study: Proposed answers based on a bibliographic study from 2005 to 2020
Why hasn't the performance measurement field professionalized and matured academically?	It is reasonable to argue that the lack of academic maturity constituted an embryonic phase for the field as of 2005 (the area was less than 15 years old). After another 15 years have passed, performance measurement is roughly a 30-year-old field, and it has shown some signs of academic maturity and selective concentration on fewer disciplines
Why is the performance measurement research community so dependent on a limited number of works from a limited number of contributors?	In the absence of empirical evidence to support a rationale for this phenomenon, one possible explanation is a general lack of empirical validation to many of the frameworks developed in the past 15 years (will be discussed further in a later section)
Has the field had its day?	Based on data from this study, it can be argued that the field has not had its day yet; there is evidence of continuous academic momentum, which is evident in the steadily increasing rate of publications
Why have no new dominant ideas or breakthroughs emerged since 1995? Is there a deeper-rooted problem?	Similar to the inquiry regarding dependence on a limited number of contributions: one possible explanation is the lack of empirical validation for these works. Another is perhaps the complexity of the generated frameworks so far when compared to the balanced scorecard (for example)

Regarding the questions geared toward the research community, the intention was not to assess the validity of the research agenda questions proposed by Neely's 2005 update. Nor was it to comment on their integrity in advancing the research agenda. The goal, however, was to shed light on the progress made in each of those question areas and provide insight into whether or not they endured to provide direction for future research and, therefore, should be explored. In support of the insights provided, selected studies from the literature review were highlighted. The extent of the studies could provide initial evidence of the meaningfulness (or lack thereof) of the progress made in performance measurement research from 2005 to 2020. The insight was summarized and structured as shown in **Table 9**.

Table 9 Insights into progress of performance measurement research

From Neely's 2005 update: Questions related to advancing the research agenda	From this study: Insight into the progress of the research in the field
How to design and deploy enterprise performance management rather than measurement systems?	Evidence for meaningful progress as many studies have explored the transition from performance "measurement" to "management." Studies include (Broadbent *et al.,* 2009; Brun *et al.,* 2009; Folan *et al.,* 2005; Melnyk *et al*., 2014; Neely *et al.,* 2006; Radnor *et al.,* 2007; Taticchi *et al.,* 2008a/b; Taticchi *et al.,* 2010; Yadav *et al.,* 2013)
How to measure performance across supply chains and networks rather than within organizations?	Evidence for meaningful progress based on Term Map analysis showed supply chain being among the most frequently used terms in the papers included in the dataset. Additionally, several studies aimed to examine the suitability of performance measurement for the extended enterprise (Bai *et al.,* 2008; Beamon, 1999; Behrouzi *et al.,* 2011; Beske-Janssen *et al.,* 2015; Chang 2009; Cocca *et al.,* 2010; Diran Wickramatillake *et al.,* 2007; Lehtinen *et al.,* 2010; Olugu *et al.,* 2011; Ren *et al.,* 2006; Sillanpää, 2015; Zegordi *et al.,* 2009; Zhong *et al.,* 2010;)
How to measure intangible as well as tangible assets for external disclosure as well as internal management?	A limited number of studies indicated little evidence of meaningful progress (Corona, 2009; Johnson, 2006). Additionally, neither the Term Map nor Keyword co-occurrence analyses revealed significant findings relative to tangible, intangible, or external disclosure
How to develop dynamic rather than static measurement systems?	Some evidence of meaningful progress was indicated by a limited number of studies, including (Bianch *et al.,* 2015; Bisbe *et al.,* 2012; Grillo *et al.,* 2018; Hasegan *et al.,* 2018; Irfani *et al.,* 2019). Additionally, neither the Term Map nor Keyword co-occurrence analyses revealed significant findings relative to dynamic or static performance measurement systems
How to enhance the flexibility of measurement systems to cope with organizational changes?	Very little evidence of meaningful progress. One main paper was by (Bititci *et al.,* 2011). In the paper, the authors posed vital questions like "how do performance-measurement systems evolve in response to changes in the organization's inner and outer operating environment?" and "how do network-based performance measurement systems evolve in response to changes in networks' inner and outer operating environment?" A study worth noting (although not included in the dataset) was Kennerley *et al.* (2003), where authors discussed the challenges of measuring business performance amid changing business environments. Advancing this work would contribute to exploring the flexibility enhancement of measurement systems.

