University Library



A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Al-Nawab, Hadeel

Title:
Evaluation for Evidence-Based Performance Management: Understanding and Measuring Performance Managers' Perceptions

Date:
2020

Persistent Link:
https://hdl.handle.net/11343/243009

Terms and Conditions:
Terms and Conditions: Copyright in works deposited in Minerva Access is retained by the copyright owner. The work may not be altered without permission from the copyright owner. Readers may only download, print and save electronic copies of whole works for their own personal non-commercial use. Any use that exceeds these limits requires permission from the copyright owner. Attribution is essential when quoting or paraphrasing from these works.

Evaluation for Evidence-Based Performance Management: Understanding and Measuring Performance Managers' Perceptions

Hadeel Faisel Al-Nawab

https://orcid.org/0000-0002-8287-6553

Submitted in total fulfilment of the requirements of the degree of Doctor of Philosophy

April 2020

Melbourne Graduate School of Education

The University of Melbourne

Abstract

A common claim is that high-performing organizations use evidence-based practice to manage staff performance, herein called performance management. The literature showed that the implementation of performance management policies is crucial because even well-designed performance management models fail if they are not implemented as intended (Armstrong, 2015). Given that behavior can be mediated by perception, this thesis focused on the perceptions held by implementers of performance management that might mediate their implementation of performance management policies. This is important because, despite the research on evidence-based performance management, there remains a gap in understanding and measuring the perceptions held by the implementers. Moreover, there is a sizeable gap between performance management research and practice due to a plethora of obstacles like poor access to reliable research. Three sequential research stages were conducted focusing on item generation, scale development and scale refinement and validation, respectively. An initial set of 130 items was developed, based on a thematic analysis in the narrative literature review and a scoping literature review study (Stage 1). These items were reduced to a set of 55, then 41 items in Stages 2 and 3, respectively, using exploratory and confirmatory factor analyses, Cronbach's alpha calculations and examination of conceptual importance of data from a combined 589 survey respondents. Furthermore, the effect of several antecedent factors was tested using multivariate analysis of variance.

The thesis findings enhanced the extant understanding by presenting perception of performance management as a broad concept encompassing two higher order factors (Perception of Efforts and Perception of Results) and eight lower order factors (perception of Performance Evaluation, Documentation, Organizational Support, Supervisory Support, Climate, Turnover,

Critical staff Withdrawal and New staff Withdrawal). The findings also added to the literature on the effect of antecedent factors, particularly age, work role, organizational size, industry sector and workforce experience. Finally, the thesis further narrowed the gap between research and practice by creating a theoretically grounded and empirically validated performance management perceptions scale for organizations to use in in evidence-based performance management.

Declaration

This is to certify that:

1. This thesis comprises only my original work towards the degree of Doctor of Philosophy except where indicated in the preface.
2. Due acknowledgement has been made in the text to all other material used.
3. This thesis is fewer than 100 000 words in length, exclusive of tables, maps, bibliographies and appendices.

Preface

1. This thesis is comprised of my own original work. I have undertaken the data collection, analysis, and written all the content presented within this thesis.
2. None of the work presented has been submitted for other qualifications.
3. No third party editorial assistance, either paid or voluntary, was used to produce this thesis. The supervisory team provided detailed feedback, commentary, and edits, and the University of Melbourne Academic Skills Team provided advice about structure and voice based on earlier drafts.
4. As a domestic graduate researcher I received a fee remission scholarship from the Australian Government.

Acknowledgements

Academic Team

Professor John Polesel, Advisory Committee Chair

Many thanks to John for his guidance and for creating a safe space for meaningful discussion on my perspective of the PhD process. He came on board on short notice because my previous Chairperson left the university abruptly, and he was incredibly kind in taking on this extra work.

Associate Professor Shelley Gillis, Co-Supervisor

Shelley was valuable in many ways, most of all was helping me to navigate through the administrative labyrinth that accompanied the PhD. She was also kind enough to remind me, constantly, that accountability goes both ways, and that I am well within my right to hold her accountable to deadlines as much as she did with me.

Laureate Professor John Hattie, Principle Supervisor

It is safe to say that this thesis would not have come to fruition without John. I had worked with him professionally prior to asking him to join my supervision team and knew that we could work well together. But I had no idea just how amazing it would be. He received a jaded, tired, skeptical student who was at the end of her tether and ready to leave it all behind to join the circus.

Once he came on board, I finally felt like a PhD candidate who was doing *actual* research and discovering new things. Despite all the literature I had read on the power of feedback, I only truly felt it when I started working with John. His straight-shooting, constructive and engaging feedback could pull any person out of the proverbial ‘pit’ and help them regain perspective. I am forever in his debt.

Family and Friends

I thank my mom, Dr. Alia Al-Beer, for her continued support and prayers, and, along with my late father, Dr. Faisel Al-Nawab, for risking their necks (literally) to pull us out of a war zone to safety and a better life.

I thank my sister Teeba Al-Nawab for many things, most of all is procreating, because now my mother has grandchildren, and I can be childfree. It was with great parenting that she and her husband Peter Neale raised three wonderful children. The holidays would have been insufferable had they given life to tiny terrible terrors. Lucky for me, they are tiny terrific tikes, who are only terrible when their iPads are put away and they have to eat broccoli.

I thank my surrogate sister Dr. Anna Dabrowski and my two nieces for keeping me in check whenever the anxiety would take over.

I thank my partner in crime, Kellie Picker, who encouraged me to keep going when matters seemed hopeless.

I thank my friends Timoci O'Connor, Grace Kenwrick, Dr. Amy Berry, Dr. Luke Rowe, Dr. Daniel Cloney, and Scott Bolton for more things than I can count. They have made my life in Australia and my PhD work worth the blood, sweat and tears.

I thank Dr. Rosie Anderson for taking me out of my comfort zone, and making sure I keep flying the plane whilst learning to fly.

Last but not least, I thank my husband, Dave Matheson; the handsome fool who agreed to marry me. How he put up with the messy house, the mood swings and, the woman who transformed into a gremlin over the last eight months, I will never know! My guess is that he is either a monk disguised as a bearded musician, or simply a man with very low self-esteem. Either way, I am grateful that our paths crossed five years ago.

Dedication

For Dave

Table of Contents

Abstract ii

Declaration iv

Preface v

Acknowledgements vi

Dedication viii

Table of Contents ix

List of Tables xiv

List of Figures xv

List of Acronyms xvii

Chapter 1: Introduction 1

1. Context 2
2. Research Problem 9
3. Epistemology and Theoretical Perspective 14
4. Significance 16
5. Research Scope, Goals and Questions 18
6. Thesis Organization 22

Chapter 2: Literature Review 23

1. Organizational Performance 23
2. Human Resource Management 25
   - 2.1 Strategic Alignment 27
   - 2.2 HRM Strength 28

2.3 Theoretical Underpinnings 29
2.4 Evidence-Based Management 36
3. Performance Management 49
3.1 Requirements of Effective Staff Performance Management 50
3.2 Performance Metrics 57
4. Evaluation 61
4.1 Participatory Evaluation 66
4.2 Utilization-Focused Evaluation 69
4.3 Theory of Change 72
4.4 The Personnel Evaluation Standards 75
5. Perception and Performance Management Implementation 85
5.1 The Perceiver. 88
5.2 The Situation. 94
6. Perception Targets 97
6.1 Perception of Evidence-Based Practice 98
6.2 Perception of Organizational Communication 99
6.3 Perception of Organizational Support 102
6.4 Perception of Work Engagement 103
6.5 Perception of Others' Behaviour 106
6.6 Perception of Propriety, Utility, Feasibility and Accuracy. 110
7. Performance Management Perception Logic Model 110
8. Chapter Summary 113
Chapter 3: Methods 115

1. Research Design 115
1.1 Stage 1: Item Generation 117
1.2 Stage 2: Scale Development 120
1.3 Stage Three: Scale Refinement and Validation 120
2. Study One: Scoping Literature Review 121
2.1 Study Purpose 121
2.2 Inclusion Criteria 121
2.3 Search Terms and Databases 122
2.4 Search Strategy 123
2.5 Publication Selection and Data Charting 124
3. Study Two: First Online Survey 124
3.1 Study Purpose 124
3.2 Study Participants 124
3.3 Data Collection 126
3.4 Data Analysis 128
4. Study Three: Second Online Survey 131
4.1 Study Purpose 131
4.2 Participants 131
4.3 Data Collection 133
4.4 Data Analysis 133
5. Ethics and Funding 135
6. Chapter Summary 135
Chapter 4: Scoping Literature Review 137

1. Purpose 137
2. Search Strategy Results 138
3. Results 140
3.1 Publication Description and Characteristics 140
3.2 Relevant Scales 144
4. Items for the EMI Scale 148
5. Chapter Summary 166
Chapter 5: First Survey Results 168
1. Purpose 168
2. Data Collection and Coding 168
2.1 Demographics 169
3. Data Analysis 170
3.1 Exploratory Factor Analysis 170
3.2 Multivariate Analysis of Variance 183
4. Construct Comparison 193
5. Chapter Summary 197
Chapter 6: Second Survey Results 198
1. Purpose 198
2. Data Collection 198
2.1 Demographics 199
3. Data Analysis 203
3.1 Confirmatory Factor Analysis 203
3.2 Multivariate Analysis of Variance. 210

4. Chapter Summary 230
Chapter 7: Discussion 232
1. Research Problem and Thesis Scope 233
2. Research Summary 236
3. Substantive Contributions 242
3.1 Performance Management Perception Targets 242
3.2 The EMI Scale 247
3.3 Antecedents 252
3.4 Perceiver Characteristics 253
3.5 Situation Characteristics 263
4. Future Directions 266
5. Concluding Comments 273
Reference List 275
Appendix A: Scoping Study Results Charting 332
Appendix B: Survey One 337
Appendix C: Survey Two 346
Appendix D: The EMI Scale 352
Appendix E: EMI Scale – Item Orientation 355
Appendix F: EMI Scale – Item Subject 357

List of Tables

Table 1. Survey 1 Factor Pattern Matrix……..176
Table 2. Survey 1 Factor Correlations…….181
Table 3. Survey 1 Factor Means, Standard Deviations and Estimates of Reliability……..182
Table 4. Survey 1 Statistically Significant Effects of Work Role……184
Table 5. Survey 1 Statistically Significant Effects of Location……...187
Table 6. Survey 1 Statistically Significant Effects of Organization Size…….190
Table 7. Survey 1 Non-Statistically Significant Effects……..193
Table 8. Survey 2 Factor Means, Standard Deviations and Estimates of Reliability……..206
Table 9. Survey 2 Factor Correlations…….207
Table 10. Survey 2 Statistically Significant Effects of Work Role……..211
Table 11. Survey 2 Statistically Significant Effects of Participant Age……...214
Table 12. Survey 2 Statistically Significant Effects of Organization Size……...221
Table 13. Survey 2 Statistically Significant Effects of Workforce Experience……...224
Table 14. Survey 2 Non-Statistically Significant Effects on the Eight Lower Order Factors….229
Table 15. Survey 2 Non-Statistically Significant Effects on the Two Higher Order Factors…..229

List of Figures

Figure 1. Conceptual Hierarchical Relationships in Organizational Performance……..…..19
Figure 2. Framework of Evidence-Based Management……20
Figure 3. Components of an HRM System……...….26
Figure 4. The Propriety Standards…….……....78
Figure 5. The Utility Standards……..…79
Figure 6. The Feasibility Standards…….……..80
Figure 7. The Accuracy Standards…….....82
Figure 8. Simplified Logic Model of the Conceptual Perception Targets and Characteristics of the Perceiver and the Situation…….……....112
Figure 9. Research Design……...116
Figure 10. Item Generation Steps……..……..118
Figure 11. Overview of Research Stages and Empirical Studies…….136
Figure 12. Results of the Scoping Literature Review Search Strategy……139
Figure 13. Items for Perception of Evidence-Based Practice……..152
Figure 14. Items for Perception of Organizational Communication……154
Figure 15. Items for Perception of Organizational Support…….156
Figure 16. Items for Perception of Work Engagement……158
Figure 17. Items for Perception of Others' Behavior……..160
Figure 18. Items for Perception of Propriety……...162
Figure 19. Items for Perception of Utility……163
Figure 20. Items for Perception of Feasibility…….164
Figure 21. Items for Perception of Accuracy……...165
Figure 22. Survey 1 Exploratory Factor Analysis Scree Plot……..172
Figure 23. Survey 1 Statistically Significant Effects of Work Role……186
Figure 24. Survey 1 Statistically Significant Effect of Participant Location…….……..189
Figure 25. Survey 1 Statistically Significant Effects of Organization Size…….192
Figure 26. Relationship between the Hypothesized Conceptual Performance Management Perception Targets and the Eight Latent Factors……...……..196
Figure 27. Survey 2 Respondent Demographics; Gender and Employment Status…….200

Figure 28. Survey 2 Respondent Demographics; Gender and Workforce Experience……202
Figure 29. Survey 2 Standardized Paths of the Eight-Factor Model……...205
Figure 30. Survey 2 Standardized Paths for the Higher Order Factor Model…….……209
Figure 31. Survey 2 Statistically Significant Effects of Work Role……213
Figure 32. Survey 2 Statistically Significant Effect of Participant Age……..216
Figure 33. Survey 2 Statistically Significant Interaction Effect of Age Group and Gender on Perception of Results……...218
Figure 34. Interaction Effect of Gender and Employment Status on Perception of Documentation……220
Figure 35. Survey 2 Statistically Significant Effect of Organization Size…….....223
Figure 36. Survey 2 Statistically Significant Effect of Workforce Experience……...226
Figure 37. Survey 2 Statistically Significant Effects of Industry Sector on the Lower Order Factor Model……228
Figure 38. Summary of Research Stages and Results……..241
Figure 39. Simplified Logic Model of the Empirically Tested Factors……...249

List of Acronyms

AHRI	Australian Human Resource Institute
AMOS	Analysis of Moment Structures
ANOVA	Analysis of Variance
CEBMa	Center for Evidence-Based Management
CFI	Comparative Fit Index
CMIN/DF	Chi-square Mean/Degree of Freedom
EMI Scale	Evaluation for Evidence-Based Management Scale
HR	Human Resources
HRM	Human Resource Management
KMO	Kaiser-Meyer-Olkin Measure of Sampling Adequacy
MANOVA	Multivariate Analysis of Variance
ProA	Prolific Academic
RMSEA	Root Mean Square Error of Approximation
SEM	Structural Equation Modelling
SPSS	Statistical Package for the Social Sciences
SRMR	Standardized Root Mean Square Residual
UWES	UTRECHT Work Engagement Scale

Chapter 1: Introduction

Technological advancement, changes in consumer demographics, and labor migration continue to force organizations to adapt to ensure sufficient performance within their respective market. Organizational performance describes an organization's capacity to achieve its objectives in a sustainable, reliable and, measurable way through intentional actions (Hunter & Nielsen, 2013). In general, organizations use four types of resources; physical, technological, financial, and human. The shift in the dominant economy from manufacturing to knowledge services led to an increased reliance on Human Resource Management (HRM) for two reasons: (a) human resources constitute a significant portion of organizational costs but also contribute the most revenue (Cascio, 1991; de Brito & de Oliveira, 2016; Tracey & Hinkin, 2008) and, (b) the intellectual contribution of knowledge workers requires a more nuanced approach to staff performance management, herein called performance management. Despite the proverbial 'black box' in determining causality, the literature suggested the presence of a noteworthy link between HRM and organizational performance because staff performance is intrinsically linked to the performance of an organization (Den Hartog et al., 2013; Katou, 2015; Vermeeren et al., 2014). Given the complexity of staff behavior, the need for evidence-based practice in HRM is rising at exponential rate in the literature, though not as much in practice (Barends et al., 2017). In general, performance management, be it evidence-based or otherwise, can fail in two ways; design and / or implementation (Armstrong, 2015; Armstrong & Taylor, 2014). This thesis focused on perception as one of the main influencers of implementation behavior, because as Armstrong and Taylor (2014) observed, even well designed policies can be poorly implemented if they are perceived negatively. This thesis aimed to contribute to evidence-based management,

using lessons from evaluation, specifically, by enhancing the knowledge of performance managers' perceptions of their organization's performance management approach. This chapter introduces the research context, problem, scope, goals and questions. It follows with an outline of the remaining six chapters within the thesis.

1. Context

The Economy. Economic shifts over the last century have redistributed market power, changed the dynamics of competition and increased employment opportunities on a global scale (Brown et al., 2017; McGovern et al., 2007). In most developed countries like Australia, The United States, and The United Kingdom, the dominant economy became what is known as the 'knowledge' economy in which cognitively complex services like scientific research, finance, and medicine overshadowed manual industries like farming or manufacturing (Curd, 2011; Kellerman, 1985; Kenessey, 1987; Thakur, 2011). There are several noteworthy characteristics of knowledge services: (a) intellectual property is the most valuable commodity, despite a lack of physical embodiment, (b) the workforce is highly educated, and (c) even the design of tangible goods requires more technical skills and innovation than ever before (McGovern et al., 2007). This shift in the economic structure has led to the exportation of knowledge services and the outsourcing of manual labor to other countries like China or India. Moreover, technological advancement over the last few decades has led to the globalization of the labor market so, labor migration, especially that of highly skilled labor, has never been so easy and increased the number of companies in competition for market dominance (Bloom et al., 2011; Dimaya et al., 2012; McGovern et al., 2007).

In response to the changes in market structure, many organizations started to adapt through cost reduction measures; typically via increased automation, outsourcing to countries with cheaper labor, internal downsizing, and increasing their market flexibility through non-standard employment, like fixed-term contracts (Kim et al., 2012; McGovern et al., 2007; Vosko et al., 2009). Though effective in reducing operational costs in the short term, such responses tend to damage market power in the knowledge industry as they can disengage talented staff because of poor perceptions of organizational support (Boudreau & Ramstad, 2002, 2007; Breuer et al., 2011). Within the context of the knowledge economy, talented staff are those that are highly skilled in problem-solving relationship-building and innovation, as such, they tend to be highly educated and demand more income than those who work in manual labor (Brown et al., 2017). This can be problematic for organizations in an expanding global market in terms of attracting and keeping such workers. Cost-cutting measures tend to increase the risk of employee disengagement, burnout, and withdrawal from any organization rises, but this is a greater threat with talented knowledge workers because they are highly employable, and their option for early retirement is significant due to the financial security they have from their relatively high income (Brown et al., 2017). Therefore, as employee commitment lowered and the risk of losing talented staff increased, some organizations began to emphasize the use of HRM.

Despite the call for business models to abandon downsizing and focus on staff development, labor cost-cutting remains pervasive in many organizations, even those with extensive HRM systems, be it in the form of internal restructuring, redundancy or, the overuse of precarious employment (Kalleberg, 2003; Robertson, 2016). This suggests that, while employers recognize the benefit of HRM interventions like performance management, they still value market flexibility and may invest their HRM efforts strategically with some employees more so

than with others (McGovern et al., 2007). For example, secure employment tends to be more common in work roles that require talented staff with complex tasks and, therefore, are difficult to define, monitor and manage.

HRM. The move from tangible products to intangible contributions to organizational performance pushed HRM, and performance management in particular, to become more of a core business function than it was in its infancy during the early 20th century (Armstrong & Taylor, 2014). It also introduced a new level of complexity in performance management, which, in turn, led to an upsurge in theories, models, practices and requirements for performance data. At its core, the shift in HRM is connected to an increased appreciation for the value of staff and their skills leading to a move from control to commitment-based HRM. The underlying premise of commitment-based HRM is that performance management will work in concert with organizational performance management to improve an organization's competitive power better than control practices (Appelbaum et al., 2000; Armstrong, 2015). As McGovern et al. (2007) stated, persuading employees to believe in and commit to their organization's goals tends to be far more effective than commanding them to obey rules based on organizational goals.

In essence, the increased reliance on knowledge workers mandated a change in the way staff are valued as well as how their performance is defined, measured and managed (Beatty et al., 2003). Over the years, within scholarly literature, the concept of staff performance became more nuanced as indicated by the surge of new terms like extra-role behavior, organizational citizenship behavior, commitment behavior, and withdrawal behavior. To clarify, for the purposes of this thesis, staff performance pertains to employee behavior that is required to achieve an organization's goals, typically documented in job descriptions (Jalalkamali et al., 2016).

Performance Management. Performance management is an HRM initiative used to ensure the continual achievement of organizational performance through sustainable stellar employee performance (Armstrong, 2015). The implementation of staff performance management involves various players like executives, HRM professionals, and performance managers (Op de Beeck et al., 2016).

A common finding within the substantial body of literature dedicated to HRM, performance management and employee behavior is that performance management must adopt evidence-based practices in order to be effective (Armstrong, 2015; Barends et al., 2014; Society for Human Resource Management, 2017). Another recurrent finding is that organizations with effective performance management tend to have superior levels of employee and organizational performance than those that do not (Beatty et al., 2003; Huselid et al., 1997; Kehoe & Collins, 2017; Kehoe & Wright, 2013). Importantly, the shift in HRM has seen a commensurate increase, in academia, in the value of performance management over the sole use of performance measurement.

According to the Australian Human Resource Institute (AHRI), performance management has six components; policies, objectives and standards, appraisal, management of underperformance, reward, and development (Australian Human Resources Institute, 2019). Therefore, performance management is, fundamentally, a resource-intensive intervention, requiring time and financial investment, which some companies may not be willing or able to devote (Baker, 2013; Falcone & Tan, 2013). Furthermore, considering that there are no government mandates or accountability structures, at least in Australia, for proper implementation of performance management policies, it is not surprising to find idiosyncratic management practices, both within and between organizations (West & Blackman, 2015). Recent

research found that one of the common features of well performing organizations is having staff who perceive performance management as a legitimate part of their core business and effective in enhancing their performance (Blackman et al., 2014; West & Blackman, 2015). It is unknown, however, whether favorable perceptions lead to positive performance management outcomes or vice versa, nonetheless, as the literature indicated, the two are positively correlated enough to merit research into performance management perceptions. There are, unfortunately, troubling trends in the findings of such research. For example, recent results of the Australian Public Service employee census showed that employee confidence in performance management was low, particularly in its overall effectiveness (Australian Public Service Commission, 2018). Moreover, according to a recent census from the Australian Public Service, only 45% of the 103,137 respondents felt that their overall experience of performance management was useful for their own performance (Australian Public Service Commission, 2018). These findings supported similar ones from an AHRI survey conducted a decade earlier, which showed that almost half of the 1,605 respondents perceived their organization's performance management to be, at best, somewhat effective, while a third perceived it to be ineffective (Australian Human Resources Institute, 2008, n.d.-a). Though it is unknown whether the AHRI provided a definition of effectiveness in the survey, a shared understanding of effectiveness between the respondents can, at least, be assumed as all the respondents were HR professionals and members of the AHRI.

The AHRI survey also found that the perceived fundamental purpose of performance management was not performance improvement, as it should have been (Australian Human Resources Institute, 2008). In particular, 46% of the 1,605 respondents to the AHRI survey stated that the main purpose of performance management was the strategic alignment of their performance indicators with organizational goals (Australian Human Resources Institute, 2008).

In addition, almost one in five respondents stated that the main role of performance management was, simply, the measurement of staff performance (Australian Human Resources Institute, 2008). Such findings are concerning because they indicate that many employees perceive performance management to be, primarily, a one-sided intervention with questionable effect, instead of an effective intervention with benefits to both the organization and the employee. Of critical concern is that these perceptions are not new; data from 2009, 2010, and 2011 showed similar trends of performance management perceptions (Blackman et al., 2012). In Australia, a decade's worth of results from the annual State of the Service Reports indicated that, although well-designed performance management systems may be in place, there was considerable room for improvement (Blackman et al., 2014; Blackman et al., 2012).

