### Employee Performance Management: Charting the Field from 1998-2018

<table>
<thead>
<tr>
<th>Journal:</th>
<th><em>International Journal of Manpower</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>IJM-10-2019-0483.R2</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keywords:</td>
<td>Performance management, Employee behaviour, Employee productivity, Human resource management</td>
</tr>
</tbody>
</table>
Employee Performance Management: Charting the Field from 1998-2018

Abstract

Purpose
The purpose of this paper is to study the development and consider the future of one of the most controversial areas of human resource development - employee performance management (PM).

Design
Through bibliometrics, a multiple correspondence analysis identifies the main research directions of PM studies and provides a map of descriptors and a list of authors, along with a framework to track PM literature over 20 years (1998-2018).

Findings
Scholars have attempted to address some of the questions raised by earlier researchers. However, critical questions remain unanswered, and there is increasing dissatisfaction with the process. The most glaring yet unaddressed problem with PM is poor employee acceptability of the process.

Practical Implications
If the research gaps are addressed, the lack of acceptability of the PM could be resolved, and more effectively managed in the future.

Originality
The study particularly addresses poor employee ‘acceptability’ of the PM process, a subject that has received limited attention by scholars.
Employee Performance Management: Charting the Field from 1998-2018

1. Introduction

The purpose of this paper is to study the development and consider the future of the employee performance management (PM) process. In order to do this, the PM literature extensively examines the leading management and psychology journals over a 20 years period. The paper focuses on the broad and comprehensive topic of employee PM as opposed to the narrow approach of the performance appraisal (Claus and Briscoe, 2009). There is a distinct and essential difference in these two processes, yet they are frequently confused by both practitioners and academics alike.

Performance appraisal refers explicitly to the supervisor-employee interview where employees are typically evaluated once a year using a given set of dimensions and assigned a score to that assessment (DeNisi and Murphy, 2017). Moreover, performance appraisal has a very chequered history associated with control, hierarchical management and, more recently, complex information technology processes that have resulted in mounting dissatisfaction in both supervisors and employees (Adler, Campion, Colquitt, Grubb, Murphy, Ollander-Krane, and Pulakos, 2016).

PM was introduced in the early 1990s to address the well-documented limitations of performance appraisal (Arvey and Murphy, 1998). The process of PM encompasses a much broader range of management practices, including career management, training and development, regular feedback and reimbursement considerations (Aguinis, 2007). PM is a continual process, as opposed to the once-a-year event of appraising performance expectations. At the same time, “PM is seen as an integrated process in which managers work
with their employees to set expectations, measure and review results, and reward performance,“ (Den Hartog, Boselie and Paauwe, 2004, p 4).

This study explicitly examines the comprehensive process of PM and, therefore, includes the role of employee performance appraisal as a subset of PM. The study organises, reviews, critiques and synthesises representative literature on PM in an integrated way so that new frameworks and perspectives on the topic can be generated (Torraco, 2005). The study takes an innovative and rigorous bibliometric approach. To the best of our knowledge is the first bibliometric PM literature review of its kind. The study responds to the call for research on the current state of PM practices in established and influential academic journals (Gorman, Meriac, Roch, Ray and Gamble, 2017). Ultimately, we identify and address poor employee ‘acceptability’ of the PM process and make a compelling case that despite its present-day mounting adverse publicity (Cappelli and Tavis, 2016; Buckingham and Goodall, 2015), PM can have a valuable place in the contemporary organisation. Thus, this study tackles poor employee ‘acceptability’ of the PM process, a subject that only been indirectly identified by scholars but not unpacked and resolved. The review also leads to the generation of an empirical-based model for future research that rejects an exclusive yet flawed performance appraisal approach and demands a return to a more inclusive, systematic approach to PM research.

The paper is structured as follows: First, we present a brief overview of the historical development of PM literature between 1970 and 1998 (see Table 1). Second, we present a detailed description of the methodological approach utilised to uncover PM research trends and gaps. Third, we offer a detailed description of the analysis and the resulting trends of PM content and its evolution between 1998 and 2018. Fourth, we offer a discussion surrounding
the contributions of our review and conclude with a model focused on avenues for broadening the breadth and scope of future research on PM.

***** Insert Table 1 About Here *****

Table 1: Historical Trends of PM Research 1970–1998

2. Historical Background of Extant PM Literature

Over the past half-century, researchers and practitioners have written an enormous amount on PM research. It is useful to provide a brief overview of some of the traditional treatments and models. Typically, most early researchers focused on performance appraisal. However, the appraisal is just one activity of the PM process (Claus and Briscoe, 2009). The early literature established a vast quantity of conceptual thinking is categorised into measurement thinking (i.e. Landy and Farr, 1980) later cognitive characteristics of the appraisal interview (i.e. DeNisi, Cafferty and Meglino, 1984; Feldman, 1981; Landy and Farr, 1980). Early social cognitive research helped to clarify and refocus research from measurement to information processing and cognition (Landy and Farr, 1983). However, the models adopted at this time made few contributions to the practice of appraisals in organisations (Murphy and Cleveland, 1991).