7. INSIGHTS FOR AN OUTLOOK ON THE FUTURE PERFORMANCE MEASUREMENT RESEARCH

Studying the evolution of performance measurement research from 2005 to 2020 revealed mixed findings. On the one hand, the field showed some signs of academic maturity and professionalism. This finding was reflected in the concertation on fewer disciplines where a significant portion of the academic research activities occurred (namely, management/business, operations management, and supply chain). On the other hand, there wasn't adequate evidence to argue the maturity of the theoretical foundations; the field still relied on a limited number of works. This finding lent itself directly to the yet-to-be-answered conundrum Neely (2005) pointed out: the why? Although findings could inform more possible explanations, as alluded to earlier, the main question remains. However, perhaps it is more worthwhile for the research community to investigate how scholars can develop new ideas and frameworks to ensure their effectiveness and "stickiness."

Additionally, very few studies -in general- in the field of performance measurement were designed based on a holistic case study approach in which an organization is the unit of analysis. More of this empirical validation is needed. This need is critical based on the general lack of theoretical validation of the proposed frameworks found in the literature. For example, an interdisciplinary team of researchers can develop a framework to synergize aspects related to evolving performance measurement systems. Such a move would be a step forward in assessing the maturity of the theoretical foundations in the field. The goal should be to uncover organizational barriers and hurdles that may contribute to the inability of real organizations to effectively evolve those metrics (in a sense, this is an exercise of empirically validating the theoretical frameworks). There may also be industry-related or contextual environment-related aspects that need to be understood. This notation is critical because if real organizations are not provided with a practical roadmap to evolve those performance metrics, theoretical validation, l*et al*one proving implementation feasibility, will endure as a significant challenge for the field. In terms of advancing the research agenda, this study proposes the following set of informed questions:

- *How can performance measurement systems be evolved to support the development of new business models for organizations?*
- *How does an organization practically evolve its business metrics to enable it to plan for a future that is not here yet?*

Considering the implications of the fourth industrial revolution and the fusion of technologies that are diminishing barriers across industries, how does an organization contemplating traditional vs. new business directions develop a robust business performance measurement system that helps it survive today and thrive tomorrow?

How do we ensure strategic alignment between the corporate level strategy and the business performance metrics, particularly as the latter continue to evolve?

Another proposed area for future research is to conduct a similar deep dive analysis in other disciplines concerning performance management. For example, it would be insightful to understand the impact of the research evolution on various fields, including but not limited to supply chain.

Megatrends are shifting the competitive landscapes for organizations across a multitude of industries. Organizations are finding themselves competing with traditional and non-traditional competitors. As the latter typically operate under different business models and to different profit formulas (driven by other performance metrics and measurement systems), the former face the unique challenge of tightening the existing business model and creating the new one. Evolving the performance measurement systems becomes the craft of the successful, so the research community is urged now more than ever to advance the knowledge in the field.

8. CONCLUSION

This review of the performance measurement literature demonstrates the dynamic evolution of this area of inquiry. While academic maturity can be argued (demonstrated by the focus on operations research and supply chain), the field can still be characterized by being multidisciplinary to a large extent. There is also a lack of novel theoretical contributions. This study aimed to accomplish five goals: (1) to understand the intellectual structure of current literature, (2) to explain how the literature has evolved, (3) to investigate whether and how the evolution in the research of performance measurement could contribute to the overall development of the field, (4) to provide informed responses and insights into the research agenda inquiries as proposed by Neely in 2005, and (5) to propose an informed agenda for future research. Findings from this paper suggest that the literature is divided into six main areas of inquiry, with operations research and supply chain being the main ones. Further, it makes three main contributions to the performance measurement literature. Findings from this study could inform the performance measurement literature in 3 main ways; by highlighting the absence of adequate examination of *effectiveness*, an essential complement to the extensively studied *efficiency* in the current literature*;* by indicating 'practical implementation' as the current stage of research evolution of the field; and by emphasizing the need to adopt novel theoretical perspectives. Future performance measurement studies could benefit from considering the evolution of business models informed by the fourth industrial revolution and the practical strategic alignment between corporate strategy and the ever-evolving performance measurement systems.