Such findings are not exclusive to Australia, for example, in the United States, a survey from Watson Wyatt[1] showed that a mere 30% of employees perceived their organization's performance management as beneficial for their performance (Pulakos et al., 2008). In addition, a survey conducted in 2010 with 750 WorldatWork[2] members, mainly senior HR managers,

[1] Watson Wyatt Worldwide, Inc. was a global management consulting business which merged with Towers Perrin in 2010 to forge Towers Watson.
[2] WorldatWork is a global non-profit association for HRM professionals.

found that while many believed their organization's performance management to be important, they perceived numerous challenges to its success in achieving the desired staff performance outcomes (WorldatWork & Sibson Consulting, 2010). The challenges noted in the survey included poor goal-setting and managers' hesitation to have, potentially, difficult performance discussions with their staff (WorldatWork & Sibson Consulting, 2010).

A focus on perception is important because poor perceptions of performance management are positively correlated with poor performance behavior; authors like Falcone and Tan (2013), noted that the disconnection between performance management models and real-life implementation behavior as a significant contributor to meager perceptions and performance outcomes. This position is prominent among many HRM scholars and practitioners (Armstrong, 2015; Bandaranayake, 2003), in particular, it was stated that partial implementation of performance management is common. It is also widely known that the most commonly implemented component of performance management is the annual performance review, which, when implemented in the guise of performance management, rarely leads to desirable performance outcomes commensurate to its operating costs (Falcone & Tan, 2013; Kruse, 2012).

Several reasons for such partial implementation were presented over the last few decades in both academic and professional literature. First, because annual performance reviews have existed for at least 60 years prior to contemporary performance management, they became indoctrinated in organizational functioning to such a point that they are seen as analogous to performance management (Baker, 2013; Falcone & Tan, 2013; Seiden & Sowa, 2011). Second, the legal obligations in performance management are also heavily focused on performance appraisal and mitigating underperformance rather than performance management and improvement in performance (Fair Work Ombudsman, n.d.). This leads to management practices

for the sake of legal compliance rather than effectiveness, specifically because documents associated with performance appraisals provide a basis for legal defense of disciplinary actions, should the need arise (Armstrong & Baron, 1998; Baker, 2013). Third, performance management, no matter how well designed, remains contingent on the staff tasked with its implementation, herein called implementers or performance managers. To elaborate, if implementers lack the skills, motivation or positive perceptions of performance management, they are far less likely to implement performance policies as intended (Australian Human Resources Institute, 2019; Den Hartog et al., 2004). In essence, implementers need to trust that their efforts in fully implementing their organization's performance management policies will lead to comparable value. The aforementioned reasons for partial implementation of performance management bring forward the argument for understanding and measuring perception as it can influence implementation behavior. As such, efforts into evidence-based performance management should consider the implementers' perceptions, of which there are three components; the perceiver, the target, and the situation. These components were focal points in this thesis and are discussed further in this chapter as well as Chapter 2.

2. Research Problem

Proponents of evidence-based management suggested that a robust evidence-basis increases the likelihood of accurate, relevant and effective performance management (Barends et al., 2014; Kovner, 2012; Rousseau & Barends, 2011). To be evidence-based, performance management needs useful information from a variety of reliable sources to inform decision-making (Barends et al., 2014). Reliable sources include academic research and evaluation data gathered by external or internal evaluators. Despite the well-known importance of evidence-

based management in practice, it is not advancing at the rate it should (Armstrong, 2015; Barends et al., 2012; Bedford et al., 2009).

Challenges. Indeed, evidence-based management faces significant challenges; for one, certification or an academic qualification in management is not mandatory so knowledge of evidence-based management can be vastly diverse (Rousseau, 2012; Rousseau & Barends, 2011). Second, the use of reliable information is variable, in fact, the partial use of academic research in performance management is well documented world-wide due to issues of access and translation (Bezzina et al., 2017; Sanders et al., 2008). Third, as research has shown, there is insufficient use of evaluation approaches and data to inform performance management decisions (Dopson et al., 2012; Marshall, 2016; Pattyn, 2014). This is especially concerning because evaluation can be done internally or through external consultants, and it is highly adaptable to the specific contexts, needs and real-life situations (Alkin, 2013).

Evaluation for Evidence-Based Practice. Evaluation refers to a systematic inquiry used to form educated judgments of value, effect, or significance to support decision-making (Cousins et al., 2014). Its uses extend from program improvement (Lobo et al., 2010; Lobo et al., 2014), to evidence-based policy and practice (Edwards et al., 2016; Naccarella et al., 2007; Schwarzman et al., 2018). While the effectiveness of performance management has been a focal point in the literature, the approaches to ascertaining effectiveness refer to measurement, which differs from true evaluation models, which relate to improvement (Alkin, 2013; Armstrong, 2015; Armstrong & Taylor, 2014; Carden & Alkin, 2012; Christie et al., 2005).

Several reasons were proposed for the limited use of evaluation in performance management, including; insufficient support from organizational leaders, issues of methodological knowledge, and measurement difficulty (Derven, 2012; Griffin, 2012; Marshall & Rossett, 2014). Numerous findings have also shown resource limitations as significant barriers to evaluation use, not just in terms of financial resources but also time and skills needed to conduct evaluation (Brug et al., 2011; Francis & Smith, 2015). Further, poor or erroneous perceptions of evaluation were cited as barriers to its use as a source of evidence for evidence-based management (Lobo et al., 2010; van Koperen et al., 2016).

In general, unfavorable perceptions of evaluation stem from flawed information about evaluation within the performance management discipline and negative experience with evaluation (Azzam & Levine, 2015; Chouinard, 2013). Evaluation, as a term, has been used erroneously within much of the performance management literature. For example, Armstrong's (2015) work stressed the importance of evaluation in performance management but the information presented related to measurement practices not evaluation. Furthermore, experience with evaluation can vary greatly, for one, the benefit of evaluation in practice depends on the purpose of its use. To elaborate, if evaluation is implemented as a function of accountability instead of improvement, it can be perceived as the proverbial 'witch hunt', thereby reducing the likelihood of meaningful engagement with the evaluation process and data (Cousins et al., 2013; Fournier, 2005; Mathison, 2005; Shadish & Luellen, 2011). It is important to acknowledge that the aforementioned barriers to evaluation usually occur together, for instance, limited leadership support could be due to budgetary constraints, poor perceptions of the value of evaluation, or both. Whatever the case may be, it has long been understood that the extent of evaluation use in

organizations can be highly dependent on perceptions and organizational politics (Azzam & Levine, 2015; Chelimsky, 1995; Ferris et al., 2008; Weiss, 1993).

Data for Evaluating Performance Management. Performance management, unlike other organizational functions such as marketing or finance, does not produce the analytics or metrics sufficient for an evaluation of its effectiveness (Fink & Sturman, 2017). Hence, data from staff performance reviews are often used as the barometer of the effectiveness of performance management (Fink & Sturman, 2017; Handa & Garima, 2014; Marler & Boudreau, 2017). Whilst performance appraisals are useful sources of data, they cannot be the sole indicator of performance management effectiveness because they provide a unidimensional view of a multidimensional phenomenon (Baker, 2013; Boris & Kopczynski Winkler, 2013; Boudreau & Ramstad, 2007; Carlson & Kavanagh, 2018).

Since the 1980s research has been devoted to HR metrics leading to advances in HR data to give organizations more decision support mechanisms (Riley et al., 2012). In addition to staff performance results, general HRM metrics now include administrative costs per employee, time needed to fill a vacant position, staff satisfaction surveys, turnover rates, unplanned lateness rates, unplanned absence rates, level of staff qualification, and the number of interviews per hire (Boudreau & Ramstad, 2002). Though helpful, these metrics are predominantly used in a summative form of inquiry which reduces their utility. To elaborate, whilst most metrics give data on one outcome or another, some of them have predictive capacity but only if they are used in a timely fashion (Marler & Boudreau, 2017). For example, using a survey to assess staff perceptions of fairness in the workplace gives the same information whether it is administered to current employees or soon to be former employees (typically through a staff exit survey). However, the timing of such a survey has implications on the utility of its data. Specifically,

assessing fairness perceptions from current employees gives an organization the opportunity to intervene, whereas little can be done for soon to be former employees. Indeed some metrics, like end-of-year financial reports, lend themselves more to summative inquiry, however, metrics like unplanned absence reports can be used several times during a financial year, and give an organization the opportunity to reduce such withdrawal behavior before the year is completed. This is, fundamentally, the difference between summative and formative evaluation (Alkin, 2013; Carden & Alkin, 2012). Given that employee behavior is a complex construct and that there are numerous performance management practices needed to manage it, performance management has, understandably, lagged behind in the analytics domain.

Despite the barriers to evidence-based performance management, the need for evidence remains prevalent in management discourse. As Bedford et al. (2009) showed, high-performing organizations use more and various measures to examine staff performance than lower performing organizations. Nevertheless, there is a gap in formative evaluation of performance management effectiveness, specifically in evaluative inquiry into the implementation of performance management. As proponents of formative evaluation stress, maximum potential for improvement is better achieved through evaluation of processes as well as outcomes (Mathison, 2005, 2008; Shadish & Luellen, 2011). In this case, instead of only measuring the outcomes of performance management implementation, it would be prudent to also investigate its potential drivers to allow for timely intervention (Mertens & Wilson, 2012). This thesis focused on implementers' perceptions of performance management that may influence implementation and aimed to identify the antecedent factors that may influence such perceptions. Given the many challenges to evaluation use for evidence-based performance management, this research also aimed to create a theoretically grounded and empirically tested scale for measuring perceptions

of performance management, for organizations to use in evaluation for evidence-based performance management. In this way, it intended to narrow the gap between academic research and performance management in practice.

3. Epistemology and Theoretical Perspective

This research was guided by empiricism, theory of change, and utilization-focused evaluation. The empiricist epistemology emphasized the role of empirical evidence, which formed the foundation of knowledge acquisition in this research. Because of the aforementioned obstacles to evaluation use and evidence-based performance management, a theory of change along with utilization-focused evaluation provided invaluable guidance to increase the utility of the research findings.

A theory of change explains how an intervention like performance management can achieve its goals, and guides an evaluation by explicating the areas of inquiry and the manner in which they are most likely to link together (Frechtling, 2015). It is customarily accompanied by a visual representation in the form of a logic model (Rogers, 2014). For this thesis, a theory of change and a logic model provided a framework for a thematic analysis of the literature. The resultant themes were then used to generate scale items, form correlation hypotheses and support the interpretation of empirical data. This research is predominantly exploratory and as is common in exploratory empirical research, initial data analysis requires no a priori hypotheses; however, this is, arguably, an unrealistic requirement. To elaborate, after an extensive literature review, researchers rarely have no expectations of data patterns. Thus, it is more accurate to state that exploratory research demands analysis without adherence to a priori hypotheses, which was the case in this thesis.

Utilization-focused evaluation is an approach based on the premise that valuable evaluations are both useful and used; therefore, it requires careful consideration of the effect of each evaluation step on data use (Christie et al., 2005; Patton, 2000; Ridde et al., 2005). It mandates the identification of the intended uses, the intended users, and the potential obstacles to data utility in order to mitigate them as much as possible. Guided by utilization-focused evaluation, a scale to measure implementers' perceptions of performance management was developed for organizations to use, irrespective of their evaluation capacity. Importantly, the scale was intended to be used as a triage tool. To clarify, because perception is a vastly complex construct, it is only with great hubris can one assert that a scale can measure all aspects of perception. As such, the intended use of the scale is to help organizations identify priority areas for further inquiry and intervention.

Despite the fact that all staff have a stake in their organization's performance management and its evaluation, the thesis presented performance managers as the primary intended users. Because of the potential restrictions on funds, time or evaluation skills within any given organization, the scale needed to be easy to use and interpret. It was also designed to work with existing metrics to give organizations a more comprehensive view of performance management implementation. Though utilization-focused evaluation requires active collaboration with the primary intended users to identify the content, methods, and uses for the evaluation (Patton, 2000), this was outside the scope of this thesis. However, the research findings link well to future research and evaluation projects on performance management, discussed in Chapter 7 Section 4.

4. Significance

The significance of this research is threefold; first, as the literature showed, there is a need for empirical determination of performance management perceptions that might influence the relationship between performance management policies and implementation. Second, a theoretically sound and empirically tested measure for the perceptions can contribute significantly to evaluation for evidence-based performance management. As there is no shortage of performance metrics for organizational use, creating another measure to capture the same information would not be as useful as one that works with existing metrics. Given that the literature contains many HRM metrics like Kirkpatrick's (1959) model, or the balanced scorecard from Kaplan and Norton (1996), this research, by focusing on performance managers' perceptions, provided a missing piece in the proverbial data puzzle.

Many studies have shown that the gap between the academic and the practitioner communities in evidence-based management is sizable due to many reasons. One of which pertains to the practical relevance of academic research in performance management practice (Cohen, 2007; Rynes et al., 2007; Sanders et al., 2008). In essence, practical relevance can be explained in three forms; conceptual, legitimating, and instrumental relevance (Nicolai & Seidl, 2010). Conceptual relevance, most common in the literature, refers to the theories and conceptual findings that can help practitioners understand a particular phenomenon. Legitimating relevance refers to authority that research has in validating a certain practice. Instrumental relevance relates to the ability of research findings to resolve real-life practitioner problems (Wang et al., 2020). Though it can be argued that this thesis addressed conceptual and legitimating relevance by combining theory and empirical data to identify the perception targets, instrumental relevance was most pertinent because the Evaluation for Evidence-Based Management scale (EMI for

short) was developed, specifically, to provide a valid and useful instrument for organizations to use in their evidence-based management efforts.

Third, the exploration of the characteristics of the perceiver and the situation and the extent to which they influence perception added to the significance of the thesis' contribution for two reasons. First, the literature about which characteristics of the perceiver and / or the situation influence perception were mixed so this research aimed to contribute to clarifying the effects of such characteristics. Second, most of the literature focused on perceptions of performance management held by all staff (Andersson Arntén et al., 2016; Baluch, 2017; Bos-Nehles & Veenendaal, 2019; Burke et al., 2015). In addition, multiple studies indicated an assumption that senior managers are aware of HRM implementation to such a degree that the respective research samples recruited senior managers (Bangchokdee & Mia, 2016; Gerhart et al., 2000; Yunos et al., 2015). Other studies focused on non-managerial employees because their experience with HRM was deemed to be a driver of performance outcomes (Wang et al., 2020). While both streams of research are meaningful, their target populations added to the research gap. Therefore this research focused on implementers and provided a deeper insight into the perception of performance management and its antecedents in evaluation for evidence-based management.

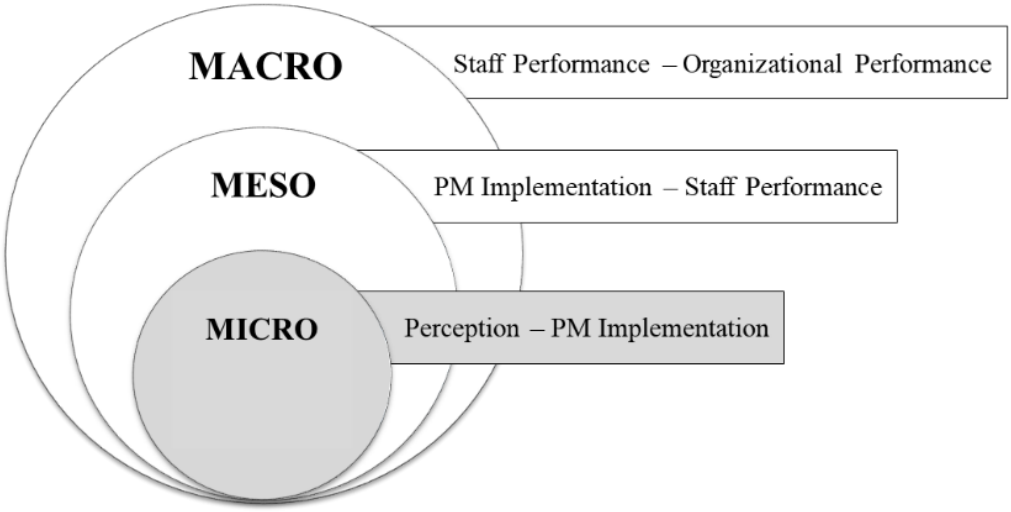
5. Research Scope, Goals and Questions

There are two lenses with which the research scope was established; the first regards three relationships within the organizational performance phenomenon were conceptualized and the focus of this thesis resided in only one of them as the following describes. The first relationship can be seen at a broad or macro level and encompasses the relationship between organizational functions and organizational outcomes. In this relationship, staff performance is thought of as an influential factor between organizational functions and performance outcomes.

The second relationship can be seen at a middle or meso level, which refers to the association between performance management policies and staff performance. In this relationship, performance management implementation is theorized to be an influential factor between performance management policies and staff performance outcomes. The scope of this research was within the third relationship, which is at a small or micro level. This level refers to the connection between performance management policies and performance management implementation. In this theoretical relationship, the three components of perception were considered; the target of perception (performance management), the perceiver (performance managers), and the situation (the organization). Perceptions of performance management are theorized as important antecedents that can influence the relationship between performance management policies and implementation behavior, while characteristics of the organization and the implementers can moderate the extent of perception. As shown in Figure 1, the scope of this thesis is within the micro level.

Figure 1

Conceptual Hierarchical Relationships in Organizational Performance

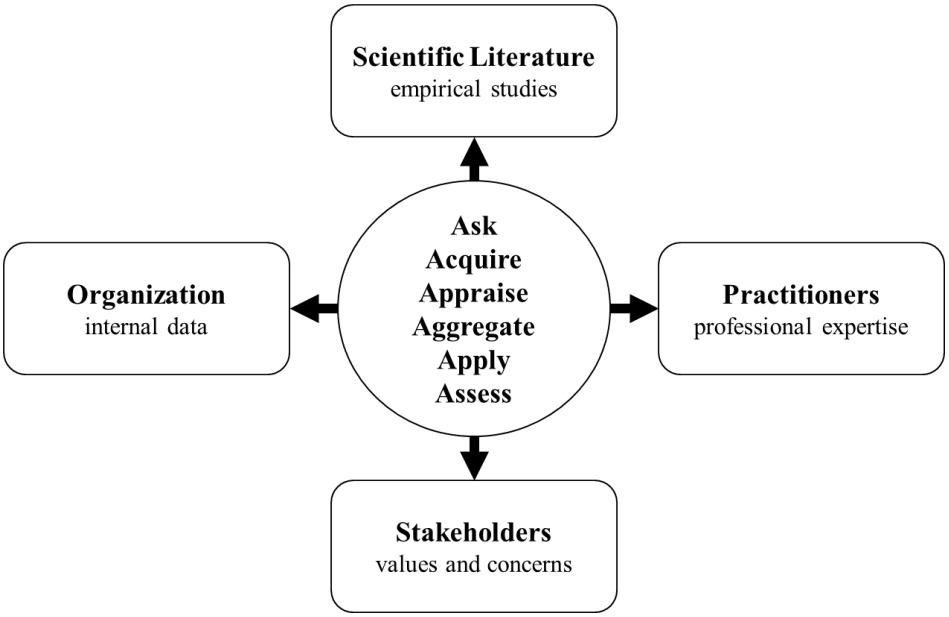


Note. Research scope (shaded) within three conceptual relationships of organizational performance. PM = performance management.

The second lens used to establish the scope relates to evidence-based management. As shown in Figure 2, evidence-based practice has a clearly structured framework. Rynes and Bartunek (2017) synthesized the definition from the Center for Evidence-Based Management (CEBMa) and stated that "Evidence-based practice is about making decisions through the conscientious, explicit and judicious use of the best available evidence from multiple sources by: asking an answerable question; acquiring research evidence; appraising the quality of the evidence; aggregating the evidence; applying the evidence in decision-making, and assessing the outcomes of the previous steps" (p. 239)

Figure 2

Framework of Evidence-Based Management



Note. Figure adapted from CEBMa illustrating the steps and evidence sources required for evidence-based management (Center for Evidence-Based Management, 2019).

It is important to note that the thesis does not present 'low performance' as a given, rather that gaining insights of performance managers' perceptions can be useful for organizations that wish to evaluate their performance management systems. As such, the research, sits within the 'Acquire' step, in relation to evidence from the organization / internal data.

In sum, the definitive goal of this thesis was to bridge the gap between evidence-based performance management and academic research. To do so, it aimed to identify implementers' perceptions of performance management that may influence the relationship between performance management policies and implementation behavior. Second, it aimed to create a

theoretically grounded and empirically tested scale for organizations to use, irrespective of their evaluation capacity, to measure those perceptions.

Therefore, the first research question guiding this thesis was:

RQ 1: What perceptions of performance management should be measured in evaluation efforts for evidence-based performance management?

In addition, two sub-questions were added in reference to characteristics of the perceiver and the situation as follows:

> RQ1a: "what perceiver characteristics influence perception of performance management and to what extent?"
>
> RQ1b: "what situation characteristics influence perception of performance management, and to what extent?"

The second research question guiding this thesis was:

RQ 2: How can these perceptions be measured in a valid and reliable measurement scale?

With respect to the second research question, the validity in question was construct validity and reliability referred to internal consistency. To address the research questions, three sequential research stages were done, namely; item generation, scale development, and scale refinement and validation, which are discussed in detail in Chapter 3.

6. Thesis Organization

This thesis contains seven chapters, including this one, and six appendices. The next chapter provides the literature review, which contextualizes the research and expands on this introductory chapter by detailing organizational performance, HRM, performance management, evaluation, and perception as a potential influencer of implementation behavior. Chapter 3 outlines the research design, specifically the three research stages, and details the empirical studies chosen to address the thesis goals and questions. In particular, it discusses the data collection, sampling, and analytical approaches chosen for each study. Chapters 4, 5 and 6, respectively, present the results of the three sequential empirical studies; in particular, Chapters 4 and 5 provide recommendations for their respective subsequent study, while Chapter 6 presents highlights of the findings for discussion. The thesis concludes with Chapter 7, which offers a summary of the research and a thorough discussion of the results, significant contributions and specific recommendations for future directions.

Chapter 2: Literature Review

This chapter expands on the description of the research context and specific research problem presented in Chapter 1. In particular, it discusses HRM, performance management, organizational metrics and evidence-based management in greater detail. This is followed by an explication of the specific evaluation approaches and theory appropriate for the thesis, and presents perception as an important factor in employee behavior detailing the specific perception targets that should be measured to enhance evidence-based performance management. The chapter ends with a summary and recommendations for the first study and the overall research.

1. Organizational Performance

As stated in Chapter 1, organizations use HRM to attain and retain talented staff to increase their competitive edge in their respect market. The rise of commitment-based HRM and fall of control mechanisms is most prevalent in organizations aiming to retain talented workers in the knowledge economy. While there are myriads of philosophies, theories and approaches for HRM, there is consensus on the necessity of HRM, a core component of which is the improvement and management of staff performance (Armstrong, 2015; Boxall, 2007; Boxall et al., 2007). Despite the diversity in organizational characteristics like funding sources, levels of social responsibility, government regulation, or industry-specific pressures, the end goal is the same; to realize organizational goals in a cost-effective and sustainable manner (Hunter & Nielsen, 2013; Richard et al., 2009). To do so, organizations often rely on; financial resources, like government funding or private revenue; physical resources, like building facilities and transportation; technological resources, like information systems; and human resources. What has changed markedly over the last few decades, in developed countries is the nature of

organizational performance in terms of type of service and the labor needed, organizational operations and performance measurement practices (de Grandbois, 2016; Hunt & Samman, 2019; Thakur, 2011). In essence, the concept of performance became more nuanced and complex due to changes in population demographics, labor migration, technological advancement, and socio-political environments (McGovern et al., 2007). For example, with the digital age came a shift in the economy of developed countries from manufacturing to knowledge services so performance became less about tangible output and more about intangible intellectual contribution. Moreover, social media increased the capacity for global communication, including that between organizations and the workforce (Brown et al., 2017; Parveen et al., 2015). For example, the LinkedIn platform, designed for professional networking as well as staff recruitment, increased the visibility of the workforce and work opportunities on a global scale (LinkedIn Corporation, n.d.).