Following the interest in cognitive studies, researchers looked at the issues surrounding a deficiency of trust, fairness and feedback and the negative impact of politics on performance appraisal. In general, a good appraisal system increased the level of acceptability, which established trust and fairness as a crucial component of the appraisal process. However, Longenecker and Gioia (1988) claim that interview research has been micro-analytic and unitary. Nevertheless, by the mid-1990s, an effective appraisal system was considered as being a continuous process and a tool for managing future performance, not just for scoring
past performance. PM is the answer to the limitations of the narrow performance appraisal, by increasing both employee and management acceptance. Thus, the early literature solicits the following questions: 1) To what extent has PM research developed beyond the narrow PA accuracy? and 2) Has the broader PM increased the acceptability of the process?

3. Methodological Approach

First, a content analysis of the published articles dealing with PM took pace. Content analysis provides for an objective, systematic and quantitative consideration of published articles. It also allows for an interpretation of the direction in which journal editors, reviewers and authors are taking the field as it reflects the evolution of their priorities over time (Furrer and Sollberger, 2007; Furrer, Thomas and Goussevskaia, 2008).

In terms of bibliometrics, scholars consider the Web of Science (WoS) to be one of the leading academic databases for studying research contributions. WoS covers more than 16,000 journals and 70,000,000 articles and, in general, expectations are that the material included in WoS holds the highest quality standards. However, many other databases exist, some of them internationally known (e.g., Scopus, Econ Lit and Google Scholar). Following the systematic search methods found in the review articles (Terjesen, Hessels and Li, 2013), we searched for relevant literature in online databases in the WoS Core Collection.

Bibliometric studies use a wide range of methods, the most popular being those that consider the number of publications and the number of citations. We read the first collection of 312 articles and exchanged notes among three co-authors to reach a consensus on the articles for exclusion from our final sample. We have focused on full-length articles published in WoS.
Core Collection indexed academic journals, as they are considered knowledge-certified by peers.

**Data Sources**

The first step in our analysis was to select the articles to be analysed. Because of the specific focus on PM, we selected every article published between 1998 and 2018. We decided to study articles that had the most impact on the field and, therefore, focused on what Ramos-Rodríguez and Ruíz-Navarro (2004) call ‘certified knowledge’. To do so, we decided to only retain articles from journals indexed in the WoS database, as they can be considered certified. Indeed, the WoS database comprises the most relevant journals for PM research. For example: *International Journal of Human Resource Management; Personnel Review; Human Resource Management; Public Personnel Management; International Journal of Manpower; European Journal of International Management; Human Resource Management Review; Performance Management Systems: A Global Perspective; Journal of Applied Psychology;* and *Journal of Occupational and Organizational Psychology.* In addition, we referred to other high-level outlets for PM scholars in management and psychology journals. It is pertinent to note here that while the certified knowledge technique (Rodríguez and Ruíz-Navarro, 2004) is well established in bibliometric methodology and typically, a good paper could likewise be located in a high impact leading relevant journal (Barney, 1991; Venkatraman, 1989); there can be exceptions. For example, on some occasions, a worthy and well-cited contribution can be found in a journal of less impact.

Articles in the database were retrieved using the search function and the lexemes ‘performance’ and ‘performance appraisal’. The first search returned 103 articles; the next search, which used the lexemes ‘performance’, ‘performance management’, ‘performance
appraisal’, ‘performance planning’, ‘monitoring’, ‘evaluation’, and ‘employee performance management’, raised over 900 articles. However, this string of lexemes raised a substantial and unmanageable per cent of spurious manuscripts; two researchers checking the abstracts verified this. The subsequent search used three lexemes: ‘performance’; ‘performance management’; and ‘performance appraisal’, and raised 120 articles, all of which were relevant and deemed by the research team to be the most accurate research stream. We performed a secondary content analysis and discovered another ten significant articles. We also did a final manual check of late-entry 2018 papers in May 2019. Thus, the total number of relevant articles extended to 140.

Themes Selection and Coding Procedure – Descriptors

We used WordStat software for content analysis in order to obtain the list and frequency of nouns, verbs and compound forms extracted from the articles’ content. In order to code and analyse the content of the articles, and following the procedure proposed by Dabic, González-Loureiro and Furrer (2014), a list of descriptors was derived from the keywords provided by the authors of the articles. The outcome of this process was a list of 38 descriptors most commonly used in research on PM (see Table 2).

***** Insert Table 2 About Here *****

List of Descriptors

The top five most cited descriptors were ‘Across borders’ (f 208), ‘PM’ (f 191), ‘Research Schema’ (f 177), ‘Effectiveness’ (f 163) and ‘Quantitative’ (f 157). The three most common descriptors included 35.8% of the occurrences; the top ten included 57.6%, and the top 20 included 84.0%. The average number of descriptors per article was 9.75, and thus a multivariate approach to data analysis is likely to be more meaningful and valid than
univariate analyses (Hair, Anderson, Tatham and Black, 1998). Thus, following the recommendations of Furrer and Sollberger (2007) and Furrer and colleagues (2008), a multiple correspondence analysis (MCA) was conducted. Following the methodology described by Hoffman and De Leeuw (2011), a matrix with the 38 descriptors was constructed, computing a ‘1’ when each of the latter topics was present in each of the 102 articles, and ‘0’ otherwise.

The MCA provided coordinates in a two-dimensional space for each descriptor. Representing a figure with a large number of words would not be interpretable and, therefore, each of the 38 descriptors represents several keywords (see Table 2). The coordinates of each descriptor correspond to its relative position based on the number of its co-appearances with other descriptors in the sample (Bendixen, 1995).