REFERENCES

Aramyan, L. H., Oude Lansink, A. G. J. M., van der Vorst, J. G. A. J., & Van Kooten, O. (2007). Performance Measurement in Agri-food Supply Chains: A Case Study. *Supply Chain Management: An International Journal*, 12(4), pp. 304–315.

Argyris, C. (1952), *The Impact of Budgets on People*, Controllership Foundation, New York, NY.

Bai, S., and Liu, X., (2008). Application of Data Envelopment Analysis in Supply Chain Performance Measurement. ICPOM2008: *Proceedings of 2008 International Conference of Production and Operations Management*. Vol. 1-3.

Baines, T., & W. Lightfoot, H. (2013). Servitization of the Manufacturing Firm. International *Journal of Operations & Production Management*, 34(1), pp. 2–35.

Barros, C. P., Managi, S., & Matousek, R. (2012). The Technical Efficiency of the Japanese Banks: Non-Radial Directional Performance Measurement with Undesirable Output. *Omega*, 40(1), pp. 1–8.

Beamon, B.M. (1999), Measuring Supply Chain Performance, *International Journal of Operations & Production Management*, 19(3-4), pp. 275-292.

Behrouzi, F., Wong, K. Y., & Behrouzi, F. (2011). A study on Lean Supply Chain Performance Measures of SMEs in the Automotive Industry. 2011 IEEE International Conference on Industrial Engineering and Engineering Management. Presented at the *2011 IEEE International Conference on Industrial Engineering and Engineering Management* (IEEM). https://doi.org/10.1109/ieem.2011.6117914.

Beske-Janssen, P., Johnson, M. P., & Schaltegger, S. (2015). 20 Years of Performance Measurement in Sustainable Supply Chain Management – What Has Been Achieved? *Supply Chain Management: An International Journal*, 20(6), pp. 664–680.

Bisbe, J., & Malagueño, R. (2012). Using Strategic Performance Measurement Systems for Strategy Formulation: Does It Work in Dynamic Environments? *Management Accounting Research*, 23(4), pp. 296–311.

Bititci, U., Garengo, P., Dörfler, V., & Nudurupati, S. (2011). Performance Measurement: Challenges for Tomorrow. *International Journal of Management Reviews*, 14(3), pp. 305–327.

Bottani, E., & Rizzi, A. (2006). A Fuzzy TOPSIS Methodology to Support Outsourcing of Logistics Services. *Supply Chain Management: An International Journal*, 11(4), pp. 294–308.

Bourne, M., Mills, J., Wilcox, M., Neely, A., and Platts, K. (2000), Designing, Implementing and Updating Performance Measurement Systems, *International Journal of Operations & Production Management*, 20(7), pp. 754-771.

Bracci, E., Maran, L., & Inglis, R. (2017). Examining The Process of Performance Measurement System Design and Implementation in Two Italian Public Service Organizations. *Financial Accountability & Management*, 33(4), pp. 406–421.

Broadbent, J., & Laughlin, R. (2009). Performance Management Systems: A Conceptual Model. *Management Accounting Research*, 20(4), pp. 283–295.

Chang, H. H. (2009). An Empirical Study of Evaluating Supply Chain Management Integration using the Balanced Scorecard in Taiwan. *The Service Industries Journal*, 29(2), pp. 185–202.

Charles, A., Cooper, W.W. and Rhodes, E. (1978), Measuring Efficiency of Decision-Making Units, *European Journal of Operations Research*, 2(6), pp. 429-444.

Chenhall, R. H. (2003). Management Control Systems Design within Its Organizational Context: Findings from Contingency-Based Research and Directions for the Future. *Accounting, Organizations and Society*, 28(2–3), pp. 127–168.