Changes in international politics have also expanded vocational opportunities and labor migration, thereby increasing the number of competitors for talented knowledge workers (Lischka, 2015). Moreover, the rise of the aging population not only affected the demand side of a market but also the supply side, specifically, in terms of pension costs and business models. To elaborate, as life expectancy increased, many countries have raised the retirement age to cope with the demands on the retirement pension. As a consequence, the workforce started ageing as well, and organizations needed to adapt their business models to the needs of their workers. This was especially compounded by the perpetual evolution of technology as it made training important not only for new staff, but tenured staff as well (Brown et al., 2017).

The aforementioned developments have led to new demands on organizations to increase their market competitiveness, particularly in terms of human capital. Therefore, in addition to traditional performance indicators like financial profit or consumer satisfaction, the measurement of organizational performance now includes the capacity to recruit and retain talented staff (Richard et al., 2009). As a consequence, the role of HRM has become more crucial to an organization's core business than ever (Beatty et al., 2003). The next section describes HRM and its theoretical underpinnings. It follows with a deeper exploration of evidence-based practice and the challenges it faces, thereby highlighting the room for significant contribution from this research.

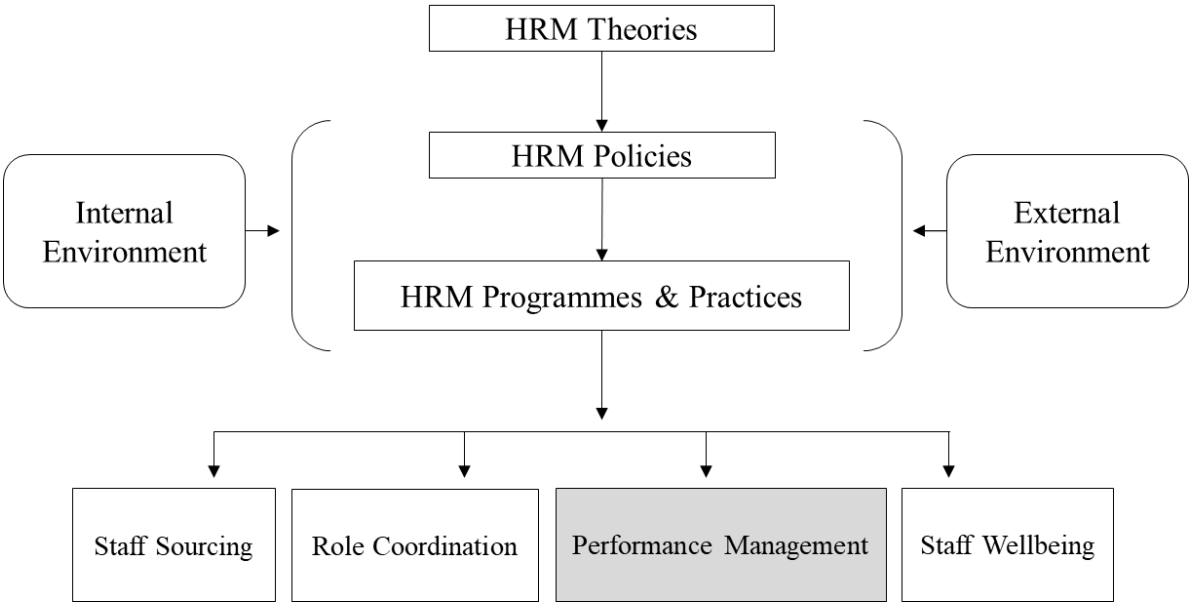
2. Human Resource Management

HRM entails a comprehensive approach to the employment and development of staff with the aim to affect staff behavior in a way that is advantageous to both the organization and the workers (Armstrong, 2015; Armstrong & Baron, 2005). Guest (1991) suggested several objectives for HRM, which included promoting functional flexibility to manage innovation, attaining and retaining skilled employees, and systematically integrating staff performance into organizational strategy. Given the complexity of HRM, the prolific writer and former Chief Examiner of the Chartered Institute of Personnel and Development, Michael Armstrong, proposed that HRM should be viewed as a system in its own right (Armstrong, 2015; Armstrong & Taylor, 2014). Shown in Figure 3, an HRM system consists of interconnected and supportive components that influence the design and achievement of HRM outcomes. As Armstrong (2015) proposed, an HRM system sits within the context of an organization's internal and external environments. The external environment encompasses pressures on an organization that stem from its social, legal, economic, and political surroundings (Armstrong, 2015).

The internal environment presents pressures that stem from the organization's size, type of service, history, and staff characteristics (Armstrong & Taylor, 2014). As shown in Figure 3, to achieve its objectives, HRM requires multiple complex practices, broadly categorized under four components; staff sourcing, role coordination, performance management, and staff relationships and wellbeing (Armstrong & Taylor, 2014; Katou, 2015; Katou et al., 2014). The focus of this thesis was on the performance management component as highlighted in Figure 3.

Figure 3

Components of an HRM System



Note. An HRM system adapted from Armstrong (2015). The focus of this thesis is within the shaded box (performance management).

2.1 *Strategic Alignment*

The recognition of the need for HRM to be a core strategic operation in organizations gave rise to the concept of strategic HRM (Armstrong & Taylor, 2014; Storey, 2001). Strategic HRM is grounded in the specific propositions that; (a) human capital plays a strategic role in the organization's success, (b) employees are key to policy implementation, and, (c) systematic planning, implementation, and evaluation of HR strategies are essential to HRM effectiveness and organizational performance (Armstrong, 2015; Collins & Clark, 2003).

In general, strategic management involves using a process of strategy development and modification over time to achieve specified organizational goals (Boxall & Purcell, 2011). Thompson & Strickland (1999) argued that the main strategic management activity lies in instilling the organization with direction, long term-goals and a roadmap for achieving said goals. Others have built on Thompson and Strickland's (1999) work and set at least four core activities of strategic management; goal setting, strategy formulation, implementation, and evaluation (Armstrong, 2015; Collins & Clark, 2003; Lengnick-Hall et al., 2009). Importantly, practitioners and researchers alike seem to agree that strategic management necessitates the continuous monitoring and improvement of strategies to ensure the organization can adapt to any socio-political, legal, and technological changes (David, 2013; Noe et al., 2015).

There are two potential obstacles facing strategic management; the first is that, in practice, strategy formulation is not as linear nor rational as it is in theory (David, 2013; Thompson & Strickland, 1999). Because decisions are limited by people's problem-solving capacity, their perceptions, emotional reactions, and potential susceptibility to the influence of organizational politics, strategies can be difficult to discern (Harrison, 2009; Lengnick-Hall et

al., 2009). The second is that implementation of strategy is contingent on the actions of many staff members within the different levels of the organization's hierarchy so, unless all staff have a mutual understanding of the strategy and its importance, implementation is likely to be idiosyncratic (Al Shirawi, 2015). Therefore, strategy formulation and implementation for decision-makers and implementers with a common purpose, who share an economic logic and value evidence-based decisions (Al Shirawi, 2015; David, 2013).

2.2 *HRM Strength*

A strong HRM system is more likely to bring about desired employee behavior and organizational performance (Bowen & Ostroff, 2004). Following pivotal work in the 1990s like that from Becker and Gerhart (1996), Huselid (1995), and Ichniowski et al. (1997), the relationship between HRM and organizational performance became a focal point in many research publications. Moreover, mediators and moderators within this relationship form an integral part of HRM academic discourse (Ostroff & Bowen, 2016; Wright & McMahan, 1992; Wright et al., 1994).

According to Bowen and Ostroff (2004), strong HRM systems require communication mechanisms that make HRM content coherent, distinctive, and understandable enough to create consensus among the employees (Ostroff & Bowen, 2016). The seminal work from Bowen and Ostroff (2004) has been used many times over in the research on HRM and performance. However, HRM, in terms of its strength as a system, has been incorrectly conceptualized and operationalized by some (Ostroff & Bowen, 2016). The authors noted that rather than focusing on the higher order-level construct of HRM strength, what some investigations have used was in fact an individual-level construct. Ostroff and Bowen (2016) commended authors like Delmotte et al. (2012) as they were among the few who explicated their perspective by using the term

'perceptions of HRM system strength' (Wang et al., 2020). Therefore, since this thesis related more to perception as an influential factor in performance management, the scope was confined to a specific and individual-level construct (perception of performance management), and the potential perception discrepancy between groups of performance managers/implementers.

2.3 *Theoretical Underpinnings*

The complexity of staff performance has driven substantial academic work,, and development in HRM. By and large, theories regarding staff performance can be categorized by their perspective; employee motivation or organizational motivation. Though, in practice, the theoretical underpinning of an organization's HRM is rarely explicated, by observing an organization's HRM system, one can see the influence of any or a combination of the following theories. This section provides a description of ten HRM theories that were helpful in determining the perception targets used in this thesis.

Employee Motivation Theories. Six HRM theories from the perspective of employee motivation are presented next; expectancy theory, equity theory, commitment theory, the AMO theoretical framework, social exchange theory and, organizational support theory.

Expectancy Theory. Vroom's (1964) expectancy theory relates to the cognitive precursors of employee motivation and how these precursors relate to each other (Lunenburg, 2011; Parijat & Bagga, 2014). The theory suggests that employees expect their hard work to lead to commensurate performance, which, in turn, leads to proportional reward (Armstrong & Taylor, 2014; Frenkel & Bednall, 2016). In effect, motivation to exert effort is directly influenced by the employee's expectation that by working hard, their tasks can be successfully done and that their performance will lead to appropriate reward (Lunenburg, 2011). Parijat and Bagga (2014) suggested that the theory can be a useful predictor of the choices employees will make when

given different tasks, and as such remains a popular theory in many organizations. One of the shortcomings of this theory is that definitions of *effort* or *performance* and perceptions of the *value* of rewards can vary from one person to the next, making them difficult to quantify (Parijat & Bagga, 2014). In addition, it could encourage overuse of extrinsic rewards, which in unionized organizations, are heavily regulated and not necessarily linked to individual performance (Lunenburg, 2011; Parijat & Bagga, 2014). Though a negative effect of extrinsic rewards on motivation remains to be seen, some have posited a concern for organizational capacity to sustain increasing extrinsic rewards (Hegney et al., 2006; Songstad et al., 2012). Therefore, HRM based on this theory should focus on intrinsic rewards, like employee autonomy, in addition to extrinsic rewards based on performance.

Equity Theory. Adams' (1965) equity theory explains how fair treatment in the workplace helps to motivate workers (Koontz et al., 1986). In essence, Adams proposed that employees seek a fair balance between their effort and compensation in comparison to fellow workers. This theory is often in effect in unionized organizations, where equity and equality are highly valued (Carpenter et al., 2012). This theory's strengths lie in its ability to provide an additional layer for consideration when examining employee motivation that expectancy theory lacks, namely, team dynamics.

Given that teamwork is a common and critical feature in most organization, equity theory enhances the body of knowledge of how performance can be influenced by employee perception of equitable treatment in direct comparison to their teammates (Armstrong & Taylor, 2014; Lawler, 1994; Shields, 2007). For example, when employees feel that the organization's decisions are fair, they tend to respond with commitment behavior like helping teammates without external prompts (Colquitt et al., 2001; Ghosh et al., 2014). Similar to expectancy theory,

a limitation of equity theory is the definition and measurement of *effort*. To elaborate, as this theory is based on perceptions of equity, what one individual may see as low level effort may differ to another, thereby creating a gap in reward expectation (Koontz et al., 1986). As such, HRM based on this theory needs explicit, transparent, and consistent performance measures, standards, and rationale for rewards. Again, similar to expectancy theory, compensation can be in the form of intrinsic or extrinsic rewards and as such, companies must be aware of the long-term effects of the type of compensation they provide.

Commitment Theory. Proponents of commitment theory, like Guest (2002) and Walton (1985), suggest that promoting commitment instead of imposing control can benefit an organization in the long-term. This is because a committed employee needs no external impetus to take initiative in their work, and is less likely to request financial compensation for such initiative (Armstrong, 2015). Law (2007) concisely acknowledged that HRM can represent a mechanism of control because of the mandatory nature of its practices and the power it gives to the supervisor over the employee. Neither of which is conducive to meaningful employee commitment. HRM based on this theory tend to invest in more collaborative means of performance goal setting, measurement, and improvement to address the imbalance of power.

The AMO Theoretical Framework. The Ability, Motivation, and Opportunity (AMO) framework explains the link between HRM and staff performance (Appelbaum et al., 2000; Bailey et al., 2001; Boxall & Purcell, 2011; Marin-Garcia & Martinez Tomas, 2016). The framework posits that performance is a product of the joint relationship between a worker's ability, motivation, and opportunity to perform. As Armstrong and Taylor (2014) stated, "the formula provides the basis for developing HR systems that attend to employees' interests,

namely their skill requirements, motivations and the quality of their job." (p.7). The framework notably stresses that to get optimal staff performance, the three components are needed in unison.

To elaborate, a skilled employee can, generally, perform well, but if they are also motivated and given autonomy over their work, they will most likely exceed performance expectations. The framework equally stresses the fact that if even one of the three components is missing, then HRM efforts are diminished (Marin-Garcia & Martinez Tomas, 2016). For example, if a skilled and motivated employee perceives opportunities in their workplace to be based on favoritism, they are far less likely to perform well. The framework provides a much needed structure to understanding the relationship between HRM and staff performance, however, it omits the influence of other contributors to employee behavior like personal beliefs and perceptions (Marin-Garcia & Martinez Tomas, 2016). Therefore, HRM practitioners that use the AMO framework need to understand employee perceptions because they can influence their behavior.

Social Exchange Theory. Social exchange theory, established by Blau (1964), suggests that a worker will tailor their contribution at work in line with what they perceive others have contributed to them. The theory advises that social relationships, including those at the workplace, are fundamentally guided by reciprocity (Armstrong & Taylor, 2014). Put simply, workers who obtain something of value will most likely feel morally obliged to behave in kind (Gould-Williams & Davies, 2005). Some research findings have supported social exchange theory in showing that employees tended to remain committed and productive as long as they perceived their organization to be contributing to their career goals and development (Erdogan et al., 2004; Worley, 2006). Similar to expectancy theory and equity theory, this theory refers to the relationship between the worker and the organization but also with their colleagues. HRM based

on this theory tend to focus on team dynamics as well as building relationships with the organization.

Organizational Support Theory. Organizational support theory, initiated by Eisenberger et al. (1986), explains the development of organizational commitment. Like social exchange theory, it relates to the impact of perceptions of organizational support on staff performance attitudes and subsequent performance behavior (Rhoades & Eisenberger, 2002). Nonetheless, an important dissimilarity between the two theories is that, contrary to social exchange theory, organizational support theory suggests that the obligation to reciprocate is not morally based as much as it is economically driven. That is, when employees have favorable perceptions of their organization's support, they reciprocate in order to capitalize on that support. A good example of this was noted in research by Rahman and Nas (2013). The authors discovered that positive perceptions of organizational investment in staff skill development resulted in lowered rates of turnover intention because the employees valued the development of their skill set and overall employability. HRM based on this theory tends to invest in formal support mechanisms like professional development, as well as internal career development / promotion.

Organizational Motivation Theories. Four HRM theories from the perspective of organizational motivation are presented next; human capital theory, resource-based theory, agency theory and, contingency theory.

Human Capital Theory. Proponents of human capital theory, like Thompson and Harley (2007), stated that the theory relates to the way workers contribute to enhance organizational performance, and how their contribution is valued by the organization. The theory asserts that human capital, like any resource, can be bought (by offering competitive salaries), or made (by training existing employees) (Armstrong & Taylor, 2014). The theory emerged in defense of

valuing human resources over exploiting them. It also became increasingly popular as the economy shifted from primary industries, like mining, to tertiary industries, like knowledge services (McGovern et al., 2007). HRM based on this theory tends to focus on attaining and retaining talented staff as a direct mechanism of increasing the organization's market competitiveness.

Resource-Based Theory. Resource-based theory, also called the resource-based view, was developed by Barney (1991). It merges concepts from organizational economics and strategic management. The theory is similar to human capital theory in the emphasis it places on the value of employees for an organization's competitive power in its respective market. Associated with human capital theory, resource-based theory also states that an organization's competitive power is increased when its resources, including human resources, are unique and costly to duplicate (Armstrong & Taylor, 2014). However, the value of an employee to an organization can be difficult to measure considering that concept of value differs in meaning from one person to the next. HRM based on this theory, like human capital theory, focuses on the value of human resources to increase the organization's market advantage. This is especially the case in knowledge industries where intellectual property is the main commodity, and is thus fostered by the organization and retained as its own property rather than that of the employee (Breuer et al., 2011; Brown et al., 2017; Huselid et al., 2005; Noe et al., 2015).

Agency Theory. Organizational policies are often dictated by the executive team, initiated by senior management, designed by HR professionals, and implemented by middle and junior managers due to their direct working relationships with the rest of the staff (Soldan & Nankervis, 2014). Agency theory, also called principal-agent theory, posits that, in most organizations, a perceptible separation exists between the principals like senior executives, and their agents like

line managers (Al Shirawi, 2015; Armstrong & Taylor, 2014; Piening et al., 2014). This theory highlights the fact that simply because a policy is well developed; it may not be implemented as intended if there is no support from the implementers. Agency theory is controversial; its critics state that it implies that staff cannot be trusted (Armstrong, 1996; Gomez-Mejia & Balkin, 1992). On the other hand, its proponents posited that the theory emphasizes the importance of clear communication and collaboration to gain support from staff across the many levels of organizational hierarchy (Baumeister, 2016; Soltani, 2005; Wright & McMahan, 1992). They further stressed that lack of support from *agents* is not necessarily malicious, rather, a common result of diverse perceptions because people assign their own meaning, for example, to the purpose and feasibility of policy implementation. As such, this theory must be construed with care as it can lead to either controlling management practices or collaborative processes depending on its interpretation.

Contingency Theory. Contingency theory, pioneered by Legge (1978), states that organizational operations are reliant on the organization's environment and circumstances. Essentially, it highlights the need to be aware of and form contingencies for factors that influence causal relationships between policies and outcomes (Armstrong & Taylor, 2014). This theory influenced HRM models to consider the external and internal environments that can affect HRM effectiveness (see Figure 3). HRM based on this theory focuses on performance outcomes with great consideration for the context in which they occurred.

These theories have informed much of the conventional wisdom and assumptions in HRM research and practice (Armstrong & Taylor, 2014). While HRM should be educated by social science theories, it must not be solely based on them. As in other disciplines like medicine, a need for evidence-based practice gained traction over the last few decades. The next section

discusses evidence-based management and the challenges it faces despite the consensus on its importance for effective practice.

2.4 *Evidence-Based Management*

Traditional management practice tends to rely on conventional wisdom based on experience, intuition, anecdotal advice, the internet, and extant organizational processes to make decisions. However, as stated in Chapter 1 (see Section 5), evidence-based management requires a more organized approach to decision making.

Similar to the early proponents of evidence-based medicine, advocates of evidence-based management, like Briner and Barends (2016) and Rousseau (2012), continue to strengthen the importance of its use and its need for systematic use of reliable evidence to inform decisions (Barends et al., 2014; Hodgkinson, 2012; Sackett et al., 1996). However, evidence-based management faces considerable challenges because it lacks the standards and protocols within evidence-based medicine in which a consensus was reached on what constitutes reliable evidence (Rousseau & McCarthy, 2007). Therefore, evidence-based management requires a certain level of mastery developed through years of practice in which an appreciation of scientific principles in decision-making is fostered as well as that for critical thinking, and ethical practice (Rousseau, 2012). As a consequence, evidence-based management is not the norm in HRM practice (Rousseau, 2012).

The earlier movement in evidence-based HRM stressed the use of systematic reviews, as the pinnacle source of evidence (Barends et al., 2012). The issue with this was highlighted by Bartlett (2011) who stated that when systematic reviews challenge the behavior of practitioners, especially those in powerful positions, the results will almost always be rejected due to the lack of knowledge of the value of systematic reviews as well as issues in translating the findings into

practice. It was further stressed that this is because of the psychological tendency to reject or ignore information that is inconsistent with one's worldviews, especially if they had no ownership over or relationship with the research (Munro, 2010). Proponents of participatory evaluation approaches like Patton (2000) or Cousins and Chouinard (2012) stressed time and time again that when clients are involved in evidence generation, they are far more likely to accept and use the data even if it challenges their understanding of the phenomena at hand. This is never the case with systematic reviews or experimental studies because of the bias that accompanies stakeholder involvement. Indeed, there is potential for the quality of data to be compromised by subjectivity the further one moves from systematic reviews or randomized experiments, but the value of data is restricted if they are not used. However, evidence construction should not move away from objective methods to an extreme degree, rather, because most HR decisions are more socio-political than empirical in nature, an *understanding* of the extant political dynamics, subjective perceptions, and emotional components within organizations must be considered in evidence construction (Barends et al., 2017; Briner & Barends, 2016; Hodgkinson, 2012). The need to consider the internal political climate of an organization is not a new suggestion. Baritz (1960) demonstrated that the notion of successfully applying scientific research through an apolitical approach is profoundly flawed. The careful identification of the specific stakeholder with whom to engage as well as the development of mutual trust between the stakeholders and the proponents of evidence-based management goes a long way in successfully implementing evidence-based management (Hodgkinson, 2012). This is strongly supported in the field of evaluation (Cousins & Earl, 1992; Fetterman et al., 2014; Patton, 2000; Ridde et al., 2005). Moreover, evidence-based management requires meaningful engagement of a wider cohort of organizational stakeholders including representation from line

managers, administrative staff, trade unions and organizational staff affected by the phenomena in question (Baughman et al., 2011).

The contemporary view of evidence-based management, as exemplified by Barends et al. (2014) expands it to include robust empirical research, as well as practitioner expertise, contextual information, and stakeholder data, like those from client satisfaction surveys (Hodgkinson, 2012). As such, the new evidence-based management movement emphasizes evidence construction using a combination of scientific inquiry and problem-solving methods with acknowledgement of the highly politicized context in which the evidence applies (Bartlett, 2011; Briner et al., 2009).

Many authors have emphasized two main barriers to evidence-based management; the first is the divide between scholarly research and management practice, and the other is the limited use of evaluation as a source of evidence (Bartunek, 2007; Bartunek & Rynes, 2014; Cascio, 2008; Cohen, 2007). The remainder of this section outlines the two barriers and the factors that contribute to them.

The Academic-Practice Divide. According to research by Dopson et al. (2012) into manager access and use of research in six leading health organizations "research-based knowledge and particularly management journals appear as the lowest source of interest and influence for most managers" (p.7). Six contributors to the academic and practitioner divide were prominent in the literature (Werner, 2012). The first concerns time limits and the general conceptualization of time. Restrictions of time to solve problems in practice make evidence-based management almost impossible for many HR practitioners (Zanardelli, 2012). To elaborate, when decisions need to be made quickly, the search for scientific evidence is unlikely to be the first recourse because managers are seldom trained in the search and critique of

academic evidence, instead, many opt for advice from colleagues, professional publications, or the internet (Zanardelli, 2012).

The conceptualization of time is often different between academics and practitioners, because the former seldom work on immediate problems, like those for which practitioners seek information (Zanardelli, 2012). To clarify, most scholars write for other scholars, with the aim of progressing theory or changing discourse for an eventual impact in practice (Werner, 2012). Practitioners, on the other hand, work within an immediate timeline, and so, they flock to information specifically regarding their day-to-day work (Leung & Bartunek, 2012).

The second contributor to the academic research and practice divide is the language used in academia versus the language within professional HR publications. To elaborate, though academic articles often include a section outlining the implications of the findings, these sections rarely form the bulk of the information provided, in fact, they are usually smaller than an article's abstract (Cohen, 2007). This is most likely because the practical implications of the findings are not usually the reasons for writing the article in the first place and are thus added like an afterthought to provide closure to the research problem at hand (Leung & Bartunek, 2012). In addition, the implications section commonly refers to implications for future research not current practice.