In order to relate the content and method of research on PM, additional insights emerged, as proposed by González-Loureiro, Dabic and Furrer (2015), by mapping descriptor frequency and distance from quantitative and qualitative research markers. The quantitative marker was designed by using keywords such as: quantitative, statistical distribution, performance appraisal meta-analysis, models, model, predicting turnover, Promethee method, quasi-experiment, item cluster subcomponents, questionnaires, questionnaire, surveillance, survey, 5-factor model, interpretivism, meta-analysis, patterns, alignment research agenda, ANP, AHP, Promethee, system factors, systems, fit, MCDA, visual techniques analytic hierarchy process, measurement organisational behaviour, variables, predictors ratio reinforcement schedules, response rates, internal consistency, dimensions, distribution weighting systems, surveillance, validity, benchmark, inventory, comparison orientation, analytic hierarchy process, multilevel model. The qualitative marker was designed by using keywords such as:
qualitative, interpretivism, concept maps, Foucault, visual techniques, paradigm, schemata, and comparative. The larger the distance between the descriptor and the marker indicates that fewer articles use the specific approach, either quantitative or qualitative. Mapping this distance with the occurrence of each descriptor indicates the nature of the research on PM.

Themes Selection and Coding Procedure – Clusters

Finally, the descriptors were distributed straightforwardly into clusters. As a way of validating the clustering technique, the classification of clusters was initially performed independently by two researchers. Then the clusters were compared and found to be very similar. When creating the clusters, the researchers expended the qualitative coding method described by Glaser (1978), which considers the Six C’s – causes, contexts, contingencies, consequences, co-variances, and conditions. As a cross-check, to increase the reliability of the cluster, the newly labelled clusters were compared with the extreme labelled descriptors as described Dabic et al. (2014) and González-Loureiro et al. (2015). According to these authors, the cluster should be labelled concurring to the most-extreme-located descriptors. Therefore, our clustering technique combined two methods.

Clusters represent themes of articles in a literature network, and a cluster can be seen as a group of well-connected articles in a research area with limited connection to papers in another cluster or research area (Leydesdorff, 2011). Clustering permits the topological analysis of networks, identifying topics, interrelations and collaboration patterns. The researchers agreed their two clustering techniques correlate and generated four significant groups or clusters from the data: 1) Accuracy; 2) Firm-level; 3) Across borders, and 4) Employee acceptability (see Table 3).

***** Insert Table 3 About Here *****
The 4 Major Clusters

The PM literature reveals a steady increase from 1998 onwards with two peaks in publications, in 2007 and 2011. In their recent 100-year review of PM in Journal of Applied Psychology (JAP), DeNisi and Murphy (2017) consider that the literature had its heyday between 1970 and 2000. However, their review is largely confined to I/O Psychology Journals (in particular JAP), which have traditionally focused on the measurement aspect of performance appraisal. We believe their article to be a comprehensive historical review rooted in performance appraisal research and grounded from an I/O psychology perspective. While the DeNisi and Murphy (2017) study take an independent subject review of performance appraisal over 100 years in JAP; the present study takes a broader subject approach of the PM literature over a narrow period within the broader scope of management and I/O psychology journals.

4. Analysis and Key Results

Connection Among Descriptors

The ultimate goal of this analysis is to find a low-dimensional representation of the original high-dimensional space (i.e. the matrix of articles and descriptors). The MCA provided a pair of coordinates in this two-dimensional space for each of the 38 descriptors. On the map of Figure 1, the sizes of the points are proportional to the number of articles associated with a descriptor, and the proximity between descriptors corresponds to their shared content (Bendixen, 1995; Hoffman and Franke, 1986). Descriptors are close to each other when they share a large proportion of articles discussing them and are distant from each other when they appear together in a small number of articles (Furrer et al., 2008). The centre of the map represents the core of PM research in the field. For example, the descriptors ‘Across Borders’, ‘Research Schema’, ‘Effectiveness’ and ‘Quantitative’ are close to the centre. Conversely, the descriptors ‘capabilities’, ‘360’, ‘TQM’, ‘ethics/CSR’ and ‘crisis’ are located
away from the centre, indicating that issues related to these subjects have perhaps only recently or marginally attracted research attention (i.e. CSR and crisis). An alternative view is that they could be declining (i.e. TQM, 360).

The dimensions of the map (Figure 1), which resulted from the MCA, have also been interpreted using the four ‘major codes.’ The first, horizontal, dimension separates the code ‘Firm-level’ (on the left) to the horizontal dimension on the right ‘Across Borders’. The vertical dimension separates the code ‘Accuracy’ (on the top) from those focusing on the major code ‘Employee acceptability’ (at the bottom).

***** Insert Figure 1 About Here *****

Figure 1 Structure of Selected Research

5. Most Influential Papers

In every scientific field, some publications assume fundamental roles in the evolution of the field. These articles, owing to their impact, are accelerating factors in the development of the field (Berry and Parasuraman, 1993). Therefore, in order to gain a better understanding of the directions of the future development of research in the field, it is critical to identify which are the most influential articles on PM published between 1998 and 2018. To measure the impact of an article, we used the generally accepted method of summed citation counts (Bergh, Perry and Hanke, 2006; Furrer et al., 2008). The top 10 most influential papers were initially identified based on the number of appearances in the Social Science Citation Index, accessed through WoS (see Table 4).