Chenhall, R. H. (2005). Integrative Strategic Performance Measurement Systems, Strategic Alignment of Manufacturing, Learning and Strategic Outcomes: An Exploratory Study. *Accounting, Organizations and Society*, 30(5), pp. 395–422.

Cocca, P., & Alberti, M. (2010). A Framework to Assess Performance Measurement Systems in SMEs. *International Journal of Productivity and Performance Management*, 59(2), pp. 186–200.

Corona, C. (2009). Dynamic Performance Measurement with Intangible Assets. *Review of Accounting Studies*, 14(2–3), pp. 314–348.

Cvetkoska, V. and Eftimov, L. (2021). Bibliometric Analysis for Performance Measurement in Business, Management, and Accounting Subject Area. *Journal of Engineering Management and Competitiveness*, 11(1), pp. 51-63.

Diran Wickramatillake, C., Lenny Koh, S. C., Gunasekaran, A., & Arunachalam, S. (2007). Measuring Performance within the Supply Chain of a Large Scale Project. *Supply Chain Management: An International Journal*, 12(1), pp. 52–59.

Dixon, J., Nanni, A. and Vollmann, T. (1990), *The New Performance Challenge*, Business One Irwin, Burr Ridge, IL.

Drucker, P. (1954), *The Practice of Management*, Harper, New York, NY.

Fitzgerald, L., Johnston, R., Brignall, S., Silvestro, R. and Voss, C. (1991), *Performance Measurement in Service Business*, CIMA, London.

Folan, P., & Browne, J. (2005). A review of Performance Measurement: Towards Performance Management. *Computers in Industry*, 56(7), pp. 663–680.

Franco-Santos, M., Kennerley, M., Micheli, P., Martinez, V., Mason, S., Marr, B., Gray, D., & Neely, A. (2007). Towards A Definition of A Business Performance Measurement System. *International Journal of Operations & Production Management*, 27(8), pp. 784–801.

Franco-Santos, M., Lucianetti, L., & 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.

Gomes, C. F., Yasin, M. M., & Lisboa, J. V. (2011). Performance Measurement Practices in Manufacturing Firms Revisited. *International Journal of Operations & Production Management*, 31(1), pp. 5–30.

Grillo, H., Campuzano-Bolarin, F., Mula, J. (2018). Modelling Performance Management Measures through Statistics and System Dynamics-Based Simulation. *Dirección y Organización*., pp. 20-35.

Hasegan, M. F., Nudurupati, S. S., & Childe, S. J. (2018). Predicting Performance – A Dynamic Capability View. *International Journal of Operations & Production Management*, 38(11), pp. 2192–2213.

Hayes, R.H. and Abernathy, W.J. (1980), Managing Our Way to Economic Decline, *Harvard Business Review*, July-August, 67-77.

Henri, J.-F. (2006). Management Control Systems and Strategy: A Resource-Based Perspective. *Accounting, Organizations and Society*, 31(6), pp. 529–558.

Hollis, A. (2001). Co-authorship and the output of Academic Economists, *Labour Economics*, 8, pp. 503– 530.

Irfani, D. P., Wibisono, D., & Basri, M. H. (2019). Integrating Performance Measurement, System Dynamics, and Problem-Solving Methods. *International Journal of Productivity and Performance Management*, 69(5), pp. 939–961.

Ittner, C. D. (2014). *Performance Measurement, Implementation Issues* in C. Clubb (ed.), The Blackwell Encyclopedia of Management: Accounting, 2nd Edition. London: Blackwell Publishing, 2005. Reprinted in C. Clubb and S. Imam (eds.), Wiley Encyclopedia of Management, 3rd edition. Chicester, UK: Wiley.

Ittner, C. D., Larcker, D. F., & Randall, T. (2003). Performance Implications of Strategic Performance Measurement in Financial Services Firms. *Accounting, Organizations and Society*, 28(7–8), pp. 715–741

Jardioui, M., Alami, S., & Okar, C. (2015, October). What Are the Critical Success Factors for the Implementation of Supply Chain Performance Measurement Systems in SMEs. 2015 International Conference on Industrial Engineering and Systems Management (IESM). *2015 International Conference on Industrial Engineering and Systems Management (IESM)*. https://doi.org/10.1109/iesm.2015.7380180.