The language in academic research publications also contain a level of abstraction that business periodicals avoid (Werner, 2012). The latter emphasize the practicality of the information they contain, and are written in a language common to most management practitioners (Leung & Bartunek, 2012). However, the vernacular used in academic articles is typically that of at least a graduate school level, and includes jargon, like 'regression analysis' or 'constructivist epistemology', which is alien to the majority of managers (Bartunek, 2007; Leung

& Bartunek, 2012). If one were to view this through the lens of Bloom's Taxonomy, one would note that academia produces information with data collection, analysis, and synthesis as the main foci, whereas trade books and other practitioner publications focus on application (Bloom, 1956). Therefore, HR managers do not typically read academic publications, rather they opt for trade books and periodicals like the Harvard Business Review, because they are far more understandable and relatable (Kovner, 2012; Leung & Bartunek, 2012). Though in publications like the Harvard Business Review, the aim is to provide easily consumed information to help improve practitioners' practice, more often than not, the writers' treat data rigor and analysis as an afterthought, if they are mentioned at all (Cohen, 2007; Werner, 2012). Arguably, there is little incentive for academics to produce publications tailored for practical use as they are not usually rewarded for non-academic writing, which in turn, is a common job requirement for role advancement in academia (Leung & Bartunek, 2012; Werner, 2012). Many higher education institutions push their academic staff to not only publish for academic platforms but also for journals with the highest Q ranking (Quality factor) which are seldom open-access (Bennis & O'Toole, 2005). As several authors noted, the format of scholarly research publications can be chaotic for non-scholars, and the relationship of its often dense language to real-life managerial issues can be tenuous at best (Cascio, 2008; Markides, 2007; Vermeulen, 2007). Ostensibly, the format of such publications seems to lack a foundation of practical utility by non-scholars.

The third contributor to the academic research and practitioner is the education and certification requirements for managers. Like medical or law schools, business schools are classified as professional schools so education in management does not emphasize training in scholarly literature. However, unlike business students, medical or law students read academic publications and may even write their own (Barends et al., 2012; Leung & Bartunek, 2012).

Moreover, management roles do not legally require certification as is the case for doctors or lawyers; many managers become managers based on their skills specific to the organization's service and then learn management skills through experience or professional development (Leung & Bartunek, 2012).

The fourth contributor to the academic and practitioner divide is the poor access to research evidence. The storage of academic data is often complex making it difficult to search and access relevant publications (Werner, 2012). Unlike physical libraries, cyberspace does not provide a librarian to help with the search, leaving the practitioner to navigate the many, and often unnecessarily complicated online databases on their own (Werner, 2012). One means of accessing online research is through subscription-based databases like Business Source Complete, Academic OneFile and JSTOR with which journal publishers like Taylor and Francis, have strict contracts on cost and scope of access (Werner, 2012). Searching through the petabytes[3] of text is a skill in its own right requiring patience and stamina. Unlike the business discipline, some disciplines like medicine have their own research consolidated in one database; PubMed. So, users of such databases have to learn how to work in one system, whereas users of business databases have to learn multiple systems, which is a harrowing task for most non-scholars (Werner, 2012). Capitalizing on this issue, Google offered Google Scholar[4], but for the most part it only offers access to citations that link the user to content that must be purchased.

A second point of access is through databases that belong to professional organizations like the Society for Human Resource Management, or academic ones like the Academy of

[3] One petabyte is 1000 terabytes.
[4] Google Scholar was launched in 2004 as a freely accessible database.

Human Resource Development (Academy of Human Resource Development; Society for Human Resource Management, n.d.). A third way to access relevant information is through business-oriented consulting organizations, which offer access to their subscribed members (Society for Human Resource Management, n.d.; Werner, 2012). This can, however, be limited in its propriety because information from business firms is tailored to promote their own products and services rather than the advancement of evidence-based management knowledge (Lawler, 2007).

Despite the move of book publishers to surrender their work to peer review in university presses like SAGE or Routledge, not all do so (Rousseau, 2012). Some opt for the commercial route to reach wider audiences and publish in popular platforms like HarperCollins[5] (Wilbourne, 2001). This makes it hard for the average reader to distinguish between robust research publications from commercial ones. By the same token, university presses, due to financial pressures, are extremely selective about their publications, frequently opting for those that gain more market value, as such well researched manuscripts seldom get published (Wilbourne, 2001).

Academic databases are akin to VIP lounges, offering access to full-text publications subject to pay-per-view rights, or full subscription at exorbitant costs so, unless an organization has subscriptions to such databases, an individual manager faces a significant barrier to academic research (Werner, 2012). Alvarez and Mazza (1998) colorfully compared the gap between academic and practitioner literature to that between high fashion and off-the-rack clothing, the former being elite, and the latter more accessible to the common man (Terpstra & Rozell, 1997).

[5] HarperCollins Publishers L.L.C. is one of the Big Five English-language publishers along with Simon & Schuster, Hachette, Penguin Random House, and Macmillan.

Fortunately, the movement towards open access scholarly publications has enhanced practitioner access to academic research, but it needs to be paired with active encouragement from academic institutions for scholars to publish in open access databases (Odlyzko, 2002).

Internet developments over the last decade have also increased practitioner access to information. Examples include professional social networking sites like LinkedIn (LinkedIn Corporation, n.d.), and open online collaborative projects platforms like WikiProject Business (Wikimedia Foundation Inc., n.d.). In the latter, volunteer editors evaluate and improve the thousands of business related articles in Wikipedia. Other sources of information on the internet include websites from governmental bodies, expert associations, and online communities of practice. The Australian FairWork Ombudsman website has a page dedicated to performance management (Fair Work Ombudsman, n.d.). It is a reliable source of general and legal information about the topic, with specific relevance to the Australian context and it provides templates for practical use. It, however, focuses heavily on under-performance and does not give guidance on evaluation for evidence-based management. Similarly, other governmental bodies like the Australian Human Rights Commission or the New South Wales Government – Industrial relations site provide guidance but with a strong focus on managing under-performance rather than evidence-based management (Australian Human Rights Commission, n.d.; NSW Industrial Relations, n.d.).

The AHRI, a membership-based organization, provides some information publicly, but restricts most of its information to its members (Australian Human Resources Institute, 2019). It provides members with professional development, access to its own research findings, full access to the monthly HRM magazine (Australian Human Resources Institute, n.d.-b), certification opportunities, and AHRI: ASSIST, which is an online resource center, providing a range of

practical tools (Australian Human Resources Institute, n.d.-a). The World federation of People Management Associations is another membership-based organization representing approximately 700,000 HR professionals from over 95 organizations around the world. Its site provides a limited amount of open access resources (World Federation of People Management Associations, n.d.). Despite the membership costs, more practitioners join these associations because, in addition to published resources, they offer professional development and networking opportunities.

The issue of access is linked to visibility, the fifth contributor to the research and practice divide (Lawler, 2007). The almost invisible research domain is competing with juggernauts in the business literature market. The distribution of business books is a significant industry; Audible Australia[6] has over 22,800 audiobooks under its business section, BookDepository[7] has more than 1,052,700 books under the business & management category, Booktopia[8] has over 152,000 books, electronic books and audio-books for management. These platforms are far more visible and accessible to practitioners, though they do have costs; their respective online interface is far more user-friendly than most academic databases. Unfortunately, publications in commercial platforms are not rated by their data quality or research rigor, but by user ratings of popularity, which led to publications like *The Five Dysfunctions of a Team* by the business author and screenwriter Lencioni (2002), which is literally a fictional tale, to be among the top

[6] Data relevant as of December 2019
[7] Data relevant as of December 2019
[8] Data relevant as of December 2019

publications in management, commonly categorized under the 'non-fiction' heading in online bookstores[9].

The sixth contributor to the academic and practitioner divide concerns the development of evidence. Evidence in social science is socially construed and guided by individual interests and understanding (Hodgkinson, 2012). The perceptions of value may be different between academics and management practitioners in that the latter place high stock in experiential knowledge more so than empirical research (Dopson et al., 2012). Multiple studies discovered that HR managers have inconsistent beliefs about scholarly HR topics and that they tend to rely on what they think would work best (Bezzina et al., 2017; Rynes et al., 2002; Rynes et al., 2007).

A study of 1,500 practitioners in the United States, Belgium, and the Netherlands, examined the underlying reasons for such practitioner behavior (Barends et al., 2017; Bezzina et al., 2017). The study supported the position that the desire to generate academic publications might be an issue rather than a virtue for organizations (Bartunek, 2007). Many organizations prefer to engage the services of management consultancy firms rather than academic researchers due to their preference of strict confidentiality and the perception that confidentiality may be compromised by the academic's desire to publish their findings (Rynes et al., 2007). Therefore, this means that organizations may be collectively limiting their knowledge base in two ways: (1) the publicly available literature becomes limited in practical knowledge and, (2) because consultancy firms have financial and reputational stakes in the results, their research may not be

[9] In Booktopia, this book is featured among the Business Books for Excellence, and ranked number 6 in Amazon UK for Human Resource Management.

as rigorous or unbiased as academic research, which is subject to ethics regulations (Bartunek & Rynes, 2014; Cohen, 2007).

Some organizations, fortunately, are making headway in developing a community of practice between academic scholar and practitioners, for instance by employing staff with scholarly degrees to work with practitioners (Cohen, 2007). An example is the Center for Effective Organizations at the University of Southern California, which has researchers who conduct academic research as well as collaborate with practitioners to improve organizational management (Leung & Bartunek, 2012). Such collaboration not only offers credible data, but skill development for practitioners.

Evaluation Use. The second barrier to evidence-based management relates to the limited evaluation for evidence-based management, which has three key contributing factors (Derven, 2012; Griffin, 2012; Marshall & Rossett, 2014). The first of which is the inconsistency in understanding evaluation. The term evaluation has been used speciously within much of the business and management literature as the content is more accurately related to measurement than evaluation (Armstrong, 2015).

In addition, the evaluation literature is filled with different terms, approaches and constructs, most of which overlap to such a degree that the process of authoritative categorization is almost impossible (Alkin, 2013; DuBow & Litzler, 2019). For instance, the array of the terms that have been used to describe evaluation led by theory include chains of reasoning, logic analysis, impact-pathway analysis, program theory analysis, program theory-driven evaluation, and theory-based evaluation (Donaldson & Lipsey, 2006; Weiss, 1997a; Wildschut, 2014). Having so many terms that essentially refer to one branch of evaluation approaches can be confusing to most readers. Moreover, the terms ‘approach’ and ‘model’ have

also been used interchangeably in the literature (Bennett, 2003; Stufflebeam et al., 2002). This thesis used the term approach in reference to a broad concept driving the attitude and perspective guiding an evaluation, and the term model in reference to a narrower construct illustrating the ideal set of processes within an evaluation approach.

Many publications on evaluation discuss it in terms of program evaluation because evaluation of programs was the genesis of the profession (Scriven, 1974; Shadish & Luellen, 2011). HRM and performance management, for the purposes of this thesis, are categorized as an intervention. Though quite similar to a program, the main point of distinction is that programs have an end date, whereas interventions may be mainstreamed and become a sustainable operation. Last, it is important to acknowledge that almost all of the evaluation literature discusses evaluation in terms of external evaluation, that is an independent evaluator is hired to conduct the evaluation, rather than an internal evaluation conducted by the program's personnel (Skinner, 2001).

As stated in Chapter 1 Section 1, specious perceptions of evaluation present a significant barrier to its use in evidence-based management (Lobo et al., 2010; van Koperen et al., 2016). Negative experiences with evaluation reduce the likelihood of meaningful engagement with the evaluation process and acceptance of its data (Cousins et al., 2013; Mathison, 2008). Such negative experiences include instances where audit and other accountability practices were used under the guise of evaluation. This goes against what proponents of evaluation state as the purpose of evaluation, which is to improve rather than prove performance (Mathison, 2005; Stufflebeam et al., 2002).

The second contributor to poor evaluation use is its expense. Evaluation can be a costly endeavor because causal links between HRM and performance outcomes can be hard to determine, let alone improve, as they are not as straightforward as they are in other disciplines like medicine (de Brito & de Oliveira, 2016). To elaborate; HRM requires a multi-faceted management approach to improve employee performance and consequently organizational performance. That, combined with external variables that effect employee performance such as illness or family matters, makes it nearly impossible to ascertain the direct cause of a particular behavior change (Bowen & Ostroff, 2004; Gibb, 2000). In addition to the complexity of HRM, a dominance of isolated rather than strategic HRM practices contributes to the difficulty in ascertaining causa links. Such isolation of practices can lead to duplication of processes, jeopardize efficiency, and make measurement that more difficult (de Brito & de Oliveira, 2016).

The third contributor to poor evaluation use in organizations is the limited capacity to do and use evaluation. As Bourgeois and Cousins (2013) noted, conducting an evaluation requires specific skills in evaluation approaches, research methods and communication skills. As such, organizations that value evaluation tend to either use external evaluators, while others have internal evaluators or evaluation units (Bourgeois et al., 2015; Carman, 2007; McKegg et al., 2016). In addition, Bourgeois and Cousins (2013) also noted that organizations need the capacity to use evaluation data to inform decisions and lead to meaningful improvement. Increasing an organization's capacity to do and use evaluation requires significant support from senior leadership, therefore the extent of evaluation in organizations can be highly dependent on perception of its value and organizational politics (Chelimsky, 1995; Weiss, 1993).

The barriers to using academic research and evaluation data have led to partial evidence-based practice. In particular, information from the internet, professional expertise, and

organizational metrics constitute the bulk of the evidence used to inform HRM decisions. Organizational metrics like end of year financial reports, provide much of the data used to inform HRM decisions on resource allocation and performance management however, they are prone to measurement error (Gerhart et al., 2000; Huselid & Becker, 2000). The next section discusses staff performance management paying particular attention to the requirements of effective practice. This is followed by a description of the prevalent organizational metrics used to inform organizational decisions related to performance management.

3. Performance Management

For the purposes of this thesis, performance management refers to an HRM initiative with the specific intention to improve and maintain the performance of existing staff (Armstrong, 2015; Armstrong & Taylor, 2014). It is "a systematic process for improving organizational performance by developing the performance of individuals and teams. It is a means of getting better results by understanding and managing performance within agreed framework of planned goals, standards, and competency requirement" (Armstrong & Taylor, 2014, p. 9). Historically, the precursor to performance management is the annual performance review which has been a common element in the workforce, at least, since 1914, when Lord and Taylor Co.[10] instituted a formal performance review system (Law, 2007). It was not until the 1960's and 1970s that management-by-objectives came along with assessment techniques like behaviorally anchored rating scales (Milkovich & Wigdor, 1991). The term 'performance management' did not gain recognition until the 1980's (Armstrong, 2015). Like other proponents of performance management, the AHRI stated that performance management is a more expansive construct than

[10] Lord & Taylor is a luxury department store in the United States founded in 1826

performance review (Australian Human Resources Institute, 2019). In fact, performance management encompasses performance measurement and gives it meaning and context (Lebas, 1995). Unlike the traditional performance review, performance management is not an annual ritual, rather it is a continuous process for improving staff performance that requires goal setting, strategic alignment, planning, monitoring, feedback and support (Armstrong & Baron, 1998; Australian Human Resources Institute, 2019).

With the rise of strategic HRM came the push for strategic performance management which require the explicit alignment of employee performance with organizational performance (Lengnick-Hall et al., 2009). In 2009, a survey conducted by CPA Australia and the University of Technology – Sydney found that organizations with high performance management systems combined a broader range of measures, to evaluate staff performance than other organizations (Bedford et al., 2009). The research findings stressed the importance of emulating such practices because using only outcome measures of performance gives organizations insight into one part of the situation and, usually, in an untimely way for effective intervention (Bedford et al., 2009).

For the purposes of this thesis, effectiveness describes the degree to which something can achieve the desired outcomes. Effective performance management can communicate to employees the strategic goals of the company, specify the implications on individual performance expectations, and align individual career goals to the organization's goals (Armstrong, 2015; Shields, 2007).

3.1 *Requirements of Effective Staff Performance Management*

The effectiveness of performance management has been a topic of professional and scholarly research for decades however, the utility of said research is highly debatable. As Sharma et al. (2016) noted, the literature regarding performance management research is diluted

by weak theoretical foundations of measure development, measurement inconsistencies and flawed interchangeable use of certain terms (Furnham, 2000). These issues have drastically limited the utility of performance management research (Cawley et al., 1998; Seiden & Sowa, 2011). The muddy distinction between related constructs has, arguably, started the proverbial domino effect of the aforementioned issues in performance management research (Keeping & Levy, 2000). Specific to this thesis, the extant literature on perceptions of performance management was limited because it did not provide a unified understanding of how said perception should be conceptualized, operationalized and measured (Biron et al., 2011).

Conceptual issues are exemplified by research like that of Dewettinck and van Dijk (2013), in which worker perceptions of performance review fairness was used as a surrogate measure for perceptions of the fairness of performance management. This is highly problematic because, though the evolution of performance appraisals to performance management systems has been gradual in both academia and practice (Claus & Briscoe, 2009), it has been clearly established that the former is only one component of the latter (Biron et al., 2011; Sharma et al., 2016). Likewise, Haines and St-Onge (2012) used a scale for performance management outcomes as a barometer for the effectiveness of performance management. While both pieces of research used important measures, conclusions about performance management based on only one of its multiple facets are misleading. To clarify, in this thesis, the terms *appraisal* and *review* refer to a measurement process, conducted at least once a year, of employee performance outcomes (DeNisi, 2000; DeNisi et al., 1984) while the term *implementation* refers to the adoption and activation of organizational policies by the intended implementers (Soldan & Nankervis, 2014). There are eight requirements of effective performance management; role

clarity, accurate measurement, involvement, feasibility, strategic utility, improvement, defensibility, and evaluation (Strebler et al., 2001).

Role Clarity. The first requirement is role clarity which states that that in order to perform one's role well, one needs to be clear about their responsibilities and role boundaries. For decades research findings showed that role ambiguity has detrimental effects on staff behavior (Jackson & Schüler, 1985; Tubre & Collins, 2000). This is because chronic confusion about one's tasks leads to anxiety and conflict between employees (Jackson & Schüler, 1985; Schaubroeck et al., 1993; Tubre & Collins, 2000). Recent research by Hassan (2013) showed that clarifying employee responsibilities and performance requirements enhanced overall role clarity which, in turn, improved performance and reduced turnover rates. Recently, the Society for Human Resource Management (2017) stated that "role clarity and understanding who is making what decisions can be the sources of either great corporate governance or unending organizational dysfunction" (p. 32).

Accurate Measurement. The second requirement of effective performance management states that there need to be measurable criteria for performance to reduce bias (Strebler et al., 2001). This requirement has become especially difficult but quite critical due to the complexity of performance in the information economy. No longer can performance be measured by the number of tangible outputs produced, rather it requires a multifaceted analysis of the intellectual and intangible contribution of staff to their team, department, and overall organization (Carlson & Kavanagh, 2018; Marler & Boudreau, 2017). While there remain some tangible outputs in knowledge services, like the number of reports completed for an organization's clients, the other aspects of performance include effective communication, stakeholder engagement, or client retention all of which require relationship building behavior which is hard to measure.

Researchers have found that high performing organizations use performance measures that require deeper analysis of performance than traditional measurement effort (Fink & Sturman, 2017; Handa & Garima, 2014). These include the 360 degree feedback process in which a staff member's performance is based on feedback from their colleagues, supervisors, subordinates, and their own self-reflections (Atwater & Brett, 2006; Solansky, 2010).

Involvement. The third requirement of effective performance management states that employees must be involved in the strategic design and implementation of their performance management (Ceschi et al., 2017; Lee, 2005; Strebler et al., 2001). In essence, it was stipulated that when workers are involved in, at least, the formation or review of their performance goals, they are more likely to become engaged in their job as they have an increased sense of control and ownership over their performance (Armstrong, 2015; Dayaram & Fung, 2014; Katsaros et al., 2014; Seiden & Sowa, 2011). This requirement was supported by findings from a survey of over 100 large US organizations (Lawler et al., 2012). Importantly, the findings showed that jointly developed strategic goals are more effective than those developed solely by senior staff. Likewise the findings showed that performance management will fail if there is no practical support for its implementers (Marin-Garcia & Martinez Tomas, 2016; Op de Beeck et al., 2016; Wright & Nishii, 2013a). It was also noted that the infrequency of performance management evaluation was deleterious to its effectiveness. Further, the findings showed that more than half of the organizations surveyed had a one-size-fits-all performance management system which was deemed as an ineffective approach to performance management (Lawler et al., 2012). This requirement was further supported by findings from the 2014 e-reward survey of performance management (e-reward, 2014). It showed the fundamental role of staff involvement, particularly of implementers, in all aspects of performance management. In line with equity theory, the

survey respondents cited fairness and transparency of performance management as essential to its effectiveness (e-reward, 2014; Flint, 1999; Heslin & VandeWalle, 2011).

Feasibility. The fourth requirement of effective performance management states that processes need to be practical and have sufficient resourcing if they are to be implemented well (Birken et al., 2015; Joint Committee on Standards for Educational Evaluation, 2008; Songstad et al., 2012). To elaborate, in order to implement performance management, as intended, the implementer needs the relevant skills and sufficient time and space. The 2014 e-reward survey supported this requirement by illustrating substantial feasibility issues in performance management (e-reward, 2014). For instance, many respondents stated that they lacked the time to properly develop, change, and/or implement performance management. Moreover, the complexity of performance management processes was a recurring theme among survey responses, stressing that in order for performance management policies to be appropriately implemented, they need to be as simple as possible to enact (e-reward, 2014; Farndale & Kelliher, 2013; Op de Beeck et al., 2016). In a similar vein, survey respondents stressed the importance of simplicity for performance goals to eliminate any blurring of the line between good and poor performance (e-reward, 2014). Too much bureaucracy and paperwork were cited in other research as contributors to negative perceptions of performance management, which lead to poor or limited implementation (Kim et al., 2015; Op de Beeck et al., 2016).

Strategic Utility. The fifth requirement of effective performance management states that performance management, like all HRM interventions, must be clearly useful to the organization's strategic directions if it is to be incorporated as a core business function, rather than an add-on or separate operation (Collins & Clark, 2003; Strebler, 2001; Strebler et al., 2001; Wright & Nishii, 2013b). This was supported by the findings of a 2010 survey of 750

WorldatWork members, mainly senior managers in HR, about the effectiveness of performance management (WorldatWork & Sibson Consulting, 2010). More than half of the survey participants stated that one of the contributors to performance management ineffectiveness was that, in their organizations, performance management was viewed as an HR activity more so than a vital organizational function. When it came to the strategic alignment of performance goals, the survey results showed that it was customarily done at the senior management level, and decreased in prominence the further down the organization's hierarchy one looked (WorldatWork & Sibson Consulting, 2010).

Improvement. The sixth requirement of effective performance management states that its primary goal must be improvement rather than measurement of performance (Armstrong, 2015; Seiden & Sowa, 2011). This requirement was also supported by the findings of the 2010 WorldatWork survey (WorldatWork & Sibson Consulting, 2010). Two thirds of survey participants noted a reason for ineffectiveness in performance management was that that the main purpose of performance management in their organizations was to justify the distribution of rewards.

Defensibility. The seventh requirement of effective performance management states that decisions must be justifiable, credible, and transparent to the plethora of stakeholders (Strebler, 2001; Strebler et al., 2001). This requirement highlights the influence of perception on behavior for example, the acceptance of performance feedback is more likely to occur if staff perceive it be fair and credible. As the Society for Human Resource Management (2017) recently asserted, "feedback is more acceptable when recipients believe the feedback is fair and accurate" (p. 21). Moreover, recent research from Järlström et al. (2018) found four essential aspects of sustainable HRM that are akin to defensibility; justice, transparency, profitability, and staff well-being. Their

work illustrates that decisions need to be justifiable on multiple organizational levels to be effective in achieving employee and organizational outcomes.

Evaluation. The eighth requirement of effective performance management states that it must be regularly evaluated to ensure its continued utility and effectiveness. HR scholars and professionals alike seem to agree that evaluating performance management is critical for its effectiveness, much like the importance of performance measurement in performance management. This requirement was supported by findings of the e-reward survey, in which a prominent theme was that the regular evaluation of performance management was an important precursor to improving performance management (e-reward, 2014).