In an attempt to further examine and increase our understanding of the trend, we split the most influential papers into two periods. In the period 1998–2005, five papers were
represented in the most influential papers: 1.) Levy and Williams (2004); 2.) Fletcher (2001); 3.) Den Hartog et al. (2004); 4.) Poon (2004); and 5.) Latham, Almost, Mann and Moore (2005). The leading three papers in this period focused on literature reviews and research agendas that could indicate the need at this time for researchers to reflect on the early literature and may be symptomatic of the perplexity of the early literature, driving researchers to seek guidance for future PM research efforts. For example, Levy and Williams (2004) suggested a trend of increasing awareness of the importance of social context, while Fletcher (2001) advocated the need for more practical guidance. In addition, Den Hartog et al. (2004) determined that front-line managers play a crucial mediating role in the implementing of PM and affect employee performance, which in turn affects firm performance. Their model also addresses a concept of reversed causality. In other words, the model proposes that besides PM influencing performance, there is also a reverse loop and a healthy bottom-line may have a positive effect on the willingness to invest in PM practices. The authors conclude that unravelling ‘cause’ and ‘effect’ is challenging. The remaining papers in this period include a conceptual paper by Latham et al. (2005), which creates a model that spotlights the benefits of management coaching, and an empirical paper. For instance, Poon (2004) surveyed the influence of politics on turnover intention.

On the other hand, in the period 2006–2018, there is evidence of an increased tendency to focus on employee-level issues. For instance, Kuvaas (2006; 2007), using a data set from the Norwegian banking sector, set out to explore alternative relationships between PM satisfaction and employee outcomes in the form of self-reported work performance, organisational commitment and turnover intention (Kuvaas, 2006) and the relationship between employee perceptions of developmental PM and self-reported work performance (Kuvaas, 2007). Also focusing on the employee, Gruman and Saks (2011)
address the lack of employee acceptability of the PM process by fostering and managing
employee engagement to increase job performance through PM. This conceptual paper
presents a coherent model and process for promoting the engagement of employees that goes
beyond the use of engagement surveys focusing on aggregate levels of psychological
engagement as self-reported by employees. Gruman and Saks (2011) argue that there is very
little conceptual and empirical work on how the PM process can enhance performance by
fostering employee engagement and identify the importance of motivating the employee in
the PM process, perhaps representing significant progress in the field. Likewise, Brown,
Hyatt and Benson (2010) discover that employees with low-quality PM experiences are more
likely to be dissatisfied with their job, be less committed to the firm and have an increased
intention to quit.

Similarly, Bouskila-Yam and Kluger (2011) create a model to address the lack of stakeholder
acceptability in the PM process by creating a strength-based and goal-setting process. This
method incorporates ideas both from PM practitioners’ and from psychology scholars’
perspectives using a variation of the feedforward interview reflecting a positive self-
perception. Chang, Ferris, Johnson, Rosen and Tan (2012) consider employee acceptability of
the PM process from the perspective of how the process considers the importance of
employee self-worth, capabilities and job satisfaction. In order to improve these outcomes,
they create a practical core self-evaluation model. Hence, splitting the analysis into two
sections is useful in that it suggests that the most recent highly cited articles suggest a change
may be imminent from a firm-level approach to an employee-level approach.

***** Table 4 About Here *****

Top Ten Papers
6. Where to Now?

PM Research Development beyond PA Accuracy

Our first research question inquired: *To what extent has PM research developed beyond the narrow PA accuracy?* As denoted earlier, our findings emerged in four distinct clusters, viz 1) Firm-level, 2) Across borders, 3) Accuracy and 4) Employee acceptability. The first research question will be addressed in relation to firm-level, across borders and accuracy – three of the chief clusters. The fourth of the clusters, that is, employee acceptability, will be discussed in response to the second research questions.

1) Firm-Level

Our investigation shows that generally, the PM literature emphasis over the 20-year period has been geared toward a Firm-level analysis and that this research has expanded in various areas. For example, more research attention is given to the importance of the purpose of PM in terms of a strategic alignment and business outcome, and the emerging theme from this literature is that the PM purpose must be apparent and support the strategic goals of the firm (Maley and Kramar, 2014) and be clear to all relevant stakeholders (Biron, Farndale, and Paauwe 2011; Bouskila-Yam and Kluger, 2011; Dewettinck and van Dijk, 2013; Iqbal, Akbar and Budhwar, 2015).

2) Across Borders

Research has also advanced ‘Across borders’ and shows evidence of the growing importance of global PM and all its nuances. For instance, autonomy orientation is argued to be imperative for optimal performance of global managers (Kuvaas, 2007) and perceptions of global employees strongly link to fairness (Dewettinck and van Dijk, 2013). Individual country contextual studies support these claims, for example in India (Gupta and Kumar,
2012), Vietnam (Stanton and Pham, 2014), Botswana, (Migiro and Taderera, 2011) and China (Fee, McGrath-Champ and Yang, 2011), suggesting the ineptitude of merely utilising the PM system for administrative purposes such as pinpointing employees for promotion and salary decisions.

Moving the field forward, Engle, Dowling and Festing (2008) propose a research domain that further differentiates PM systems by considering the complexity embedded in the extent of global standardisation and local customisation in global PM. Likewise, Claus (2008) argue for PM global integration and local responsiveness. Responding to this, Maley and Kramar (2014) offer suggestions for multinational corporations to maximise the effectiveness of PM and profit simultaneously during times of global crisis by applying a Real Options approach. A four-level framework has also been conceptualised to explain how individual performance results are aggregated in multinational corporations (Engle, Festing and Dowling, 2015).