Johnson, H.T. and Kaplan, R.S. (1987), *Relevance Lost - The Rise and Fall of Management Accounting*, Harvard Business School Press, Boston, MA.

Johnson, N. B. (2006). Divisional Performance Measurement and Transfer Pricing for Intangible Assets. *Review of Accounting Studies*, 11(2–3), pp. 339–365.

Kaplan, R. S., & Norton, D. P. (2001). *The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment*. Boston, Mass.: Harvard Business School Press.

Kaplan, R.S. and Norton, D.P. (1992), The Balanced Scorecard: Measures That Drive Performance, *Harvard Business Review*, January-February, pp. 71-79.

Kaplan, R.S. and Norton, D.P. (1996a), Using the Balanced Scorecard as a Strategic Management System, *Harvard Business Review*, 74(1), pp. 75-85.

Kaplan, R.S. and Norton, D.P. (1996b), *The Balanced Scorecard: Translating Strategy into Action*, Harvard Business School Press, Boston, Boston, MA.

Kaplan, RS. and Norton, D.P. (1992), The Balanced Scorecard: Measures That Drive Performance, *Harvard Business Review*, January-February, pp. 71-79.

Keathley-Herring, H. (2017). An Approach to Quantify the Factors That Affect Performance Measurement System Implementation. *Engineering Management Journal*, 29(2), pp. 63–73.

Keegan, D.P., Eiler, R.G. and Jones, C.R. (1989), Are Your Performance Measures Obsolete? *Management Accounting*, June, 45-50.

Kennerley, M. and Neely, A.D. (2003), Measuring Performance in a Changing Business Environment, *International Journal of Operations & Production Management*, 23(2), pp. 213-229.

Lehtinen, J., & Ahola, T. (2010). Is Performance Measurement Suitable for an Extended Enterprise? *International Journal of Operations & Production Management*, 30(2), pp. 181–204.

Leuschner, R., Rogers, D. S., & Charvet, F. F. (2013). A Meta-Analysis of Supply Chain Integration and Firm Performance. *Journal of Supply Chain Management*, 49(2), pp. 34–57.

Liang, L., Yang, F., Cook, W. D., & Zhu, J. (2006). DEA Models for Supply Chain Efficiency Evaluation. *Annals of Operations Research*, 145(1), pp. 35–49.

Liang, X., Gao, Y., & Ding, Q. S. (2018). What you measure is What You Will Get? Exploring the Effectiveness of Marketing Performance Measurement Practices. Cogent Business & Management, 5(1).

Lynch, R.L. and Cross, K.F. (1991), *Measure up!* Blackwell Publishers, Cambridge, MA.

Medori, D. and Steeple, D. (2000), A Framework for Auditing and Enhancing Performance Measurement Systems, *International Journal of Operations & Production Management*, 20(5), pp. 520-553.

Melnyk, S., Bititci, U., Platts, K., Tobia, J., Andersen, B. (2014), Is Performance Measurement and Management Fit for the Future? *Management Accounting Research*, 25, pp. 173-186.

Meng, X. (2012). The Effect of relationship management on Project Performance in Construction. *International Journal of Project Management*, 30(2), pp. 188–198.

Micheli, P., & Mari, L. (2014). The Theory and Practice of Performance Measurement. *Management Accounting Research*, 25(2), pp. 147–156.

Micheli, P., Mura, M., & Agliati, M. (2011). Exploring the Roles of Performance Measurement Systems in Strategy Implementation. *International Journal of Operations & Production Management*, 31(10), pp. 1115–1139.

Najmi, M., Rigas, J. and Fan, I.S. (2005), A Framework to Review Performance Measurement Systems, *Business Process Management Journal*, 11(2), pp. 109-122.

Neely, A. (2005). The Evolution of Performance Measurement Research. *International Journal of Operations & Production Management*, 25(12), pp. 1264–1277.

Neely, A., & Al Najjar, M. (2006). Management Learning Not Management Control: The True Role of Performance Measurement? *California Management Review*, 48(3), pp. 101–114.