Most managers often rely on data generated within their organization which produces contextually relevant information, however, it can be prone to measurement error (Donaldson, 2012). Such errors can be reduced by using multiple indicators for the same phenomenon, like using financial data, staff performance data and customer satisfaction data to measure organizational performance (Baysal et al., 2013; Fink & Sturman, 2017). Nonetheless, it has been established that data on management outcomes are not as readily available as other organizational data, and that there is a tendency for organizations to focus on what is easily measured which can be counterproductive (Fink & Sturman, 2017; Kerr, 1975). The following provides a brief discussion of the organizational metrics related to HRM and performance management.

3.2 *Performance Metrics*

Recently, literature emphasized the difference between basic HRM data collection like payroll costs, with HRM metrics that aim to capture the effectiveness of HRM initiatives like performance management (Fink & Sturman, 2017). As Cascio and Boudreau (2011) postulated, analyses that simply report numbers, like cost-per-hire, merely quantify the workforce and add little value in informing conclusions about how well performance management is working as an organizational function. While basic data reporting offers a quantifiable account of the organization's operations, HRM metrics intend to offer evaluative insight on effectiveness (Cascio & Boudreau, 2011; Fink & Sturman, 2017).

There are many types of HRM and business metrics used in organizations with varying degrees of sophistication (Carlson & Kavanagh, 2018). Largely , HRM metrics can be categorized as those for efficiency, effectiveness and impact (Boudreau & Ramstad, 2007). Efficiency metrics focus on the relationships between investments and outcomes in terms of speed and relevant ratios. Such metrics include cost-per-hire and time-to-fill (a vacant position). HRM metrics are also fiscally oriented to a substantial degree, for example, measures of cost-per hire, cost of paid leave or over-time, and revenue per FTE[11]. These metrics are not without their limitations, for instance, revenue per FTE links the time and use of human capital to the organization's revenue output, and if the revenue per FTE ratio increases it indicates that more revenue was generated by said human capital (Carlson & Kavanagh, 2018). However, the ratio can increase due to a decline in FTE or outsourcing, which then provides misleading calculations (Fink & Sturman, 2017). Their pitfall, however, occurs when used in isolation, for example,

[11] Full Time Equivalent

time-to-fill can be indicative of HR staff's speed in filling a vacant role, but neglects to account for the quality of the skills needed for the role (Boudreau & Ramstad, 2002; Fink & Sturman, 2017). Effectiveness metrics tend to assess whether HRM initiatives are delivering the intended outcomes, and to what extent they are doing so. This requires data collection at various operational points and combining them to portray a comprehensive picture of effect. Importantly, it may require the collection of data beyond what is commonly measured in organizations, so an issue is presented in the design or adoption of validated and useful measurement tools (Donaldson, 2012). Impact metrics aim to ascertain the impact of HRM investments on the overall performance of the organization (Carlson & Kavanagh, 2018). For example, return analytics are used to measure the financial investments and the resulting profit or benefit (Society for Human Resource Management, 2016). Importantly, impact metrics for performance management may show a small effect size given the numerous confounding factors that influence organizational performance like the social environment, the introduction of new legislation, and the actions of competing organizations. Though this thesis focused on effectiveness, it recognized that when used alone, any of the three types of metrics become highly susceptible to their limitations, but together they are far more likely to give an accurate representation of HRM performance (Bedford et al., 2009; Fink & Sturman, 2017; Hodgkinson, 2012).

There are other non-fiscal metrics, which enhance understanding of an organization's performance and reputation. For example, staffing metrics like Time-to-Start or a position's Vacancy/Occupancy focus on the time it takes to fill a position, which is indicative of the organization's capacity to attract staff (Society for Human Resource Management, 2016). To mitigate the potential errors from using metrics in isolation, a few measurement frameworks

emerged in the management discipline, the three most pertinent to performance management are outlined next.

The Balanced Scorecard. The balanced scorecard, authored by Kaplan and Norton (1996), posits that strategy development and evaluation should be undertaken using several perspectives; internal business, finance, customer, and innovation and learning. This approach remains popular in HRM because of its strategic nature focusing on actions and deliverables from human capital, which in turn usually accounts for the bulk of an organization's costs (Beatty et al., 2003; McLean, 2005; Niven, 2014). Broadly the balanced scorecard is used in two ways; overall organizational performance with clear strategic alignment, and specific employee performance requiring a balanced scorecard for each employee (Niven, 2014).

Prior to its introduction, many organizations faced a common issue, that being the isolation of the measures they used which lead to an inability to link data in a meaningful pattern (McLean, 2005; Niven, 2014). So, rather than collecting data spanning all organizational operations, Kaplan and Norton developed the balanced scorecard to focus on the aforementioned four-perspectives (Kaplan & Norton, 1996; Niven, 2014). The balanced scorecard provided clear and simple language for organizational performance. Essentially, it requires the selection of accurate measures, implementation by trained employees, and robust data mining (Nankervis, 2006; Niven, 2014).

A well designed balanced scorecard for performance management includes a mix of efficiency, effectiveness and impact metrics (Beatty et al., 2003; Denton & White, 2000; Fink & Sturman, 2017). Given that HRM matters are often included in a balanced scorecard, its users would at least have some HR metrics, even if they were perfunctory (Fink & Sturman, 2017). Nonetheless, it is not without its limitations; first, it faces difficulty in quantitative measurement

of staff contribution, thus often only simple to measure indicators tend to be used (McLean, 2005). Second, there is great difficulty in establishing robust causal links between actions and deliverables.

Kirkpatrick's Model. Kirkpatrick's (1959) model is, perhaps, the most widely used measurement model in HRM, commonly for staff training and development initiatives. The model proposes four levels to be measured (Kirkpatrick, 1959; Kirkpatrick & Kirkpatrick, 2006). First, the reactions of the participants of a training session, for example what they felt worked and what needs improvement (Kirkpatrick & Kirkpatrick, 2006). Second, the impact of the training on their learning, commonly measured with a written test or a demonstration of the new skills. Third, the participants' subsequent behavior usually measured through observations from supervisors or self-reports (Kirkpatrick & Kirkpatrick, 2006). Last, the impact of the training on the company, which is customarily measured through assessing organizational finances against outputs. There are many reasons for using Kirkpatrick's model since the 1960s, most common of which is the elegance of its simplicity, which is appealing because the model is easy to understand and apply (McLean, 2005). Further, because of its decades' long and extensive use, it has provided organizations with a common language (Ferris et al., 2008; McLean, 2005).

Kirkpatrick's model has its limitations; first, the lack of validation from robust research on the causal links between the levels (Alliger & Janak, 1989; McLean, 2005). Second, the model assumes that staff training and organizational performance have a linear progression. As such, it should be classified as taxonomy instead of model. Last, it was designed for assessing the impact of training and staff development, so it may not be adaptable to other HRM initiatives (Kirkpatrick & Kirkpatrick, 2006).

Return on Investment Model. The return on investment model requires the conversion of all organizational inputs and outputs into financial terms in order to ascertain HR's contribution (Rauf, 2016). The model accepts Kirkpatrick's model but with a strong emphasis on the fourth level, organizational performance. However, because the causal links between the levels are not linear in a real world setting, organizations rely on educated estimates from managers, which unfortunately can be too subjective (Marshall, 2016; McLean, 2005). The appeal of this model stems from its emphasis on quantifiable data, and it is indeed a good model to use for assessing the contributions of, for example, technological resources (Armstrong & Taylor, 2014; Rauf, 2016). To evaluate performance management, however, this model along with the aforementioned metrics lack the contextual analysis and politically sensitive communication needed when examining a complex phenomenon like staff performance.

Evaluation approaches have evolved over many decades to apply robust research methods in real-life contexts. The next section discusses evaluation in general, as well as the specific approaches that guided this thesis.

4. Evaluation

Using evidence in medical practice or in any of the sciences driven by objective data, as a mold for evidence-based performance management, is problematic (Rousseau, 2012). In performance management, as in most social phenomena, the evidence-based process rarely lends itself to quantitative hypothesis testing or clinical trials. Instead, the focus should be on the use of data, from multiple reliable sources, to better enhance the understanding of causal patterns. Berwick (2008) argued that management practices are often messy as they involve a variety of leadership styles, emotional connections, and dynamic environments. Hodgkinson (2012) further asserted that evidence-based management researchers must not be blind to the variety of

contextual realities in which performance management sits. For instance, as it is not essential to have a management qualification in order to become a manager, especially in junior management, many people who implement performance management may not have the skills to do so unless the organization provides them professional development (Ferris et al., 2008; Kovner, 2012). Therefore, research into the implementation of performance management must consider the implementers' situation and characteristics. Moreover, as discussed in Section 2.3, management research is commonly written by scholars with world-views and language dissimilar to those of practitioners, which does little to advance the development of shared understanding or a community of practice. Therefore, traditional academic forms of evidence construction are limited in their contribution to evidence-based management in practice.

By contrast, holistic approaches of inquiry as well as careful communication of results have been cornerstones of the evaluation profession since its infancy, and jointly present a point of distinction between evaluation, measurement and research. For instance, as evaluation scholars have stated, the main difference between evaluation and measurement is that the purpose of evaluation is to not only verify outcomes, but to also improve them (Scriven, 1974; Stufflebeam et al., 2002). As such, evaluation and measurement are not synonymous terms, rather, evaluation is a more extensive form of inquiry that encompasses measurement.

The distinction between evaluation and research is not as simple as that between evaluation and measurement because despite the points of difference, evaluation and research have some noteworthy points of both overlap (Mathison, 2008). A key similarity is that both forms of inquiry require and use an assortment of rigorous methods for data collection and analysis (Bennett, 2003). A point of distinction is in the purpose of inquiry; the chief purpose of evaluation is to provide meaningful and timely information for stakeholders in real-life practice,

whereas the purpose of research is typically the advancement of knowledge or theory to a scholarly audience (Azzam & Levine, 2015; Chouinard, 2013). As such, evaluators conduct research methods with great care for potential power dynamics and socio-political influences within the evaluation's context (Carman, 2007; Curtis et al., 2005; Levin-Rozalis, 2003). In essence, the distinction between scholarly research and evaluation lies in the manner of communicating evidence and level of stakeholder engagement.

Scholarly work shows a value for the inescapable legacy of academia as an elite source of data. This attitude is evidenced by the esoteric language common in academic publications as well as the restricted access to academic databases (Cohen, 2007). Moreover, in the pursuit of objective truth, stakeholders are typically only engaged as far as research subjects. On the other hand, evaluation values deeper levels of stakeholder engagement, using a shared language to communicate evidence, and careful consideration of contextual or political influences on the use of evidence (Levin-Rozalis, 2003; Mathison, 2008). This is not to suggest that one form of inquiry is superior to the other, rather that evidence-based performance management could benefit from both.

Historically, the most basic distinction in evaluation was defined by the terms *formative* and *summative* as coined by Michael Scriven (1974). Formative evaluations aim to explore the development and implementation of an evaluand (subject of evaluation) in order to improve it, while summative evaluations offer an analysis of the outcomes achieved (Scriven, 1974). As such, formative evaluation activities occur on a more frequent basis than summative evaluation, and though it is more resource intensive, formative evaluation provides superior guidance for improvement in a timely fashion (Gullickson & Al-Nawab, 2017). Nowadays, evaluations tend to incorporate both formative and summative evaluation activities, though some tend to place

stronger emphasis on one over the other, contingent on the goal of the evaluation and the available capacity for evaluation (Gullickson & Al-Nawab, 2017). The majority of the metrics previously discussed (see Section 3.2) have a substantial summative orientation, with the exception of Kirkpatrick's model.

The earlier views of evaluation by authors like Stufflebeam and Alkin regarded it as an activity done to provide information to aid in decision-making (Alkin, 2013; Stufflebeam et al., 2002). Much later, in the 1980s, Patton argued that evaluation is any activity that aims to increase effectiveness through systematic data-based inquiry (Patton, 2000; Ridde et al., 2005; Skinner, 2001). The multiple approaches to evaluation can be distinguished, at least to some degree, by their specific purpose. Though all of them aim to help the clients make informed decisions that will hopefully lead to positive change, some have been more specific about additional purposes such as stakeholder empowerment, or increasing the utilization of the evaluation's findings (Cousins & Whitmore, 1998; Fetterman et al., 2014).

Much effort has been exerted to definitively categorize evaluation approaches, and whilst there may be some significant distinctions between them, the majority tend to overlap to such a degree, making such efforts intimidating (Bennett, 2003). Therefore, by and large, publications of evaluation approaches are often categorized by their main proponents because they reflect the beliefs of their main protagonists in terms of the evaluation's purpose, setting, audience, stakeholder involvement, and the evaluator's role (Alkin, 2013; Skinner, 2001). Evaluation approaches were born out of the need to evaluate educational, health, and community programs, as reflected by the abundance of literature on the subject (Mathison, 2005; Mertens & Wilson, 2012; Stufflebeam et al., 2002).

Ralph Tyler's objectives-oriented evaluation represented a mostly experimental and objective paradigm, which was dominant in evaluation for almost half a century (Christie et al., 2005), much like the metrics discussed in Section 3.2. The objectives-based approach is particularly appropriate in evaluating firmly focused projects that have clear and viable objectives (Stufflebeam et al., 2002). Such approaches advocated for true objectivity through stripping the program's contextual factors and focusing on quantifiable indicators of success. Tyler's approach, along with others like it, started to receive much criticism in the 1960s and 1970s. The main debate argued that, while such approaches strengthened the external validity of evaluation findings, they were limited in their portrayal of the variables that exist in real-world settings (Christie et al., 2005).

With time, new theories emerged to address the application of research methods in settings which have socio-political influences (Shadish & Luellen, 2011). With those theories came a plethora of evaluation approaches that aim to account for the shifts in social issues and to address the best ways for knowledge construction, communication, utility and the competencies needed to for quality evaluation (Abma, 2005; Stake, 2011). For one, Michael Scriven's goal-free evaluation, developed in the early 1970s, expanded the scope of objectives-oriented evaluation (Scriven, 1974). The main premise of which is that though ideally all programs' goals should specifically address the need at hand, not all of them do, and so an evaluation should go beyond assessment based on a program's goals and instead measure by the extent to which it meets the actual needs (Bennett, 2003). Another approach that emerged, was program theory-based evaluation, or theory-based evaluation as it is sometimes called, pioneered by Schuman, Glasser and Strauss, and Weiss (Rogers & Weiss, 2007; Weiss, 1997b). This approach explicates the need to understand and use the social science theories that apply within the intervention

being evaluated. Participatory evaluation approaches gained momentum in evaluation efforts for building the capacity of stakeholders to do and use evaluation (Cousins & Chouinard, 2012; Cousins & Earl, 1992). They are, perhaps, more suitable for evaluation efforts in evidence-based performance management. The remainder of this section is dedicated to the evaluation approach, theory and framework that guided the identification of nine perception targets related to performance management, based on which scale items were created in Study 1 (see Chapter 4).

4.1 *Participatory Evaluation*

Participatory evaluation encompasses any evaluation approach that actively involves stakeholders in the planning and implementation of evaluation (Cousins & Whitmore, 1998; Cousins et al., 2013). Participatory evaluators jointly drive the evaluation with the key stakeholders like, in the case of this research, performance managers, and the other personnel targeted by a performance management system. For such evaluations ‘‘control begins with the evaluator but is divested to program community members over time and with experience’’ (Cousins et al., 2013, p. 14). Participants are encouraged to get involved in designing the evaluation, creating instruments, collecting and analyzing data, and disseminating the findings (Shulha, 2010, November). For the purposes of this thesis, a *stakeholder* is any member of staff affected by the organization’s performance management system, while the term *primary intended user* refers to is a staff member who is intended to use the evaluation results in their performance management practice.

The rationale for stakeholder participation stems from the desire to effect positive change, and build the stakeholders’ evaluation capacity to carry on or conduct future evaluations (Christie et al., 2005). An underlying principle of participatory evaluation is that active involvement of stakeholders in the evaluation can augment the likelihood of them using the evaluation findings

appropriately and positively improving their work (Reineke, 1991). Active participation of stakeholders requires them to be engaged players in the evaluation rather than subjects of research. Participatory evaluation is distinguished from other evaluation approaches in four ways (Christie et al., 2005; Cousins & Chouinard, 2012): (1) direct and meaningful involvement of stakeholder throughout the evaluation, (2) active fostering of stakeholder ownership of the evaluation, (3) the evaluator's role is less authoritative, and (4) active development of stakeholder capacity to do and use evaluation.

The nature of stakeholder involvement in participatory evaluation ranges from shared evaluator-participant responsibility to participants' complete control of the evaluation (King et al., 2007). At the shared end of this continuum, the evaluator is heavily involved and assumes responsibility for the quality of the evaluation, while at the other end, the evaluator is minimally involved, acts as a coach, and provides technical assistance (King et al., 2007).

Participatory evaluation can be categorized by three points of divergence according to the seminal work of Cousins and Whitmore (1998) who suggested three unique features of participatory evaluation in terms of stakeholder involvement. The first being control of the evaluation process, ranging from complete researcher control of the technical (process) evaluation decisions to that exerted entirely by the stakeholders (King et al., 2007). Second is stakeholder selection for participation, which ranges from only involving primary intended users to including all legitimate stakeholder groups. Third is the depth of participation, which ranges from stakeholders providing feedback to deep involvement in all facets of an evaluation (Cousins & Whitmore, 1998; Cousins et al., 2013).

There are merits and drawbacks to either end of the stakeholder involvement spectrum. For one, an advantage of limiting stakeholder participation can save time and expense, however, this is of little benefit if there is limited consensus about the evaluand's intentions or processes between all legitimate stakeholder groups. This, in turn, can make the evaluation seem biased and irrelevant to key stakeholders (Cousins & Whitmore, 1998).

There are two distinct participatory approaches; practical participatory evaluation, and transformative participatory evaluation. They are divergent by central aims, purposes, and ideological roots (King et al., 2007). Transformative participatory evaluation explicitly targets marginalized stakeholders and gives them ownership of the evaluation thus empowering them to become the leaders of evaluation to support social change (Mertens, 2013; Mertens & Wilson, 2012). Born out of community development work, transformative participatory evaluation works under the ideology that information is power. Similar to Fetterman's empowerment evaluation, transformative participatory evaluators empower stakeholders by involving them in the process of knowledge construction (Fetterman et al., 2014).

Practical participatory evaluation evolved from a pragmatic need to increase the use of evaluation results by involving key stakeholders. Its concerns relate to improving evaluation practice by connecting with the primary intended users and providing results that they can use in precise ways to effect improvement (Cousins & Chouinard, 2012; Cousins & Whitmore, 1998). Practical participatory evaluators take a leadership role in the evaluation and are instrumental in ensuring the quality of findings (King et al., 2007; Neuman et al., 2013). In practical participatory evaluation the technical decision making is shared between the evaluator and the stakeholders with an emphasis on equal partnership, however, the diversity of stakeholders is usually limited (Hansen et al., 2013). An approach akin to practical participatory evaluation is

Patton's utilization-focused evaluation, as they both aim to support decision making and enhance the utilization of an evaluation (Neuman et al., 2013; Patton, 2000). Patton's approach was instrumental in guiding this research, especially the development of the EMI scale.

4.2 *Utilization-Focused Evaluation*

Patton pioneered utilization-focused evaluation, which also gained support from other evaluation experts and the Joint Committee (Joint Committee on Standards for Educational Evaluation, 2008; Patton, 2000; Ridde et al., 2005). Its goal is centered on the intended use by the primary intended users (Ridde et al., 2005). Every step of a utilization-focused evaluation is focused on how people in their real-life settings would use the evaluation results (Ridde et al., 2005). In other words, utilization-focused evaluation is based on the premise that the success of an evaluation is judged by the extent of its use and the difference it makes in improving the outcomes rather than by its technical excellence. Central to utilization-focused evaluation is the distinction between data use and data dissemination because although dissemination facilitates use, it does not always lead to use. Patton contended that even the most technically sound evaluations are useless if they end with an evaluation report sitting on a shelf (Alkin, 2013; Stufflebeam et al., 2002). Therefore, utilization-focused evaluators must critically understand the barriers and enablers of evaluation use in the context within which they are working to enhance the likelihood of meaningful utility of the evaluation (Ridde et al., 2005; Stufflebeam, 2001).

Baughman et al. (2011) presented the same critical point in stating that evidence has an observable but at times limited place in behavior change. The authors further stated that evidence needs to be presented convincingly in a coherent argument for the behavior change at hand to occur, especially if it challenges long-standing practices. The authors stated that one must "put evidence in its proper place, a place where it will do the most good: as a means and not an end"

(Baughman et al., 2011, p. 64). Similarly, Patton's main contention is that ultimately, the true value of an evaluation is in its practical utility (Patton, 2000). Though utilization-focused evaluation tends to be a more costly activity because, for example, of the need to train organizational staff in evaluation methods, the benefits of such professional development as explained by organizational support theory, include increased organizational commitment from the staff.

The appeal of utilization-focused evaluation stems from its pragmatic and universal nature, as it encourages the use any or as many legitimate evaluation tools or methods in the interest of ensuring the appropriate use of reliable and accurate evaluation data (Patton, 2000). The practical orientation of utilization-focused evaluation leads it to work best with a carefully selected group of representatives rather than with all stakeholder groups (Christie et al., 2005; Stufflebeam et al., 2002). This, however, presents the main limitation of utilization-focused evaluation; turnover. To elaborate, when a primary intended user with whom the evaluator had been working leaves the organization, the evaluator must spend time to engage a new stakeholder. To mitigate this limitation, utilization-focused evaluators tend to work with several primary intended users and establish a succession plan (Patton, 2000; Ridde et al., 2005).

Another limitation, like in many participatory approaches, stems from the possibility of corruption by stakeholders if they push the evaluation direction towards gaining favorable outcomes of their program (Alkin, 2013; Cousins & Chouinard, 2012; Stufflebeam et al., 2002). Therefore, utilization-focused evaluation requires a poised evaluator who can approach the situation with flexibility whilst meeting the professional evaluation standards of propriety, feasibility, accuracy, and utility, discussed in Section 4.4 (Joint Committee on Standards for Educational Evaluation, 2008).

It must be noted at this point that, though the focus is on evaluation use, utilization-focused evaluation recognizes the need to meet all the relevant evaluation standards, not just the utility standard. A strength of utilization-focused evaluation is that it is a situational and dynamic approach using methodology that evolves in response to the situational dynamics and the ongoing discussions between the evaluator and primary intended users (Braverman, 2008; Patton, 2000). As a consequence, the evaluator may have any of the following roles; trainer, technical expert, colleague, spokesperson, or negotiator (King et al., 2001; Stevahn et al., 2005; Stufflebeam et al., 2002). Another strength is in its openness to draw on any of the evaluation approaches or methodologies to maximize the evaluation impacts (Alkin, 2013; Stufflebeam et al., 2002).

Similar to Patton's suggestions in utilization-focused evaluation, Heath and Heath (2007) suggested that for data to have the most impact they must be best suited for the intended user (Patton, 2000). Bartunek (2007) expanded on this and postulated that performance management practitioners are more likely to respond positively to information that is (a) from a credible source, (b) specifically relevant to their work, (c) feasible to apply, (d) easily shared, and (e) written in a more narrative style than that of scientific journals (Rousseau & Boudreau, 2011). Collectively these form characteristics of *sticky data* in that they easily stick in a person's mind (Leung & Bartunek, 2012). As such, utilization-focused evaluation was considered ideal in guiding this thesis, along with a theory of change.

4.3 *Theory of Change*

Evaluators who use a theory of change suggest that all social interventions are based on social theory whether that theory was explicated or not (Rogers, 2014). A theory of change explains how the target of an intervention will, theoretically, change for the better. In essence, a theory of change is used to clearly outline how the intervention investments, activities and outputs will achieve the intended outcomes (DuBow & Litzler, 2019; Reisman & Gienapp, 2004). As such, a theory of change is constructed for each intervention. A theory of change must not be confused with an evaluation theory. Leeuw and Donaldson (2015) presented two typologies of theory in evaluation. The first refers to the theory of change which is contingent on the specific intervention under evaluation. The second refers to the theory that guides the evaluator's actions and approach to conducting an evaluation. For example, utilization-focused evaluators use *evaluation use theory* to guide their evaluation approach, whereas objectives-oriented evaluators use *evaluation methodology theory* to guide their evaluations (Alkin, 2013).