The importance of cultural sensitivity concerning PM is apparent in many studies and the significance of cultural sensitivity when implementing PM practices across borders is an essential and dominant theme (i.e. Cooke and Huang, 2011; Buchelt, 2015; Kang and Shen, 2016). Thus, in response to the first research question, the analysis and our resultant discussion suggest that the literature is now centring on Firm-level performance, albeit perhaps at the expense of an Employee-Level focus. At the same time, the Across Borders theme is alive and well, but there remains the need for more attention on the expatriate (Claus and Briscoe, 2009) and other forms of emerging global managers such as inpatriate managers, flexi-patriate managers and host country managers.

3) Accuracy
On the other hand, the ‘measurement’ research is far from dead and buried as predicted by Lawler (1994), and there remains a captivation with performance appraisal, accuracy and ratings (Adler et al., 2016; Chiang and Birtch, 2010; Iqbal et al., 2015). Issues such as conscious rating bias (see Spence and Keeping, 2011) and rater accountability (Harari and Rudolph, 2017) have once again emerged. According to Brutus (2010), the process of evaluating individual performance in organisations hinges on the use of numerical ratings, a proven yet relatively narrow operationalization of this process. Landy and Farr (1980) claim that there is nothing more to find in the measurement theme appraisal research and we propose that they may have a very valid point – nothing remarkable has appeared after another almost four decades of measurement research effort.

4) Acceptability

The second research question probes: Has the more extensive PM increased the acceptability of the process? By acceptability, we refer to acceptability in terms of all stakeholders – but namely the employee. The second research question will be addressed in relation to the final chief clusters: acceptability. The answer to the second research question is an obvious ‘no’, and this is by far one of our biggest surprises. In this light, it is indeed not surprising that there remain much conflict and tension in PM in the workplace that leads to the PM process remaining essentially unacceptable, and this issue deserves more attention (Rosen, Kacmar, Harris, Gavin, & Hochwarter, 2017). Another contributing factor to the lack of employee acceptability is the huge degree of intolerable politics throughout all aspects of the process (Poon, 2004). A lack of employee development may also be another area that has contributed to a lack of acceptability of the PM process. The latest stream of literature points to the neglect of employee development as detrimental (i.e. Rabenu and Tziner, 2016). For example, when a firm does invest in employee development, employees perceive that the
firm is supportive of their long-term career aspirations (Chiang and Birtch, 2010; Kuvaas, 2006) and the resultant outcome is increased employee commitment (Iqbal et al., 2015) and engagement (Gruman and Saks, 2011).

A shortfall in supervisor-employee feedback also emerged as pivotal to a more agreeable and acceptable PM process (i.e. Aguinis, Gottfredson and Joo, 2012; Dahling and O’Malley, 2011; Fletcher, 2001; Kuvaas, 2007; McCarthy and Garavan, 2007; Maley and Kramar, 2014; Tuytens and Devos, 2012). Nonetheless, new evidence points to authentic feedback being potentially perceived as harsh, generating adverse employee reactions (Bouskila-Yam, and Kluger. 2011; Luffarelli, Gonçalves and Stamatogiannakis, 2016) and being highly problematic for managers (Brown, Kulik, and Lim, 2016), a further factor that will surely negate the acceptability of PM.

It is evident that, to date, not enough is known about precisely why PM is not acceptable in the workplace. Multinational corporations such as Adobe, Colorcon, Dell, Deloitte, Gap, Google, Microsoft and PwC discard their PM processes, claiming they are no longer acceptable to their employees (Cappelli and Tavis, 2016) – yet this lack of tolerability emerges as vague and non-specific. At the same time, there is a repeated call to either scrap or reinvent the entire process (Buckingham and Goodall, 2015). Despite years of research – ‘discontent with PM may be at a record high’ (Adler et al. 2016 p 219; Pulakos, Hanson, Arad and Moye, 2015; Pulakos and O’Leary, 2011).

Consequently, evidence strongly implies that PM is not acceptable to many key stakeholders of today’s organisations. Albeit, while management consultants and practitioner journals such as the Harvard Business Review have been tracking the declining dilemma of stakeholder
acceptance; (cf. Cappelli and Tavis, 2016; Buckingham and Goodall, 2015), scholarly researcher appears to have been behind the eight ball. Notwithstanding, recent evidence (Levy, Cavanaugh, Frantz and Borden, 2015; Pulakos et al., 2015) may buck that tendency and could signify that academia may be beginning to wake up to the significant gap that exists between practice and research. The jury is out!

In summary, in our central analysis 20-year period of PM publications, we have identified the main research directions in the field and conclude that PM studies are yielding some new information. Additionally, we conclude that the extant PM literature has contributed to the field over the past 20 years, albeit in a limited capacity. However, there remain many areas for improvement, with several critical gaps still existing in the literature. The abundance of literature reviews and future research agenda offerings across the 20-year period are surely an indication of past and present ambiguity and a quest for knowledge? Accordingly, several suppositions can be drawn from this study. First, the literature is still mostly conceptual, and though it has developed some noteworthy models, these tend to be untested propositions. It is true that the empirical studies over the past two decades have moved away from student populations, but they have remained generally quantitative, and there still remains a scarcity of qualitative studies. Our findings support those of McKenna, Richardson and Manroop (2011), who make a case that PM research is a one-dimensional approach dominated by a positivist ontology. Rectifying the shortfall in qualitative industry studies could be another way to encourage more interaction between scholars and industry. Second, a key problem is that research centres on only a segment of the PM – the appraisal. As a result, the accuracy and performance appraisal take up a large amount of research time and effort. It is perhaps worth considering if this is disproportionate to its relevance and weighting. Thus, a PM process is needed that expands the discussion on accuracy to the wider domain of PM. It is
not only a question of doing enough empirical research, it is also the importance of its perceived value and relevance. Third, evidence suggests that enhancing the influence of research on PM practice will require much more collaboration between PM scholars and practitioners.