Neely, A., Adams, C., and Kennerley, M. (2002), *The Performance Prism: The Scorecard for Measuring and Managing Business Success*, FT Prentice Hall, London.

Neely, A., Mills, J., Platts, K., Richards, H., Gregory, M., Bourne, M., & 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.

Neely, A.D., 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.D., Mills, J.F., Gregory, M.J., Richards, A.H., Platts, K.W. and Bourne, M.C.S. (1996), *Getting the Measure of your Business*, Findlay Publications, Horton Kirby.

O'Sullivan, D., & Abela, A. V. (2007). Marketing Performance Measurement Ability and Firm Performance. *Journal of Marketing*, 71(2), 79.

Olugu, E., & Wong, K. (2011). A Study on the Validation of Green Supply Chain Performance Measures in the Automotive Industry. *Communications of the IBIMA*, pp. 1–14.

Otley, D. (1999). Performance Management: A Framework for Management Control Systems Research. *Management Accounting Research*, 10(4), pp. 363–382

Paradi, J. C., Rouatt, S., & Zhu, H. (2011). Two-Stage Evaluation of Bank Branch Efficiency Using Data Envelopment Analysis. *Omega*, 39(1), pp. 99–109.

Pavlov, A. and Bourne, M. (2011), Explaining the Effects of Performance Measurement on Performance: An Organizational Routine Perspective, *International Journal of Production and Operations Management*. 31(1), pp. 101-122.

Pekkola, S., Saunila, M., & Rantanen, H. (2016). Performance Measurement System Implementation in A Turbulent Operating Environment. *International Journal of Productivity and Performance Management*, 65(7), pp. 947–958.

Perkmann, M., Neely, A., & Walsh, K. (2011). How Should Firms Evaluate Success in University-Industry Alliances? A performance measurement system. *R&D Management*, 41(2), pp. 202–216.

Power, M. (1997), *The Audit Society: Rituals of Verification*, Oxford University Press, Oxford.

Radnor, Z. J., & Barnes, D. (2007). Historical Analysis of Performance Measurement and Management in Operations Management. *International Journal of Productivity and Performance Management*, 56(5/6), pp. 384–396.

Ren, C., Dong, J., Ding, H., & Wang, W. (2006). A SCOR-Based Framework for Supply Chain Performance Management. 2006 IEEE International Conference on Service Operations and Logistics, and Informatics. Presented at the *2006 IEEE International Conference on Service Operations and Logistics and Informatics*. https://doi.org/10.1109/soli.2006.328909

Ridgway, V.F. (1956), Dysfunctional Consequences of Performance Measurements, *Administrative Science Quarterly*, 1(2), pp. 240-247.

Rosado, J. O., & Relvas, S. (2015, October). Integral Supply Chain Performance Management System Design and Implementation. 2015 International Conference on Industrial Engineering and Systems Management (IESM). *2015 International Conference on Industrial Engineering and Systems Management (IESM)*. https://doi.org/10.1109/iesm.2015.7380248.

Rouse, P., Harrison, J., & Chen, L. (2010). Data Envelopment Analysis: A Practical Tool to Measure Performance. *Australian Accounting Review*, 20(2), pp. 165–177.

Saunila, M., Tikkamäki, K., & Ukko, J. (2015). Managing Performance and Learning Through Reflective Practices. *Journal of Organizational Effectiveness: People and Performance*, 2(4), pp. 370–390.

Sillanpää, I. (2015). Empirical Study of Measuring Supply Chain Performance. *Benchmarking: An International Journal*, 22(2), pp. 290–308.

Silvi, R., Bartolini, M., Raffoni, A., & Visani, F. (2015). The Practice of Strategic Performance Measurement Systems. *International Journal of Productivity and Performance Management*, 64(2), pp. 194–227.

Slater, S. F., Mohr, J. J., & Sengupta, S. (2013). Radical Product Innovation Capability: Literature Review, Synthesis, and Illustrative Research Propositions. *Journal of Product Innovation Management*, 31(3), pp. 552–566.