Weiss warned against some challenges of using a theory of change; first, some evaluations use implementation theory, which only guides activities and outputs, in lieu of a theory of change, which guides an intervention in its entirety (Rogers & Weiss, 2007). Second, because interventions rarely have an explicated theory for the evaluator to use, the evaluator must spend time and effort to either identify or construct the theory of change, which can increase costs (Weiss, 1997a). Third, the quality of an explicated theory of change can be questionable if it is based on stakeholder assumptions instead of valid social science research (Weiss, 1997b). Most initiative developers base their work on their own experiential knowledge or others' expertise, as discussed in Section 2.3 on evidence-based performance management.

Though such expertise is valuable, using it as the sole basis of a theory of change is imprudent (Rogers & Weiss, 2007).

Theory construction typically requires using information from intervention documents, staff expertise, validated research, and logical reasoning (DuBow & Litzler, 2019). As such, it is a resource intensive venture. However, the costs associated with evaluation using a theory of change can be a valuable investment. For one, the results help the stakeholders create better initiatives as they know not only what outcomes to expect, but how to get them (DuBow & Litzler, 2019; Weiss, 1997a). Furthermore, a theory of change helps to refine the aspects of the program that should be evaluated, as well as allows the evaluator to present the correct interim indicators of success to stakeholders during the evaluation (Weiss, 1997). Additionally, the emphasis on learning from extant research to inform a theory of change can be invaluable in guiding evaluation measurement and thus time is not wasted on constructing measurement methods from scratch (Donaldson & Lipsey, 2006). Typically, a theory of change is accompanied by a logic model as a illustration of the intervention and its theory of change.

Logic Models. Graphical representations of an intervention's theory of change, called logic models, include the intervention's aims, resources, activities, outputs, outcomes and impact (Rogers & Weiss, 2007). They vary in complexity and sophistication depending on the complexity of the intervention and / or the skill of the model developer (Rogers & Weiss, 2007). To be sure, the development of a logic model is a skill in its own right and requires critical evaluative thinking. The more sophisticated theory-based logic models clearly explicate the anticipated (causal) links between inputs and intended outcomes and the theories or assumptions that led to the chosen anticipated links (DuBow & Litzler, 2019). These theories are based on 'if-then' statements. For instance, if staff performance is declining, then an intervention is needed;

or if an organization needs to communicate to all staff, then communication resources are needed, and so on.

Evaluators often co-construct a logic model with key stakeholders and during this process a theory of change emerges (Ratcliffe, 2012; Wildschut, 2014). Frequently, the process is iterative making the logic model a living document that is adjusted throughout the evaluation (Gullickson & Al-Nawab, 2017). For HRM initiatives, including performance management, Zanardelli (2012) drew special attention to the need for logic models to add transparency to the critical information and assumptions on which managers base their decisions (W.K. Kellogg Foundation, 2004). It must be noted that, logic model are not infallible because, as is the case with many models that show linear relationships, they may not reflect the often unpredictable real-life situations in which an intervention is situated (Kovner, 2012). However, logic models try to mitigate this by explicating the external factors that can affect an intervention.

Creating a theory of change and its logic model, involves several steps, the first of which is an analysis of the situation at hand, that is, the need for an intervention (DuBow & Litzler, 2019). Second, a clarification of the role the intervention will play in addressing said need. This step results in explicit goals, objectives, outcomes, and impacts. The next step involves making educated assumptions on how the desired outcomes and impacts will occur and how the intervention will trigger the necessary events to achieve them (W.K. Kellogg Foundation, 2004). Working backwards from the desired impacts, also called long-term outcomes, both short term and mid-term outcomes are specified (Gullickson & Al-Nawab, 2017). This is followed by identifying measurable outputs that are presumed to be indicative of progress (Renger et al., 2019; The Evaluation Center, 2014). Then, the activities needed to achieve the outputs are specified followed by the inputs needed to carry out the activities. Finally, as previously stated,

significant assumptions and theoretical underpinnings of change are explicated followed by an acknowledgement of contextual factors that could affect the anticipated change (DuBow & Litzler, 2019).

As evaluation theories and approaches evolved, so too did the concept of quality in evaluation. Organizations like the American Evaluation Association, and the Joint Committee on Standards for Educational Evaluation, herein called the Joint Committee, developed evaluation principles and standards for conducting quality evaluation (American Evaluation Association, n.d.; Joint Committee on Standards for Educational Evaluation, 2008). Of most relevance to this thesis are the Joint Committee's standards for personnel evaluation as they specifically relate to staff performance and have remarkable similarities with the requirements of effective performance management developed by Strebler et al. (2001) (see Section 3.1).

4.4 *The Personnel Evaluation Standards*

The evaluation standards were the product of a lengthy, rigorous and interactive process involving many highly esteemed evaluators as well as evaluation stakeholders in an effort to improve personnel evaluation systems (Joint Committee on Standards for Educational Evaluation, 2008; Linn et al., 1989; Stufflebeam et al., 2002). The appeal of the standards stems from their logical and intuitive orientation, the extensive work from the Joint Committee as well as the authority it has, rather than irrefutable empirical research (Ahmady et al., 2009; Glasman & Martens, 1993). The utility of the standards was paramount for the Joint Committee so, the standards were written to prescribe sound practice rather than describe all practice (Stufflebeam, 1994).

The standards were intended for the evaluatees (staff members whose performance is managed), the personnel evaluators (staff who conduct performance management activities), and the users (others with a legitimate need for access to results). The Joint Committee (2008) recognized that because the standards were developed in the United States, they may not be as applicable internationally. It has also pointed out that there are overlaps between the standards due to the complexity of staff performance evaluation.

According to the Joint Committee (2008), a standard is a principle "mutually agreed to by people engaged in the professional practice that, if met, will enhance the quality and fairness of that professional practice" (p. 26). Though the Joint Committee did not explicate the theoretical basis for the standards, it provided a list of guiding assumptions (Linn et al., 1989). First, quality evaluation helps to provide effective services. Second, those evaluated are more likely to be receptive to the evaluation if it is constructive, clear, and free of disparaging features. Third, when evaluation activities consider the evaluatee's context (culture, religion, gender, disability etc.) they will more accurately capture performance as well as differences in performance expectations. Fourth, professional development and learning based on evaluation results will better serve the learners. Fifth, if warranted disagreements on performance arise, using them as an opportunity to better define performance expectations will improve future performance. Sixth, not all performance evaluation systems are equal in their complexity so, some of the standards will be more applicable than others. These assumptions are akin to improvement-oriented evaluation and utilization-focused evaluation in that they are based on the theory that only when evaluation is useful and focused on improvement can it be deemed to be of sufficient quality.

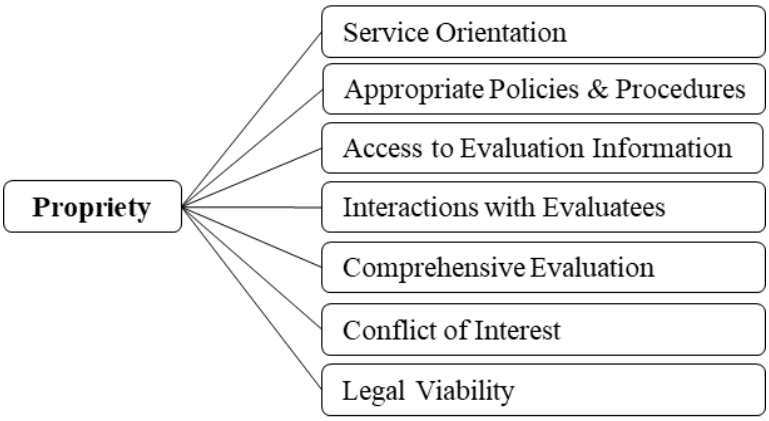
Four overarching standards, which expand to 27 targeted standards, were designed to guide the development and assessment of personnel evaluation systems in a variety of education settings (Joint Committee on Standards for Educational Evaluation, 2008). The following provides a detailed description of the standards.

Propriety. The standard of propriety states that personnel evaluations must be done in accordance with the relevant laws and ethical requirements, and with consideration for the well-being of all the evaluation's stakeholders (Joint Committee on Standards for Educational Evaluation, 2008). This standard covers seven targeted standards, shown in Figure 4. The first, called service orientation, states that staff evaluations must reflect the organization's goals and lead to the desired organizational outcomes. The second, called appropriate policies and procedures guidelines, stresses the importance of protocols for documentation and accessibility of personnel policies.

The third, called access to evaluation information, states that access to staff performance information must be confidential and limited to only those with a legitimate reason. The fourth, called interactions with evaluatees, states that those who evaluate staff performance must do so in a manner that is considerate, respectful and professional to each staff member. The fifth, called comprehensive evaluation, states that staff performance evaluation must concern both performance strengths and weaknesses. The sixth, called conflict of interest, states that any conflicts of interest must be mitigated openly and honestly. The seventh, called legal viability stresses the importance of adhering to applicable legislation. The propriety standards share the same ideology as the improvement requirement from Strebler et al. (2001); specifically the focus on evaluating strengths and weaknesses to improve performance.

Figure 4

The Propriety Standards

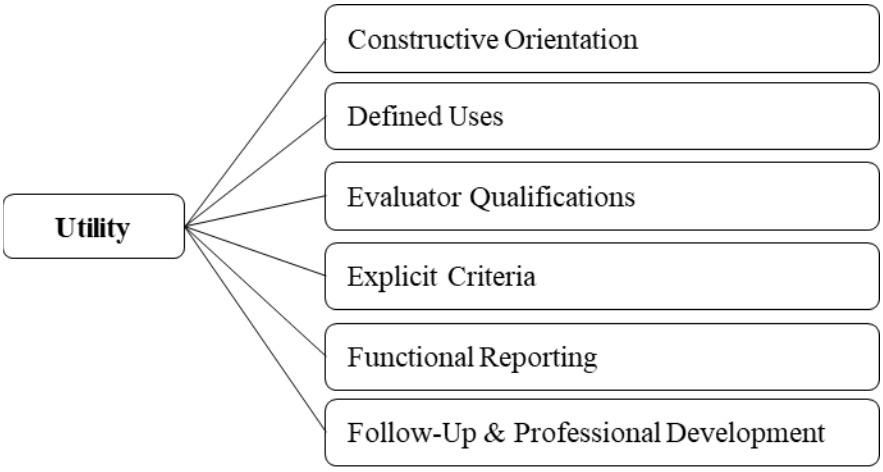


Utility. The utility standard states that performance evaluations must be informative, opportune, and influential (Joint Committee on Standards for Educational Evaluation, 2008). This standard covers six targeted standards, shown in Figure 5, the first of which is called constructive orientation and states that performance evaluations must help organizations to develop good performance and improve poor performance. The second, called defined uses, states that the intended users and uses of performance evaluations must be identified in policies and guidelines before conducting the evaluation. The third, called evaluator qualifications, stresses the importance of training those who evaluate staff performance. The fourth, called explicit criteria, states that performance evaluations must be based on clear and specific criteria directly related to the task requirements. The fifth, called functional reporting, states that evaluations must accurately reflect performance and be documented clearly. The sixth, called follow-up and professional development, states that staff evaluations must lead to appropriate professional development tailored for individual learning needs.

The utility standards share the same ideology as the strategic utility requirement from Strebler et al. (2001); specifically the strategic alignment of performance evaluation with the employee's and the organization's performance goals.

Figure 5

The Utility Standards

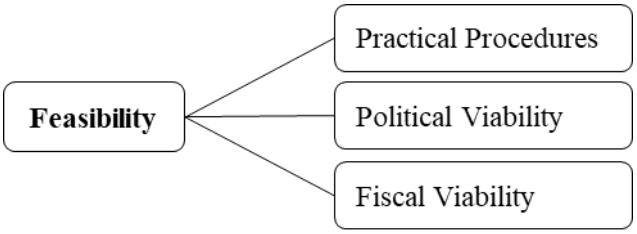


Feasibility. The feasibility standard states that in order to function well, personnel evaluation needs to be easy to implement, efficient and viable (Joint Committee on Standards for Educational Evaluation, 2008). This standard covers three targeted standards, shown in Figure 6. The first, called practical procedures, states that personnel evaluation must be practical and conducted efficiently. The second, called political viability stresses the importance of trust, inclination to engage, and organizational support required to conduct evaluations. The third, called fiscal viability, states that personnel evaluations need sufficient resources in terms of skills, time and finances.

The feasibility standards share the same ideology as the feasibility and the involvement requirements from Strebler et al. (2001); specifically that for performance interventions to be effective, they must be relevant to the stakeholders, practical and well resourced.

Figure 6

The Feasibility Standards



Accuracy. The accuracy standard states that an evaluation must be technically adequate and comprehensive to produce correct and suitable data to make sound judgments (Joint Committee on Standards for Educational Evaluation, 2008). This standard covers eleven targeted standards, shown in Figure 7. The first, called valid judgments, similar to the fifth utility standard (functional reporting), states that personnel evaluations must lead to valid conclusions about performance to minimize risk of misinterpretation. The second, called defined expectations, states that staff evaluations must be based on clear and specific role requirements.

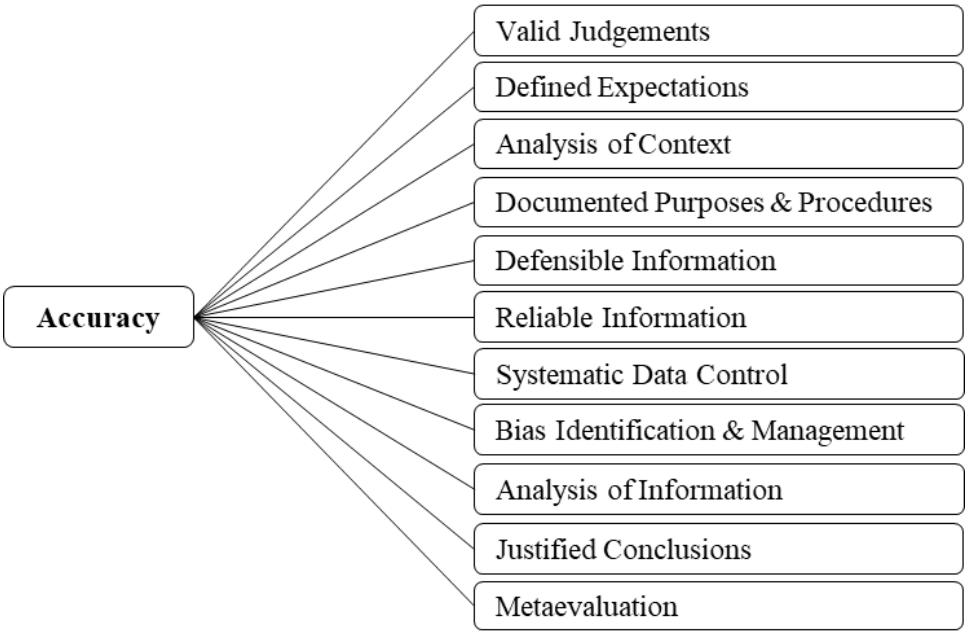
The third, called analysis of context, states that performance evaluations must take the context of performance into consideration such as illness, bereavement and other external factors that may have affected performance. The fourth, called documented purposes and procedures, also similar to the fifth targeted utility standard, states that evaluations must have clearly documented policies to achieve consistency of purpose and process regardless of the evaluator. The fifth, called defensible information, states that all evaluation results must be justifiable.

The sixth, called reliable information, states that all evaluation results must be reliable and consistent. The seventh, called systematic data control, similar to the third targeted propriety standard (access to evaluation information), emphasizes the importance of the systematic review and security of evaluation data. The eighth, called bias identification and management, states that evaluation processes must be safeguarded against bias.

The ninth, called analysis of information, states that performance conclusions must be based on accurate analysis of performance results so they can be credible. The tenth, called justified conclusions, states that performance conclusions must be justified to the legitimate stakeholders. The eleventh, called metaevaluation, states that evaluation processes must be systematically assessed in a timely manner to ensure their quality and continued utility. The accuracy standards share the same ideology as the requirements from Strebler et al. (2001) for role clarity, accurate measurement, defensibility and evaluation. In sum, the ideology is that for performance interventions to be effective, they must be based on accurate information, transparently linked to justifiable actions, without the influence of bias, and subjected to regular evaluation.

Figure 7

The Accuracy Standards



Validation of the Standards. The Joint Committee commissioned a separate validation panel to examine the validity of the standards. Specifically, the validation panel's role was to: (1) identify the assumptions underlying the standards; (2) critique and report on the validation process; (3) assess the applicability of the standards in a variety of national and international contexts; (4) confront the committee with issues and ideas from germane theoretical analyses and empirical research, and (5) publicly report on the validation of the standards (Linn et al., 1989). In 1986, a complete draft of the standards was reviewed by the 54 members of the National Review Panel and another 24 members of the International Review Panel. Combined, they provided a plethora of constructive feedback and suggested revisions (Linn et al., 1989). The standards were also subject to further review and a public hearing. Though comprehensive, the

validation of the standards was in relation to their face and content validity, and despite the mention of field tests, the nature of those tests is unknown (Howard & Harman, 2007).

Strengths and Weaknesses of the Standards. There are many similarities between the targeted standards, some of which are justified though they may seem redundant at first glance. For example, the fifth and sixth targeted accuracy standards (defensible information and reliable information, respectively) which collectively state that evaluation results must be justifiable, reliable and consistent, were split into two standards to highlight the necessity for the evaluation results to, not only be justifiable but, open to scrutiny. Another example is seen when comparing the third targeted feasibility standard (fiscal viability) and the third targeted utility standard (evaluator qualifications) in that they both stress the importance of proper investment in staff evaluation in terms of resourcing. Nevertheless, they differ in their scope; the former covers all resourcing, and the latter focuses on the skills of the evaluator.

Some of the other similarities, however, are so significant that they led to excessive overlap. For example, the sixth targeted propriety standard (conflict of interest) is similar to three of the targeted accuracy (valid judgements, bias identification and management, and justified conclusions) in that they all equally stress the need for evaluations to be fair and unbiased. The resultant overlap was considered, by the Joint Committee (2008), to be necessary to achieve prescriptive clarity. Moreover, the Joint Committee acknowledged such overlap and stated that not all the standards may need to be applied at the same point in time. To elaborate, some organizations may need to improve their performance management in terms of accuracy only; therefore, they need only use the accuracy standards for guidance. Such selective use justifies the overlap as it makes it more likely that a comprehensive evaluation for accuracy will occur. However, some organizations may need guidance to identify the specific area in which their

performance management needs to improve, essentially needing an evaluation triage, for lack of a better term. The overlap between the standards does not lend itself to parsimonious evaluation of a performance management system. This could, reasonably, hinder the use of the standards because organizations, which do not have dedicated evaluation units, tend to view evaluation as an add-on process requiring additional time and effort, the buy-in for which can be minimal. Time and resource constraints have led to poor or non-existent evaluation of performance management systems.

The theoretical basis for the standards is only implied through the assumptions listed in the Joint Committee's work as well as overall design of the standards. From a scholar's point of view, this presents the standards as no more than a set of pragmatic and intuitive guidelines, developed by evaluation experts. However, to a non-scholar, like an HRM practitioner, the standards can be an invaluable toolkit for conducting quality evaluation, instead of a set of abstract theories that are difficult to translate to real-life situations. Further, at first glance, one may see the standards as more relevant to performance reviews, which, as many experts have stressed, is only a part of performance management. However, upon further inspection of the standards, it becomes clear that they are far more related to performance management than performance measurement as evidence by their similarities with the requirements for effective performance management from Strebler et al. (2001), discussed in Section 3.1.

The standards have been applied widely over the years (Glasman & Martens, 1993; Howard & Harman, 2007; Loup & Ellett, 1997). The Joint Committee's (2008) publication, now in its third edition, offers guidance as well as practical help for applying the standards, of note are the checklist and list of key questions. The checklist, which also provides users with a rating template, has been used in many contexts. For example, in 2009, the checklist it was adapted to

the Iranian medical education context and administered to twelve medical schools (Ahmady et al., 2009). The research concluded that the standards were valid and appropriate for monitoring and refining an evaluation system. The list of key questions offered a comprehensive set of questions, the standard statement to which each question applies, and the targeted standard to which each statement applies (Joint Committee on Standards for Educational Evaluation, 2008). In essence, the Joint Committee created constructs, sub-constructs, operational definitions and scale items; a process commonly used in scale item generation (DeVellis, 2017; Gable & Wolf, 1993). The central position of this thesis is that perception of the Joint Committee's standards can be helpful in assessing overall perception of performance management. The next section describes perception as an influential factor in performance management implementation and presents three perception components that must be considered in evaluation for evidence-based management.

5. Perception and Performance Management Implementation

Thus far, research has shown that even well intentioned and well-designed performance management policies undergo a process of reciprocal adaptation, in which the objectives of performance management designers are re-interpreted to some degree by the implementers (Al Shirawi, 2015; Armstrong, 2015; Piening et al., 2014). This becomes problematic when there are discrepancies on the value of policies as they often result in a gap between what the policy mandates and what is actually done (Purcell & Hutchinson, 2007). Therefore, an understanding and measurement of the implementers' perceptions, and subsequent translation, of the policies are just as important as the quality of the policies themselves (Purcell & Hutchinson, 2007; Soldan & Nankervis, 2014; Southall, 2014).

Management academics chiefly agree that worker perceptions of HRM initiative like performance management, play a substantial part in HRM effectiveness (Den Hartog et al., 2013; Jensen et al., 2013; K. Jiang et al., 2017). The role that these perceptions play seems to depend on the level of inquiry; for instance, Liao et al. (2009)[12] found that perception of organizational support mediated the relationship between HRM and overall performance, whereas Bowen and Ostroff (2004) found that worker perceptions of HRM were antecedents of department-level performance. In any case, perception of HRM is a factor worthy of inquiry in evidence-based management (Wang et al., 2020).

Katsaros et al. (2014) provided an important definition of perception in the employment context that informed the parameters of this research. According to the authors, perception is "the process by which an employee organizes and interprets his/her impressions in order to give meaning to his/her environment and thus, it influences significantly his/her workplace behavior" (p. 37). This definition highlights an important aspect of perception, which is the interpretation of information in order to assign a meaning to it, positive or otherwise (van Dijk & Legalatladi, 2015). Therefore, objective reality is not necessarily the basis of a person's reaction to a stimulus rather, it is the interpretation of this reality (Kim & Han, 2017; Pettinger, 2010). Even so, for most people, perception constitutes reality and influences their behavior to a degree that has and continues to merit acknowledgement and investigation (McConnell, 1994; Ohemeng et al., 2018).

[12] To be specific, their study focused on high performance work systems.

For more than three decades, the literature has stressed the significance of the role of employee perception on their performance and consequently, organizational performance (Bowen & Ostroff, 2004; Glavas & Godwin, 2013; Pearce & Porter, 1986; Pihl-Thingvad, 2016; Weber & Weber, 2001). The relationship between behavior and perception was perhaps best explained by Guest (1997) who proposed that perception affects attitude, which, in turn, influences behavior (Alfes, Shantz, et al., 2013; Glavas & Godwin, 2013; Katsaros et al., 2014). With respect to HRM and performance management, Guest (2002) stated that "workers do not passively accept HR practices but actively evaluate and respond to them" (p.353). In fact, employees are likely to disregard organizational policies when their perceptions of management or the organization as a whole are unfavorable (Katsaros et al., 2014; Seiden & Sowa, 2011).

Fundamentally, perception of performance management can determine the success of its implementation because it is situated within a social context and has a personal dimension and potentially emotional experiences for employees (Lewis et al., 2015; McConnell, 1994; Mone & London, 2010; Özleblebici & Çetin, 2015; Robbins et al., 2013). Given the complexity of the relationship between perception and behavior, three components of perception were distinguished, namely; the situation, the perceiver, and the target (Ohemeng et al., 2018; Robbins et al., 2013). The situation refers to the context in which perception is developed, changed or reinforced. The perceiver is the person interpreting external information, and the target is the phenomenon that the perceiver is trying to interpret (Robbins et al., 2013).