Limitations

This study is not without limitations. While we consider through our content analysis and rigorous bibliometric method that we have captured the majority of PM papers over the past 20 years, our method is not infallible. The bibliometric databases do not cover all research areas and do not index all publications – and one of the explicit challenges in examining the PM literature is that it is spread widely across management and I/O psychology journals. The second limitation is related to the selected database. Probably by selecting Scopus or Google Scholar, we could achieve a much bigger sample, but we decided to use the most influential and more focused database in WoS. Additionally, there can also be issues with citation bias, such as inappropriate citation of an individual’s own work, that of their colleagues, or the journals in which they publish. A number of bibliometric tools allow the exclusion of self-citations; however, it is not failsafe. Lastly, as mentioned earlier, worthy and well-cited articles could on occasion be found in non-leading journals with lower impact factor. Nonetheless, we feel confident that our approach has captured the spirit of the trends in PM literature.

7. Conclusion

In terms of practice, there are several important implications emerging from this study. Top management must understand the benefits of an efficient PM process clearly and show understanding and support for the procedure. A first and desirable step would be for CEOs to
lead by example, for instance by conducting exemplary PM processes with direct reports. This then encourages her/his executives and offers direct and concise assistance to achieve the greatest benefit from the PM process. The firm's PM process should be a feature of the annual report and a regular board room topic. An effective way to perpetuate successful PM system processes is for the CEO to identify and invest in senior executives’ and managers’ PM system training. As we have established PM as a time-intensive yet immensely rewarding process, it is understood that a reward system is needed as a complementary feature to successful PM system processes.

We consider that the problem of making a manager more proficient in managing the PM process and the problem of obtaining better employee acceptability of the process are conceptually and practically distinct. Although it makes sense to attack the problem of capability first, firms must also consider factors which affect each manager's willingness to faithfully and proficiently manage the PM. One way to encourage managers disposition towards PM is to include as part of their key performance indicators. For example, measurable improvements in the 'acceptability' of the PM over time could make one such objective. Achieving an 'acceptable' PM for all stakeholders should be written into the company mission statement and form part of the annual report.

Furthermore, clarity of the purpose of the PM should be a priority in what Bowen and Ostroff (2004) term 'HR climate strength.' The construct 'strength of the HR system' refers to a strong organisational climate in which individuals share a common interpretation of what behaviours are expected and rewarded. The rationale here is that when the purpose of the PM is clear and communicated to everyone involved in the process, the strength of the HR system can help explain how employee behaviour positively affects organisational effectiveness.
We side with the argument for practice not to slay the PM system, but to create more value in the process (see Pulakos, Mueller-Hanson, and Arad, 2019; Goler, Gale and Grant 2016; Levey, Tseng and Rosen, 2017). We also determine the viewpoint that PMs can and do serve a fundamental purpose in most organisations and, despite the challenges and controversy, they should remain in situ. Having said that, we also determine ways to improve the PM process in order to lead to enhanced employee satisfaction and motivation. Consequently, our standpoint is that PM research must now go beyond the narrow measurement arena. A final comment on implications for practice is that an ‘acceptable PM’ starts and finishes at the top with the CEO - she/he is paramount to its success.

In terms of future direction for research, having examined the different approaches of PM designs and their potential, as well as trends, we do see the need for future studies to have a wider perspective and concede that PM researchers need to extend the scope of discussion in regard to the future of performance evaluation in general. For example, future PM research should consider the implications of continuing changes in the nature and function of the firm, teamwork, working from home, flexible working hours, virtual employees, and the casualization of the workforce would be a good starting point. Another consideration for future PM studies is innovation in HR. For instance, how should PM change in the advent of the increasing impact of artificial intelligence (AI). Indeed, AI permits a computer to perform jobs that generally require human reasoning and rely primarily on the increased availability of data for prediction tasks. Cappelli, Tambe, and Yakubovich, (2019) report that 22 per cent of firms say they have adopted analytics in HR; but how this impacts PM is not at all clear and evidently an area for future investigation.
While we remain cautious, at the same time, we are optimistic concerning the future of PM. Although the attention on firm-level performance is important, it should not result in employee-level neglect. It is perhaps time that researchers make an audacious decision and divert their attention to more empirical research, in particular qualitative studies that incorporate alignment with the industry practitioners. For too long, the PM research field has been detached from industry. By following these steps, stakeholder ‘acceptability’ will almost certainly be given more attention.