Sofyani, H., Akbar, R., & C. Ferrer, R. (2018). 20 Years of Performance Measurement System (PMS) Implementation in Indonesian Local Governments: Why is Their Performance Still Poor? Asian Journal of Business and Accounting, 11(1), pp. 151–184.

Spekle, R. F., & Verbeeten, F. H. M. (2014). The Use of Performance Measurement Systems in the Public Sector: Effects on performance. *Management Accounting Research*, 25(2), pp. 131–146.

Taticchi, P. and Balachandran, K., (2008a), Forward Performance Measurement and Management Integrated Frameworks, *International Journal of Accounting the Information Management*. 16 (2), pp. 140-154.

Taticchi, P. Tonelli, F., Sameh, M., Botarelli, M. (2008b). Performance Measurement and Management: What is Next? *WSEAS Transactions on Business and Economics*. 5(11), pp. 497-506.

Taticchi, P., Tonelli, F., & Cagnazzo, L. (2010). Performance Measurement and Management: A Literature Review and A Research Agenda. *Measuring Business Excellence*, 14(1), pp. 4–18.

Toor, S.-R., & Ogunlana, S. O. (2010). Beyond the Iron Triangle: Stakeholder Perception of Key Performance Indicators (KPIs) for Large-Scale Public Sector Development Projects. *International Journal of Project Management*, 28(3), pp. 228–236.

Ukko, J., & Saunila, M. (2020). Understanding the Practice of Performance Measurement in Industrial Collaboration: From Design to Implementation. *Journal of Purchasing and Supply Management*, 26(1). 100529. https://doi.org/10.1016/j.pursup.2019.02.001.

Ukko, J., Hildén, S., Saunila, M., & Tikkamäki, K. (2017). Comprehensive Performance Measurement and Management – Innovativeness and Performance Through Reflective Practice. *Journal of Accounting & Organizational Change*, 13(3), pp. 425–448.

Van Eck, N.J., & Waltman, L. (2007). VOS: A New Method for Visualizing Similarities between Objects. In H.-J. Lenz, & R. Decker (Eds.), *Advances in Data Analysis: Proceedings of the 30th Annual Conference of the German Classification Society*, pp. 299-306.

Varsei, M., Soosay, C., Fahimnia, B., & Sarkis, J. (2014). Framing Sustainability Performance of Supply Chains with Multidimensional Indicators. *Supply Chain Management: An International Journal*, 19(3), pp. 242–257.

Wouters, M., & Wilderom, C. (2008). Developing Performance-Measurement Systems as Enabling Formalization: A Longitudinal Field Study of a Logistics Department. *Accounting, Organizations and Society*, 33(4–5), pp. 488–516.

Yadav, N., & Sagar, M. (2013). Performance Measurement and Management Frameworks. *Business Process Management Journal*, 19(6), pp. 947–971.

Yang, H., & Pollitt, M. (2009). Incorporating Both Undesirable Outputs and Uncontrollable Variables into DEA: The Performance of Chinese Coal-Fired Power Plants. *European Journal of Operational Research*, 197(3), pp. 1095–1105.

Zegordi, S. H., khosrojerdi, A. H., & Jamali, S. S. (2009). Performance Measurement Framework for Location Decisions on Supply Chain Design. 2009 IEEE International Conference on Industrial Engineering and Engineering Management. Presented at the *2009 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*. https://doi.org/10.1109/ieem.2009.5373110.

Zhou, P., Poh, K. L., & Ang, B. W. (2007). A Non-Radial DEA Approach to Measuring Environmental Performance. *European Journal of Operational Research*, 178(1), pp. 1–9.

Dr. Hanan Alhaddi is an academic practitioner with over 15 years of experience in the automotive industry. Her research agenda is focused on strategy implementation and the evolution of business performance metrics. She is a published author and a faculty member in management at the Ilitch School of Business, Wayne State University in Detroit, MI, where she teaches in the areas of business and strategic management. Dr. Alhaddi completed her doctorate degree in business administration from Lawrence Tech university with a focus on strategic positioning and differentiation. She has also completed a master's degree in management with a concentration on strategic management from Harvard University and a postgrad diploma in strategy and innovation from the University of Oxford. Her academic service includes serving as a reviewer and an associate editor.