Kinnie et al. (2005) found that differences in the type of work and / or characteristics of the perceiver can influence employee perceptions. Moreover, Alfes, Shantz, et al. (2013), with particular regard to engagement, argued that contextual factors are important in explaining why employees with the same level of engagement may not behave uniformly. The vast majority of

the literature treated characteristics of the perceiver and the situation as control or independent variables, while only a few considered them as mediators, moderators or antecedent variables (Bednall et al., 2013; Carriere & Bourque, 2009; Datta & Singh, 2018; Gyekye & Salminen, 2009; Profit et al., 2014; Ree & Wiig, 2019; Villajos et al., 2018). However, considering that studies have confirmed the disparity of perceptions between employee groups, it would be specious, in this thesis, to treat these characteristics as errors to be controlled, rather they were viewed as antecedents (K. Jiang et al., 2017; L. Jiang et al., 2017; Liao et al., 2009; Wang et al., 2020; Wright & Nishii, 2013b). Though the main focus of this thesis was on perception targets, the remainder of this section will discuss the situation and the perceiver, followed by a discussion of the perception targets relevant to this research in Section 6.

5.1 *The Perceiver*

The influence of *perceiver* characteristics like employee age, gender, work role, employment status, and previous workforce experience on perception were suggested by many like, Buss and Kuyvenhofen (2011), Dello Russo et al. (2013) and McConnell (1994). For instance, in an organization, which gives moderate support to its employess, some employees may perceive the support favorably if they have experienced poor support from previous employers, whereas employees new to the workforce may perceive the support poorly as they have no other point of comparison (Datta & Singh, 2018; Eby et al., 2000; George & Jones, 2012; Snyder & Stukas, 2007; Walk et al., 2013; Zhang & Liu, 2010).

Work Role. Research has shown that staff within each hierarchical level in an organization can have significantly diverse perceptions (Corley, 2004; Kang & Busser, 2018). This diversity is largely driven by the day-to-day tasks and level of power associated with each role (Kang & Busser, 2018). In this case, power refers to autonomy as well as decision-making authority. Senior managers are at the top of the organizational ladder, have considerable power, and are heavily involved in strategy formulation and organizational finances (Kang & Busser, 2018). Middle managers have some power and are involved in operationalizing the strategies developed by senior management, which requires procedural amendments and communication between senior managers and the rest of the organization (Al Shirawi, 2015; Kang & Busser, 2018). They are an integral part of policy implementation as they are tasked with strategy implementation, resource deployment, as well as staff relationships (Al Shirawi, 2015; Buss & Kuyvenhofen, 2011; Metheny, 2013; Özleblebici & Çetin, 2015). HRM operational staff like junior managers and administrative staff generally have little power and are involved in the implementation of middle management decisions as well as managing staff relationships (Kang & Busser, 2018).

Job autonomy differs between organizational hierarchies; those in senior roles tend to have much more autonomy than those in lower ranked roles which can mitigate the negative effects of their job demands (Chaudhary & Rangnekar, 2017). Consequently the former may have more favorable perceptions of the organization and its procedures (Chaudhary & Rangnekar, 2017; Datta & Singh, 2018; Zhang & Liu, 2010).

Middle managers and HR operational staff are often situated in the middle layers of an organization that are expected to cushion any friction between senior management and other staff; this is a job demand that requires a high level of communication skills (Chaudhary & Rangnekar, 2017; Osterman, 2009). Job demands are the "physical, social, or organizational aspects of a job that require sustained physical or mental effort and are therefore associated with certain physiological or psychological costs" (Demerouti et al., 2001, p. 501). This definition highlights that job demands may not always be stressful enough to cause alarm so, the concept of job stressors emerged to distinguish between harmful job characteristics and relatively harmless ones (McGonagle et al., 2015). Job stressors are factors that invoke a cognitive process associated with perceptions of a threat (McGonagle et al., 2015). These can be specific to the type of job, like the potential physical harm that many first responders[13].

Other stressors can occur across many job types and include role ambiguity, employee conflict and work overload. In either case, when facing a stressor, people will often use job resources to cope (Demerouti et al., 2001; McGonagle et al., 2015). Job resources refer to the supportive resources used specifically to cope with a stressor and can stem from within the individual, like their level of resilience or the organization through, for example, professional development (Lazarus & Folkman, 1984; McGonagle et al., 2015).

Research on the relationship of job role and employee perceptions has also considered other related factors, like tenure (number of years in the organization) and workforce experience, as they may contribute to the effects of job demands, stressors and resources on perception (Bakker & Demerouti, 2007; Demerouti et al., 2001). Indeed, findings are mixed, with some

[13] First responders include all personnel dispatched in an emergency like paramedics or law enforcement.

showing significant relationships between tenure or work experience while others finding little to no effect on employee perception (Butitova, 2019). This indicates that there may be an effect of tenure or workforce experience but only for certain populations and / or specific constructs.

Employment Status. Because performance management is not prone to the risk of automation to the same degree as, for example, factory workers, to reduce operational costs, many organizations moved away from secure employment (e.g. permanent contracts) to more precarious employment (e.g. fixed-term contracts) so they can respond to any market fluctuations without incurring significant transaction costs (McGovern et al., 2007). This, however, has reduced income security for many workers, and consequently, their loyalty to their employers (Hunt & Samman, 2019).

Precarious employment, in addition to income insecurity, is characterized by the view that the employee is, in effect, an independent contractor, with the exception of part-time workers on a permanent contract. The use of staff on a fixed-term or casual basis comes lowered operational costs by eliminating redundancy pay and paid leave, which is less than ideal for employees (Baluch, 2017; Gilfillan, 2018; Vosko et al., 2009). Precarious employment is also characterized by lower levels of legal protection, discretional organizational investment in worker wellbeing and career development, and a gender imbalance (Kim et al., 2012; Vosko et al., 2009). More women are in precarious employment than men which could be due to their desire to work flexibly or have fewer work hours to have time to care for their families (McGovern et al., 2007). Conversely it could also be because women, especially those of a certain age, increase operational costs from maternity leave (Kim et al., 2012; New Zealand Council of Trade Unions, 2013; Vosko et al., 2009).

Gender. The many studies on the influence of gender on perception have also presented mixed findings. For instance, Marinaccio et al. (2013) found a substantial difference between men and women in perceptions of control and level of peer support. However, Burke et al. (2015) found no statistically significant variations between men and women's perceptions of learning opportunities and their supervisor's behavior. A closer examination of the literature, like Burke et al. (2015), Datta and Singh (2018), De Sio et al. (2018), Ogunkuade (2018), and Civilidag (2014) indicated that such contradictory findings may be due to variance in research contexts, the achieved sample or the level of abstraction inherent in each phenomenon of interest.

Age. Like gender, tenure and workforce experience, research on the influence of employee age on their perceptions has found mixed results both in the extent and direction of influence. For example, Baruch (2004) found that older workers did not value competition as much as their younger counterparts who are still building their careers and trying to prove themselves to their employers. As such, younger workers tended to engage more with the organization's HRM, training and development, and networking. Other researchers found that younger workers had higher rates of intention to quit because younger staff members may be more critical of their organization's investment in their professional development and look for better opportunities elsewhere (Bellou, 2010). In addition, McGonagle et al. (2015) posited that perception of job stressors improves with age because people accrue job resources over time, like knowledge, peer support, and skills, that can mitigate the effect of work stressors.

Other studies, like Meyer et al. (2002), found little variance between generational groups in their perceptions of work opportunities and organizational communication. This could be because the findings referred to generational groups that may not have had experiences that varied enough to cause different perceptions of HRM or other related phenomena (Cucina et al.,

2018). Unlike earlier research, the three generations present in the current workforce are Baby Boomers, Generation X and Generation Y, which have had very different life experiences (Lapoint & Liprie-Spence, 2017). Dissimilar to other age categories, generations refer to cohorts that share, to some extent, key social definers (Christopher et al., 2018). Most Baby Boomers, born between 1945 and 1964, experienced the Women's Rights movement, space travel, the sexual revolution and traditional single income family dynamics (Zopiatis et al., 2012). Many Generation Xers, born between 1965 and the early 1980s, experienced corporate downsizing, dual income families, the Gulf War and the beginnings of digital advancement like dial-up internet (Green, 2010). Generation Y, also called Millennials, were born between the early 1980s and the early 2000s. Most of them came of age after the 9/11 terrorist attack changed global travel and have little memory of a pre-internet economy (Zopiatis et al., 2012).

Smith and Galbraith (2012) found that Baby Boomers had a preference for phone calls while millennials preferred text messaging, indicating that the medium of communication is important as well as the content. This is because millennials grew up in an era of information and communication technologies (Krywulak et al., 2009). Moreover, they are globally educated, view themselves as self-confident (often confused with entitled), highly optimistic, idealistic and value work-life balance (Srinivasan, 2012).

An important consideration for research on generational differences is the location and associated culture during one's upbringing; in essence, an employee's age may not be as important as where they were raised. For instance, in China, children born during the enforcement of the one child policy (introduced in 1979) grew up in a substantially stronger economy and received better education and work opportunities than their parents (Kadoya et al., 2018). Therefore, Chinese Millennials tend to value flexibility and interesting jobs, while the

older generations generally value financial security more (Roongrerngsukea & Liefooghe, 2013). This may not be the case for Millennials in Thailand, South Korea or Indonesia who witnessed high rates of business bankruptcy due to the economic recession during the 1997 Asian Financial Crisis (Roongrerngsukea & Liefooghe, 2013). As such, collecting data on the perceiver's age should also collect data on their upbringing, or at least, their location, to better understand their perceptions.

5.2 *The Situation*

As contingency theory and Armstrong's (2015) HRM system model (Figure 3) showed the *situation* component of perception can be subject to the influence of the external and internal environments of the organization. For instance, technological advancements can affect performance management and / or staff performance through innovative software for performance measurement, or increasing the speed and quality of task performance in general (Lischka, 2015). Further, an organization's structure in terms of its size, funding sources and nature of its service can affect the design and implementation of its performance management. For example, organizations that manufacture goods have different performance indicators than those that deliver knowledge or creative services like consultancy firms or health promotion agencies. For one, accurate measurement of performance can be easier when work tasks produce tangible outputs like the number of cars sold (Gerhart et al., 2000).

In knowledge-based industries, however, performance can be abstract and requires more qualitative assessments, which are susceptible to subjective bias. For example, community or social care services should not be measured solely on the number of families visited per month, rather, performance measurement should also consider the quality of the service delivered in terms of community engagement, awareness raising and behavioral changes. The issue is that the

measurement of these outcomes is problematic because they take a long time to develop and are difficult to link to particular service providers let alone individual employees.

Organization Size. Well-funded and larger organizations tend to have more sophisticated HRM and performance management practices, which may result in perceptual differences when compared to poorly funded or small organizations (Datta & Singh, 2018). This is commonly because the former tend to have more channels of communication, financial resources and a greater variety of supervisory styles to suit the variety of staff (Helms & Stern, 2001; McConnell, 1994). Nevertheless, the formality of the HRM commonly found in large organizations may lack a personal touch that some employees prefer (Rhoades & Eisenberger, 2002). Dekker and Barling (1995) argued that employees feel less valued in large organizations because the numerous formal polices reduce flexibility in addressing individual needs. Therefore, the larger the organization, the more meticulous the HRM staff must be to guarantee that the relational aspects of HRM, like trust-building, are not overshadowed by bureaucratic procedures (Reynolds & Helfers, 2018).

Industry Sector. The stability of an organization in terms of its funding or external competition can influence its employees' perceptions of, for example, job security (Helms & Stern, 2001). In general, industries can be categorized under five economic sectors; primary, secondary, tertiary, quaternary and quinary. The primary sector pertains to raw materials and the earth's natural resources, which includes farming, mining, and quarrying (Auerbach, 2019; de Grandbois, 2016; Kenessey, 1987; Thakur, 2011). Most employees in this sector are called 'red collar' workers because the nature of their work predominantly requires them to be outdoors. The secondary sector pertains to roles that directly add more value to the products of primary industries, which includes manufacturing, construction and other roles commonly called 'blue

collar' jobs (Auerbach, 2019; de Grandbois, 2016; Kenessey, 1987; Thakur, 2011). Most primary and secondary sector employees perform physically arduous but cognitively simple and repetitive tasks, the success of which is highly contingent on the natural environment. Tertiary industries pertain to production and service provision like trade, transport, and hospitality. Workers in tertiary industries tend to work both indoors and outdoors, and have more interaction with the consumers (Auerbach, 2019; de Grandbois, 2016; Kenessey, 1987; Roongrerngsukea & Liefooghe, 2013; Thakur, 2011). The Quaternary sector is also known as the knowledge sector which involves professional roles in finance, healthcare and entertainment. Both the tertiary and quaternary sectors contain roles that can be outsourced like Information Technology support services and other roles commonly called 'white collar' jobs. The quinary sector concerns industries in top level knowledge creation and leadership like scientific research and government (Roongrerngsukea & Liefooghe, 2013; Thakur, 2011). Employees in the quinary sector are commonly known as 'gold collar' workers due to their high level of education as well as their social, organizational and political power (Roongrerngsukea & Liefooghe, 2013). Gold collar workers are also known to prefer flexibility and autonomy, and they are required to use creative problem solving in their intellectually complex work (Roongrerngsukea & Liefooghe, 2013; Thakur, 2011).

The industry to which an organization belongs can account for great variance in employee perceptions across organizations due to the different job demands, stressors and resources. For instance, the demands and work pressures exerted on first responders, like law enforcement officers or paramedics, create a distinct working culture in which team work is potentially life-saving (Andersson Arntén et al., 2016). Such organizations also have strict task protocols and are subject to extensive government regulation which, though necessary, can

reduce the level of employee perceptions of autonomy (Andersson Arntén et al., 2016). Another example comes from Foster and Guttman (2018) in their exploration of perceived job security. The authors stated that in addition to the influence of personal circumstances or external labor events on perceptions of job security, the market conditions and industry characteristics also have an effect on employee perceptions. They explained that, for example, industries like entertainment and media are dominated by casual employment which can decrease perceptions of job security (Foster & Guttman, 2018). Further, perceptions of organizational commitment to staff can be influenced by exposure to international competition. To elaborate, in industries where outsourcing is a distinct possibility, employees may perceive that they are easily replaceable especially if the organization is going through tough economic times (Foster & Guttman, 2018).

6. Perception Targets

While characteristics of the perceiver and the situation are important, some characteristics of a perception target can affect perception transcending the effect of the situation or the perceiver. For instance, vague job descriptions are far more likely to be perceived, by any one in any organization, as unclear and confusing than precise job descriptions (Ohemeng et al., 2018). Moreover, consistent and fair reward distribution is more likely to result in positive perceptions of performance management rather than idiosyncratic processes that may lead to perceptions of favoritism or discrimination (Robbins et al., 2016). For the purposes of this thesis, the overall target of perception is performance management, however, given the complexity of performance management, several perception targets were considered to measure one's perception of performance management overall. Therefore, in addition to perceptions of the feasibility, propriety, accuracy and utility of performance management (the four evaluation standards, see

Section 4.4), five perception targets were identified as pivotal in gauging overall perception of performance management.

6.1 *Perception of Evidence-Based Practice*

As proponents of evidence-based management stated, performance management decisions should be based on explicit and meticulous use of robust evidence (Barends et al., 2014). Specifically, policies about performance management decision-making need to encourage and support the use of a variety of evidence sources to inform decisions as well as the use of the six steps as illustrated in Figure 2. The contemporary view of evidence-based management, as exemplified by Barends et al. (2014) built on the traditional view to include robust empirical research as well as practitioner expertise, contextual information, and stakeholder data, like those from client satisfaction surveys (Hodgkinson, 2012). As such, the new evidence-based management movement emphasizes evidence construction using a combination of scientific inquiry and problem-solving methods with acknowledgement of the highly politicized context in which the evidence applies (Bartlett, 2011; Briner et al., 2009).

As discussed in Section 2.4, evidence-based practice may not the norm in many organizations and some practitioners rarely devote much attention to the quality of the evidence they use to inform their decisions (Barends et al., 2014; Marler & Boudreau, 2017). This could be because they have poor perceptions of the importance of evidence-based practice or the support from the organization to use reliable evidence. Positive perceptions of the credibility, reliability and relevance of performance management are more likely to lead to proper implementation (Armstrong, 2015; Rousseau & Barends, 2011; Skinner, 2001). Moreover, positive perceptions of the support and incentives given by the organization for using evidence

are more likely to lead to evidence-based practice (Henderson et al., 2014; Koster, 2011; Rousseau & Barends, 2011).

6.2 *Perception of Organizational Communication*

Cited by 3,328 publications[14], the undeniable influence of Bowen and Ostroff (2004) article stems from its suggested distinction between HRM content and HRM process (Ostroff & Bowen, 2016). Their work suggested that perceptions of performance management center on the distinctiveness, consistency, and consensus of communication. Distinctiveness refers to the characteristics of a performance management policy that allow its link to outcomes to be clearly observable in the workplace. As such, employees need to perceive performance management to be relevant to their work, transparent and understandable (Bowen & Ostroff, 2004; Katou et al., 2014). Consistency refers to the constancy and predictability of the relationship between performance management policy and outcomes. Consensus refers to reaching a state of unified agreement among employee perceptions of performance management, as well as perceptions of fairness (Bowen & Ostroff, 2004; Jiang et al., 2012).

The seminal work by Bowen and Ostroff (2004) asserted that, whether by design or accident, organizational actions, decisions and outcomes inform employee perceptions, therefore, the content and communication process of organizational policies and practices are vital for achieving desired employee behavior (Ferris et al., 2008; Katsaros et al., 2014). Organizational communication of performance management matters refers to "the way HR policies and practices are communicated to employees" (Li et al., 2011, p. 1826). Consistent and transparent communication adds to the perception of organizational stability which can lead to positive

[14] Information from Google Scholar current as of March 2020

perceptions of job security, performance expectations and managerial competence (Bowen & Ostroff, 2004; Ohemeng et al., 2018). During times of organizational uncertainty or significant economic stress, effective organizational communication can facilitate the management of employee anxiety thereby reducing resistance to change (Katsaros et al., 2014).

In essence, perception of organizational communication concerns the content and process of performance management communication including formal and informal communication. Not only is it important for performance management activities be perceived as useful or important, they also need to be perceived as fair to keep workers engaged in their work and loyal to the organization (Colquitt et al., 2001; Malinen et al., 2013). Whereas poor perceptions of performance management communication can lead to increased occupational stress and burnout due to task confusion or job uncertainty (Gray & Laidlaw, 2004). For example, with regards to professional development opportunities, studies revealed that positive perceptions were inversely correlated with an employee's intention to leave (Morrow, 2011; Salleh et al., 2013). Moreover, workplace gossip, a component of informal communication, especially negative gossip is detrimental to the communication environment and can lower performance efficiency as well as collegial trust (Ellwardt et al., 2012).

Quite recently, Wang et al. (2020) built on the work by Bowen and Ostroff (2004) and noted that the rationale behind HRM decisions is important. This is because employees with similar perceptions of HRM content and process may still disagree about the motivation for the HRM system in their organization. A central assumption of HRM motivation, similar to perceived organizational support, is anthropomorphism; that is, the personification of an organization (Wang et al., 2020).

Other dimensions of organizational communication were proposed in the literature including an informational dimension and a relational dimension (Gray & Laidlaw, 2004; Jalalkamali et al., 2016). The informational dimension encompasses the worker's satisfaction with the content of any communication, for example its relevance to their role (Jalalkamali et al., 2016). The relational dimension focusses on the manner in which content is communicated and relates to the media used, the source of the information, the language used, and the directionality of communication (downward or two-way) (Jalalkamali et al., 2016). The latter dimension highlights the importance of encouraging the acceptance of the information communicated and engaging with it in a meaningful way to leverage change.

Positive perceptions of communication have also been associated with productivity because when performance communication is perceived to be constructive it can increase the employee's engagement with their work and, in turn, improve their performance (Zwijze-Koning & de Jong, 2007). Moreover, accurate performance feedback informs employees of their success and any potential room for improvement. The importance of this perception target is supported by several theories like expectancy theory, equity theory, and agency theory, previously discussed in Section 2.2. Undeniably, the two dimensions are codependent, for example, while perceiving performance feedback as accurate increases the likelihood of accepting it, perceiving it as constructive connects employees emotionally to their achievements and gives them an additional sense of purpose to their work (Ariely et al., 2008; Cole et al., 2006).

6.3 *Perception of Organizational Support*

Perception of organizational support is a concept that examines how employees develop opinions about the degree to which their organization values them and their work (Eisenberger et al., 2002; Rhoades & Eisenberger, 2002). An example of positive support perceptions include feeling assured their organization will help them should they need to deal with stressful situations (Worley, 2006). This perception target is, perhaps, most supported by organizational support theory and social exchange theory, previously discussed in Section 2.2. Similar to the observation from Wang et al. (2020) about the motivation aspect of HRM perception, Eisenberger et al. (1986) suggested that the perception of support is based on an individual's tendency to anthropomorphize the organization. During times of organizational uncertainty, customarily triggered by budgetary changes, internal restructuring or external changes to the market, positive perceptions of support are crucial (Katsaros et al., 2014). This is because they impact employee perceptions of the situation at hand and can increase the likelihood of their acceptance of and implementation of any necessary organizational changes instead of withdrawing from the organization (Katsaros et al., 2014).

According to some social exchange theorists, perception of organizational support can be influenced by discretionary choice. To elaborate, if the employee thinks that favorable treatment or support from their organization is not influenced by law or union negotiations but based on discretionary choice, perception of support will be greater (Rhoades & Eisenberger, 2002). It was also posited that if the employee feels that unfavorable treatment from the organization is due to matters outside their control, then perception of support will not decrease (Rhoades & Eisenberger, 2002).

Scholars have noted that when workers see their supervisors as organizational agents, perceiving favorable treatment from them will be connected to the organization as well (Rhoades & Eisenberger, 2002). As recent research has stressed, the relationship between an employee and their supervisor is a significant contributor to their perceptions at work (Oberoi & Rajgarhia, 2013; van Dijk & Legalatladi, 2015). However, if the employee perceives the support to be at the supervisor's own discretion then the positive perception of support is connected with the supervisor only. It is worth noting that though an employee may distance the supervisor from the organization, they may never be able to separate them completely as the supervisor will always have obligations to the organization (Eisenberger et al., 1986; Eisenberger et al., 2002; Rhoades & Eisenberger, 2002).

6.4 *Perception of Work Engagement*

As Alfes, Shantz, et al. (2013) suggested, employees who view their organization's HRM favorably are likely to have high levels of engagement, task performance and, extra-role behavior like voluntarily helping others. Bakker and Leiter (2010) suggested that engaged employees are exemplified by the higher level of discretionary effort they exert compared to their disengaged counterparts. Given the many findings that support the positive relationship between engagement and performance, management practices need to target the improvement of engagement. Although the terms employee engagement and work engagement are typically used interchangeably, work engagement specifically refers to the connection between the worker and their job, whereas employee engagement is a broader term that also entails the employee's relationship with the organization (Schaufeli & Bakker, 2010; Xanthopoulou et al., 2012).

Because the focus of the thesis is on task performance (implementation of performance management), work engagement was more suitable as a target of perception.

Seminal work by Schaufeli and Bakker (2010) showed that work engagement is a positive work-related condition is comprised of vigor, dedication, and absorption. Vigor refers to high levels of mental and physical energy and resilience at work, whereas dedication refers to a state of work involvement, and feelings of significance, enthusiasm and interest in cognitive challenge. Lastly, absorption refers to being fully focused, happily captivated in work and a sense of time passing swiftly (González-Romá et al., 2006; Schaufeli et al., 2002). Engaged workers tend to discover novel ideas and solutions more so than their non-engaged peers, bringing innovation and creativity to the workplace (Venz & Sonnentag, 2015). Therefore, in the context of performance management implementation, an organization must promote the implementer's engagement with their work to increase the likelihood of proper implementation. This construct is perhaps most supported by agency theory and commitment theory, previously discussed in Section 2.2.