Conclusively, beyond its practical content and contribution to the literature, this paper calls for more focused empirical research across academic and practitioner boundaries that will help to bridge the gap created by the early PM literature. Forty years ago, Meyer, Kay and French (1981), in their landmark study commented that everybody talks and writes about PM, but nobody has done any real scientific testing of it; we still agree with this reaction. This paper also calls for originality in scholarship, rather than further incremental development of rigid conceptual prototypes. This formula will not only enable researchers to respond to the PM abandonment crusaders by both justifying and defending PM’s role in the contemporary organisation but will also make a positive contribution to long-term organisational success. Indeed, there is still much to be done, and the field has far from plateaued.
References


|------------------|-------|-------|-------|-------|

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Cognitive</th>
<th>Interview</th>
<th>Training</th>
</tr>
</thead>
</table>

PA Methods
- Subjective appraisals
- Increasing psychometric sophistication
- Hybrids and system approaches

PA Purpose
- Administration
- Counselling and development
- Planning
- Legal

Adapted from Murphy & Cleveland (1995)

Table 1: Historical Trends of PM Research 1970–1998
<table>
<thead>
<tr>
<th>Descriptor</th>
<th>x1</th>
<th>x2</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACROSS BORDERS</td>
<td>.085</td>
<td>.390</td>
<td>208</td>
</tr>
<tr>
<td>PM</td>
<td>.146</td>
<td>.512</td>
<td>191</td>
</tr>
<tr>
<td>RESEARCH SCHEMA</td>
<td>.100</td>
<td>.207</td>
<td>177</td>
</tr>
<tr>
<td>EFFECTIVENESS</td>
<td>-.016</td>
<td>-.066</td>
<td>163</td>
</tr>
<tr>
<td>QUANTITATIVE</td>
<td>-.103</td>
<td>.137</td>
<td>157</td>
</tr>
<tr>
<td>ASSESSMENT</td>
<td>-.028</td>
<td>.442</td>
<td>142</td>
</tr>
<tr>
<td>PERCEPTIONS</td>
<td>-.263</td>
<td>-.452</td>
<td>113</td>
</tr>
<tr>
<td>STRATEGY</td>
<td>.284</td>
<td>.050</td>
<td>107</td>
</tr>
<tr>
<td>EMPLOYEES</td>
<td>-.204</td>
<td>-.181</td>
<td>94</td>
</tr>
<tr>
<td>FEEDBACK</td>
<td>-.219</td>
<td>.245</td>
<td>88</td>
</tr>
<tr>
<td>ENGAGEMENT</td>
<td>-.363</td>
<td>-.899</td>
<td>82</td>
</tr>
<tr>
<td>HRM</td>
<td>.780</td>
<td>.051</td>
<td>80</td>
</tr>
<tr>
<td>RATERT ACCURACY</td>
<td>-.990</td>
<td>.261</td>
<td>76</td>
</tr>
<tr>
<td>TENSIONS</td>
<td>-1.016</td>
<td>-.538</td>
<td>72</td>
</tr>
<tr>
<td>DEVELOPMENT</td>
<td>-.116</td>
<td>-.290</td>
<td>71</td>
</tr>
<tr>
<td>PUBLIC SECTOR</td>
<td>-.062</td>
<td>-.251</td>
<td>68</td>
</tr>
<tr>
<td>THEORIES</td>
<td>-.759</td>
<td>-.580</td>
<td>59</td>
</tr>
<tr>
<td>MNC</td>
<td>1.170</td>
<td>-.339</td>
<td>58</td>
</tr>
<tr>
<td>FAIRNESS</td>
<td>-.698</td>
<td>-.1018</td>
<td>51</td>
</tr>
<tr>
<td>GOALS</td>
<td>-.255</td>
<td>.939</td>
<td>44</td>
</tr>
<tr>
<td>COMPENSATION</td>
<td>1.485</td>
<td>-.725</td>
<td>39</td>
</tr>
<tr>
<td>KNOWLEDGE MGT</td>
<td>.624</td>
<td>-.448</td>
<td>37</td>
</tr>
<tr>
<td>ETHICS/CSR</td>
<td>-.475</td>
<td>1.800</td>
<td>30</td>
</tr>
<tr>
<td>JOB</td>
<td>-1.326</td>
<td>-1.070</td>
<td>30</td>
</tr>
<tr>
<td>CONTROL</td>
<td>-.273</td>
<td>-.114</td>
<td>28</td>
</tr>
<tr>
<td>EXECUTIVE PM</td>
<td>.193</td>
<td>-.437</td>
<td>27</td>
</tr>
<tr>
<td>GLOBALISATION</td>
<td>2.107</td>
<td>-.659</td>
<td>27</td>
</tr>
<tr>
<td>HR PROCESSES</td>
<td>1.817</td>
<td>.429</td>
<td>25</td>
</tr>
<tr>
<td>QUALITATIVE</td>
<td>.138</td>
<td>.129</td>
<td>25</td>
</tr>
<tr>
<td>ACADEMIC</td>
<td>1.842</td>
<td>-.340</td>
<td>24</td>
</tr>
<tr>
<td>BOTTOM-LINE</td>
<td>1.965</td>
<td>-.092</td>
<td>20</td>
</tr>
<tr>
<td>PSYCHOLOGICAL</td>
<td>-1.947</td>
<td>-1.080</td>
<td>17</td>
</tr>
<tr>
<td>MANAGING</td>
<td>-2.356</td>
<td>-.137</td>
<td>16</td>
</tr>
<tr>
<td>CRISIS</td>
<td>.839</td>
<td>-1.664</td>
<td>14</td>
</tr>
<tr>
<td>360</td>
<td>-1.344</td>
<td>-.377</td>
<td>13</td>
</tr>
<tr>
<td>CAPABILITIES</td>
<td>.935</td>
<td>-1.562</td>
<td>11</td>
</tr>
<tr>
<td>TQM</td>
<td>-1.022</td>
<td>2.200</td>
<td>10</td>
</tr>
<tr>
<td>IR</td>
<td>1.711</td>
<td>-.175</td>
<td>7</td>
</tr>
</tbody>
</table>

Table 2: List of Descriptors
<table>
<thead>
<tr>
<th>Cluster</th>
<th>Descriptors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td>Assessment, Quantitative, Goals, Rater Accuracy, 360, Control, Goals, Quantitative, Research Schema, Managing, Appraisal, TQM, Ethics/CSR, Goals</td>
</tr>
<tr>
<td><strong>Firm-level</strong></td>
<td>Control, Effectiveness, Bottom-line, Executive PM, HR, IR HR Processes, Managing, PM</td>
</tr>
<tr>
<td><strong>Across borders</strong></td>
<td>Strategy, Across Borders, Knowledge Management, MNC, Globalisation, HR processes, Crisis</td>
</tr>
<tr>
<td><strong>Employee acceptability</strong></td>
<td>Development, Feedback, Employees, Perceptions, Engagement, Tensions, Theories, Compensation, Psychological, Job, Capabilities, Compensation, PM, Psychological</td>
</tr>
</tbody>
</table>

Table 3: The 4 Major Clusters
<table>
<thead>
<tr>
<th>Rank</th>
<th>Article</th>
<th>Title</th>
<th>Journal</th>
<th>Method</th>
<th>Findings</th>
<th>Citations*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Levy &amp; Williams, 2004</td>
<td>The Social Context of Performance Appraisal: A Review and Framework for the Future</td>
<td>JOM</td>
<td>Literature review 1994–2004</td>
<td>This review of over 300 articles suggests field is more cognizant of the importance of the social context within which the performance appraisal operates.</td>
<td>806</td>
</tr>
<tr>
<td>2</td>
<td>Gruman &amp; Saks, 2011</td>
<td>Performance management and employee engagement</td>
<td>HRMR</td>
<td>Conceptual</td>
<td>Examines critical key drivers of employee engagement and suggests a new perspective for fostering and managing employee engagement to increase job performance.</td>
<td>732</td>
</tr>
<tr>
<td>3</td>
<td>Fletcher, 2001</td>
<td>Performance appraisal and management: The developing research agenda</td>
<td>JOOP</td>
<td>Conceptual</td>
<td>The impact of multi-source feedback systems and? of cultural differences is likely to attract the research effort. However, research is needed to give practical guidance on conducting appraisals.</td>
<td>676</td>
</tr>
<tr>
<td>4</td>
<td>Den Hartog, Boselie &amp; Paauwe, 2004</td>
<td>Performance management: A model and research agenda</td>
<td>Applied Psychology</td>
<td>Conceptual</td>
<td>This paper presents a model for PM combining insights from strategic HRM and I/O psychology and adds to previous models by explicitly incorporating employee perceptions, the role of direct supervisors and possible reversed causality.</td>
<td>464</td>
</tr>
<tr>
<td>5</td>
<td>Kuvaas, 2006</td>
<td>Performance appraisal satisfaction and employee outcomes: mediating and moderating roles of work motivation</td>
<td>IHRM</td>
<td>Quantitative</td>
<td>PA satisfaction was directly related to effective commitment and turnover intention.</td>
<td>433</td>
</tr>
<tr>
<td>6</td>
<td>Poon, 2004</td>
<td>Effects of appraisal politics on job satisfaction and turnover intention</td>
<td>PR</td>
<td>Quantitative</td>
<td>Results indicated that when employees perceive performance ratings as being manipulated they express reduced job satisfaction that leads to increased intention to quit.</td>
<td>340</td>
</tr>
<tr>
<td>7</td>
<td>Brown, Hyatt &amp; Benson, 2010</td>
<td>Consequences of performance appraisal experience</td>
<td>PR</td>
<td>Quantitative</td>
<td>Employees with low-quality PA experiences were more likely to be dissatisfied with their job, be less committed to the firm and have increased intention to quit.</td>
<td>227</td>
</tr>
<tr>
<td>8</td>
<td>Latham, Almost, Mann &amp; Moore, 2005</td>
<td>New developments in PM</td>
<td>Org Dynamics</td>
<td>Conceptual</td>
<td>Finds that coaching ensures a highly trained, highly motivated workforce.</td>
<td>220</td>
</tr>
<tr>
<td>9</td>
<td>Kuvaas, 2007</td>
<td>Different relationships between perceptions of developmental performance appraisal and work performance</td>
<td>PR</td>
<td>Quantitative</td>
<td>Autonomy orientation moderated the relationship between perceptions of developmental PA and work performance.</td>
<td>155</td>
</tr>
<tr>
<td>10</td>
<td>Biron, Farndale &amp; Paauwe, 2011</td>
<td>Performance management effectiveness: lessons from world-leading firms</td>
<td>IHRM</td>
<td>Qualitative</td>
<td>Creates a framework proposing to improve PM involving: strategic elements; senior managers’ involvement in communicating and training performance raters.</td>
<td>137</td>
</tr>
</tbody>
</table>

Table 4: Top Ten Papers
Figure 1 Structure of the Selected Research Area

Abbreviations in map:
Aborders- across borders; Caps-capabilities; Develop-development; Effective- effectiveness; Exec PM- executive PM; HR -human resource processes; Percept- perceptions; Pub Sec- public sector; Schema- research schema; Strat-strategy; Qual-qualitative; Quant- quantitative.