After the many years of research on engagement, it has been suggested that engagement might be a core mechanism that explains five employee behavior concepts common in the literature (Alfes, Truss, et al., 2013; Rich et al., 2010; Schaufeli et al., 2002). The first is organizational citizenship behavior, which involves positive worker behavior beyond their core tasks without petitioning for compensation (Bateman & Organ, 1983; Tsai & Wu, 2010). One way to encourage this type of behavior is through demonstrating positive behaviors towards the employee implying commitment to worker wellbeing, personal and career development, and their performance success (Salleh et al., 2013; Tsai & Wu, 2010). Recent research found that the higher the level of organizational citizenship behaviour, the lower the level of job burnout and

turnover (Zhou et al., 2014). The second concept is organizational commitment, which is "the relative strength of an individual's identification with and involvement in a particular organization" (Mowday et al., 1979, p. 226). Studies found organizational commitment to be negatively associated with employee turnover (Caykoylu & Egri, 2011; Riketta, 2002). Though related to work engagement, organizational commitment is akin to employee engagement in that it relates to the relationship between the worker and the organization.

The third concept prevalent in employee behavior research is job satisfaction, which is perhaps the most frequently measured phenomenon in organizations, commonly through employee satisfaction surveys (Akdol & Arikboga, 2015; Musah et al., 2017). Essentially it is the affective positioning an employee has towards their job features (Lu et al., 2012). In contrast to work engagement, which implies activation (vigor), job satisfaction is passive and indicates satiation. The fourth concept is workaholism, which refers to the compulsive dedication to work and can be mistaken for work engagement (Shkoler et al., 2017). The key difference is in the nature of the dedication to work; employees engaged in their work are so because the work is fun and challenging, rather than because of an obsessive need for perfectionism or pressure to meet work targets (Bakker & Leiter, 2010; Shkoler et al., 2017). Additionally, while work engagement leads to improved performance and employee well-being, workaholism can lead to burnout.

The fifth concept is burnout, a state of exhaustion and cynicism, rather than the positive mindset associated with work engagement (Bakker et al., 2014; González-Romá et al., 2006). Unlike those who suffer from burnout, employees who are engaged in their work are energized and feel connected to their work; seeing it as a challenge, as opposed to stressful. Therefore, while burnout has been related to work engagement, they are, in fact, negatively correlated concepts.

6.5 *Perception of Others' Behaviour*

Commonly, employees engage in two types of behavior; performance or withdrawal. Much research was dedicated to exploring staff behavior, but only a fraction focused on how that behavior is perceived by co-workers and how that perception can affect the perceiver (Eder & Eisenberger, 2008; Felps et al., 2006; Smith et al., 2016). This is problematic because perceptions are formed based on what people see and experience themselves as well as hearsay (McConnell, 1994). The following describes how perception of others' performance and withdrawal can affect the perceiver.

Perception of Others' Performance. Performance can be classified as in-role and extra-role performance. In-role performance, also known as core task performance, includes achieving the tasks listed in a position description (Rasheed et al., 2015; Zhu, 2013; Ziegler & Schlett, 2016). It constitutes a more objective view of employee performance in that it deals with tasks that are articulated both verbally and in writing prior to role commencement. Its measurement is, arguably, far easier than that of extra-role performance, and usually done formally on an annual basis during a performance review.

Extra-role performance, commonly associated with organizational citizenship behavior encompasses all other actions that benefit the work environment such as helping a teammate with their work, low levels of complaining, and accepting inconvenient changes without hassle (Zhu, 2013). Though extra-role actions may not be prescribed like in-role actions, they are extremely important for the functioning of the work environment as a social system. Measurement of extra-role behavior faces challenges of quantification, hence it is not commonly assessed well, if at all (Cheng & Coyte, 2014; Zhu, 2013).

Perception of Others' Withdrawal Behavior. Employee withdrawal entails behavior that distances the employee from their work or the organization as a whole (Berry et al., 2012; Eder & Eisenberger, 2008). Within the literature on withdrawal behavior, unplanned lateness, unplanned absence and turnover are the most prevalent concepts, arguably, because they are easier to quantify than psychological withdrawal. It has been shown that these behaviors can cost organizations billions per year; in fact, it was calculated that unplanned absence can cost organizations around 8.7% of their payroll costs (Eder & Eisenberger, 2008; Klatt & Klatt, 2011; Society for Human Resource Management & Kronos, 2014). It is important to note the distinction between lateness or absence and unplanned lateness or absence. The latter are far more costly as the organization would less time to find additional labor to cover for the employee inevitably leading to delays in task performance (Berry et al., 2012; Sagie et al., 2002). For example, unplanned absence can lead to hasty redistribution of tasks, which leads to added stress and a reduction in performance quality.

Unlike unplanned lateness and unplanned absence, turnover may not always be a negative outcome. It can result in freeing up funds for more skilled employees to join an organization, however this may only be the case if the organization were losing under-skilled or under-performing employees in exchange (Bamford & Hall, 2007; Tracey & Hinkin, 2008; Winterton, 2004). Though turnover can bring about desirable changes that come with new talent acquisition, turnover of critical employees, like productive or tenured staff, can be detrimental. For one, tenured staff have so much institutional knowledge that is invaluable in organizational operations (Nagadevara et al., 2008). Also, losing productive staff not only means a reduction in overall productivity and market competitiveness but also a reduction in the morale of the remaining workers. The exact financial costs of turnover can vary between organizations but

they usually stem from the same processes like advertising a vacant position, lost productivity, temporary cover, and training new staff (North et al., 2013; Sagie et al., 2002; Tracey & Hinkin, 2008). The cost of losing productive employees is often augmented by task complexity due to the often steep learning curve that new staff experience especially in large organizations that have multiple complex operations (Tracey, 2010; Tracey & Hinkin, 2008). Of particular interest for this research is the loss of productivity associated with turnover; though it can be difficult to assess precisely, the burden of turnover is common in performance discourse. Tracey and Hinkin (2008) provided a concise account of the ways in which productivity can be lost due to turnover; echoed by many other scholars like Vnouckova and Urbancova (2016), North et al. (2013), Walker (2017), James and Mathew (2012). First the employee who will soon depart is highly likely to be less productive than others, either because they are psychologically disengaged or because they must shift their focus from task performance to task handover. Second, the learning curve required for new staff can affect productivity, not just of the newcomer but also the staff supporting them (Tracey & Hinkin, 2008).

Organizational peers can be a great source of support and motivation, however, perceptions of their withdrawal can negatively impact the performance of others (Robinson et al., 2014). As previously discussed in Section 2.2, equity theory suggests that people can be influenced by others' behavior especially when they receive similar rewards but for different levels of performance (Adams, 1965). This can, in turn, affect the perceiver's views of organizational procedural justice, and the need to adhere to policies, which, consequently, leads to a reduction in their own efforts to match their perceptions (Liden et al., 2004). Though the 'offender' in this case is the co-worker demonstrating poor performance or withdrawal behavior, the perceiver may respond by decreasing their contributions to the organization because it is seen

as the source of procedural injustice. Therefore, perception of organizational inaction becomes a catalyst for the perceiver's poor performance (Hung et al., 2009).

The focus on perceptions of others' withdrawal is imperative because research has revealed that people are more likely to be affected by negative behavior because adverse experiences are psychologically more powerful than positive ones (Robinson et al., 2014). In fact, an employee's perception can be affected by their co-workers' counter-productive behaviors, even if their own work is not directly affected. Indeed, they can be influenced by simply noticing the behavior frequently or hearing complaints from other workers (Ferguson & Barry, 2011).

It must be acknowledged that while effective performance management can reduce withdrawal, not all reasons for withdrawal are within an organization's control (Hays et al., 2012). For example, newly married employees, without children, are more likely to leave their job if their spouse receives a good work opportunity that requires them to relocate. Conversely, those with children are more likely to stay because relocation is harder if the children are established in their respective education settings (Nagadevara et al., 2008). In addition, younger employees and recent graduates are more likely to leave their jobs in pursuit of different opportunities as they are at the early stages of career exploration (Allen et al., 2003; Čikeš et al., 2018). Therefore, perceptions of the same withdrawal behavior among different co-workers can have contrasting meanings for the perceiver of their organization's performance management.

6.6 *Perception of Propriety, Utility, Feasibility and Accuracy.*

The four evaluation standards, presented in Section 4.4, were also considered as perception targets important for any performance management evaluation efforts. They can be applied in the context of performance management evaluation, or in relation to the perception of evidence-based practice, perception of organizational communication and perception of organizational support. For example, the fairness of organizational support relates most to propriety, while the relevance of organizational communication relates most to utility. The next section provides the logic model used to synthesize the components of performance management perception (targets, perceiver, and situation) and their hypothesized relationships.

7. Performance Management Perception Logic Model

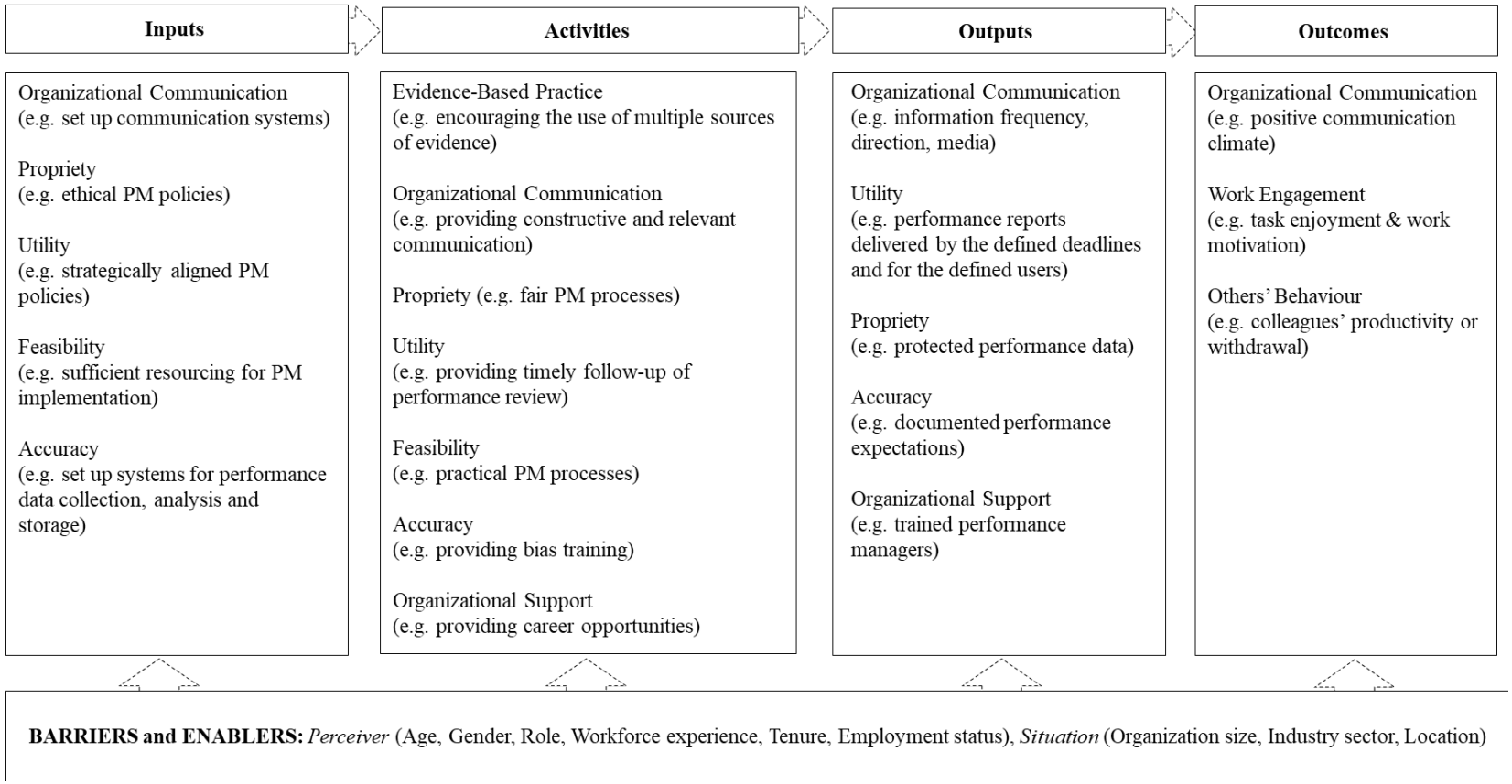
The literature suggested that the level of inquiry and the suitability of the variables measured in assessing the impact of HRM and performance management on organizational performance as an issue in the extant research (Paauwe & Richardson, 2001; Wright & Boswell, 2002). To elaborate, organization-level outcomes, when used to assess the impact of individual-level management practices on employee performance, are troublesome as they are too distal to make any meaningful inferences (Alfes, Truss, et al., 2013; Paauwe & Richardson, 2001). Further, several researchers have stressed the error of inferring HRM impact simply by examining inputs, like policy documents, and outcomes, like staff performance data or an organization's financial data (Alfes, Truss, et al., 2013; Conway & Monks, 2009; Wright & Nishii, 2013b). Lessons from evaluation, in particular the theory of change and logic models (see Section 4.3), stress the need for the inclusion of proximal indicators, like output variables, as intermediary indicators within the overall HRM-Organizational performance relationship

(Frechtling, 2015; Jiang et al., 2012; Kekahio et al., 2014; McLaughlin & Jordan, 1999; Podsakoff et al., 2007).

A logic model was used as a framework for the deductive thematic analysis of the literature to identify the performance management perception targets to use in the empirical studies in this thesis. Though a logic model has more components, the following shows the components that were most helpful for the purposes of the empirical research stages that followed this literature review. Shown in Figure 8, the perception *targets* were grouped into perceptions of performance management inputs, activities, outputs, and outcomes. Characteristics of the *perceiver* and the *situation* were categorized into the barriers and enablers of perception. These are antecedent factors that may influence the perception targets and were, thus, used in the empirical stages of this research as independent variables during data analysis. Some of these factors may be outside the control of the organization nevertheless an understanding of their influence and adapting accordingly can greatly increase the likelihood of successful organizational intervention. The arrows shown in Figure 8, which usually indicated causal relationships, have a dotted outline to further indicate that they are hypothetical, as causality was not tested in this research.

Figure 8

Simplified Logic Model of the Conceptual Perception Targets and Characteristics of the Perceiver and the Situation



Note. The arrows represent the presumed influence of one component to the next. These presumptions, educated by the literature, formed the basis of correlation hypotheses presented in the survey studies.

8. Chapter Summary

This chapter expanded on the research context set in Chapter 1 and discussed HRM, performance management, evidence-based practice and the relevant evaluation approaches in greater detail. In particular, the chapter presented a discussion of how conventional wisdom in performance management limits its effectiveness as well as the adoption of evidence-based practice. The push for evidence-based practice in performance management, which implicitly mandates frequent inquiry into not only what works, but what works best, was prompted by the frustration with the unmet expectations of performance management. However, as discussed in this chapter, evidence-based performance management remains hindered by the limited use of reliable evidence from academia and evaluation for the multitude of reasons presented in this chapter. Key among them were the practitioner and academic divide, and insufficient use of evaluation data. In order to bridge the gap between academia, performance management, and evaluation for evidence-based practice, guidance from participatory evaluation approaches, a theory of change, and Patton's utilization-focused evaluation influenced this research.

The chapter also presented performance managers' perception as an influential factor on their implementation behavior. Three components of perception were noted; the perceiver, the situation and the target. For the purposes of this thesis, the chapter presented the performance manager as the perceiver, the organization as the situation, and performance management as the target. Importantly, considering the complexity of performance management, as a target of perception, it was found that it actually subsumed several perception targets. The thematic analysis, using a logic model, resulted in nine perception targets for measurement in any inquiry into performance management.

The nine targets were recommended as the basis for the EMI scale developed by the empirical research stages presented Chapters 4, 5 and 6 in this thesis. The thematic analysis also showed that characteristics of the perceiver, like age or work role, as well as characteristics of the situation like organizational size as antecedents of the perception targets. These characteristics were used as the demographic items, and treated as independent variables in the MANOVA for Survey 1 and Survey 2.

Based on the findings of the literature review, it was evident that the scope of this thesis should prioritize the empirical determination of the performance management perception targets, the development and validation of a scale to measure them and, adding to the body of knowledge on the perceiver and situation characteristics that can affect the perception targets. Three empirical studies were proposed in alignment with three research stages of the EMI scale's development and validation. A discussion of the research design, the empirical studies and the analytical approaches is provided in the next chapter.

Chapter 3: Methods

This chapter describes and justifies the methodological approaches used in this thesis. First it presents the mixed methods research design and its three stages. Second it details the three empirical studies in terms of their purposes, sampling, data collection, and analytical approaches. The chapter concludes with a brief outline of its funding, ethics requirements, and a summary.

1. Research Design

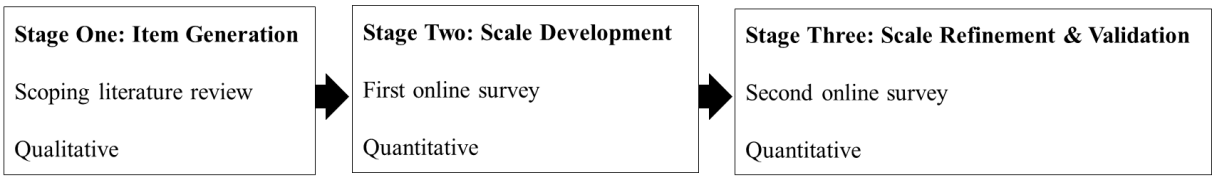
As discussed in Chapter 1, the thesis was guided by an empiricist epistemology, a theory of change, and utilization-focused evaluation. The literature review in the previous chapter showed that the identification and measurement of perceptions of performance management held by performance managers is a priority area of research into evidence-based performance management. To address this need, a sequential mixed methods research design, shown in Figure 9, with an emphasis on quantitative data was used for the purposes of triangulation, complementarity, development and initiation (Creswell & Plano Clark, 2011; Greene, 2007). The basic premise of mixing methods was that the use of both quantitative and qualitative methods offers a superior understanding of a research problem than either approach can do alone (Creswell & Plano Clark, 2011).

The purpose of this design was fourfold, as informed by Creswell and Plano Clark (2011); first, triangulation, which entails the corroboration the overall findings by using qualitative and quantitative data. The second was complementarity, which is to enrich the interpretation of the results from a qualitative and quantitative perspective. The third was development, particularly to allow the development of one study based on the findings of the previous one. The fourth was initiation, which relates to the discovery of new insights into the

perception of performance management as a social phenomenon that can influence performance management implementation (Creswell & Plano Clark, 2011; Greene, 2007). Because experimentation was beyond the scope of this thesis, the quantitative data was descriptive/correlational in nature (Creswell, 2003; Creswell & Plano Clark, 2011; Polit & Tatano Beck, 2017; Walker, 2005).

Figure 9

Research Design



Note. The three research stages are outlined along with their respective study and data type.

Using the commonly cited framework of scale development by DeVellis (2017), the following steps were undertaken to create a measurement scale for performance management perceptions; (1) determine the subject, (2) generate the items, (3) choose the measurement format, (4) review the items, (5) consider validation items, (6) administer the items, (7) assess the items and, (8) enhance scale length. It must be clarified at this point that Step 1 was done throughout the three research stages as well as the narrative literature review in Chapter 2. To elaborate, through the narrative literature review, the subject (performance management perception targets) was conceptually identified then operationalized in Study 1 (the scoping literature review) in order to generate scale items. Following which, in Studies 2 and 3 (the online surveys), the subject was empirically determined. Therefore, steps 2 to 5 were undertaken in Study 1, while steps 6 to 8 were done in Study 2 and Study 3.

1.1 *Stage 1: Item Generation*

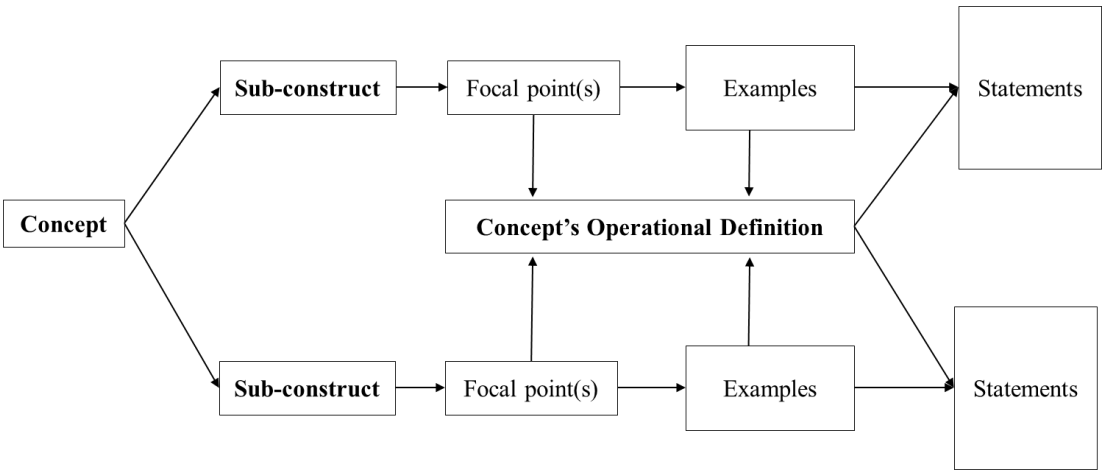
Guided by theory of change, the literature was synthesized, in accordance with the logic model, presented in Figure 8 in Chapter 2. The model provided a framework for the deductive thematic analysis from which nine performance management perception targets were identified and used as a foundation for item generation. The Joint Committee provided a detailed guide for the evaluation standards in which their sub-constructs, operational definitions, and key questions were given (Joint Committee on Standards for Educational Evaluation, 2008). These key questions were used as items in the EMI scale along with other items generated in Study 1. The work of Gable and Wolf (1993) was also used to guide item generation. Similar to DeVellis (2017), Gable and Wolf (1993) suggested that the identification of the concepts to be measured should be the basis of item development and provided their own detailed list of steps to follow. Importantly, they argued that, although the ultimate determination of the concepts to be measured must be through a thorough examination of empirical data, a theoretical basis of these concepts provides a vital foundation because it provides the reasoning for initial item generation as well as guidance for interpreting the empirical data.

For the purposes of this thesis, a few of Gable and Wolf's (1993) steps, shown in Figure 10, were used. The first step required identifying and describing the concepts to be measured, i.e. the performance management perception targets. Based on their descriptions, the second step entailed identifying and describing at least two sub-constructs for each performance management perception target. During the third step, an operational definition for each performance management perception target was derived by explicating the foci of the sub-constructs and creating a list of observable examples. In the fourth step, a list of semantically variable statements was generated for each example in line with the operational definition of each

performance management perception target. Since the evaluation standards, their respective targeted standards, operational definitions and items were provided by the Joint Committee (2008), the steps from Gable and Wolf (1993) were most applicable to the remaining five performance management perception targets (evidence-based performance management, organizational communication, organizational support, work engagement and others' behaviour).

Figure 10

Item Generation Steps



Note. Item generation steps adapted from Gable and Wolf (1993)

The four steps from Gable and Wolf (1993) were followed to create new items but only after a search for existing items showed that several new items were needed. This search was done in the first study in which a scoping review of the literature was carried out to identify valid scale items, which could be added to the list of statements. The method for Study 1 is discussed in Section 2 of this chapter.

The statements were then placed on a six-point Likert scale asking a respondent to indicate their level of agreement with each statement. Statement wording was based on three considerations; first, the survey response cognitive model, which suggests that, in responding to survey questions, a person goes through several cognitive steps; question interpretation, information retrieval, judgement formation, response formatting and response editing (Sudman et al., 1996). The number of words per question, content like jargon, response format, survey layout, item order, as well as the number of items to complete can affect the cognitive steps. The second consideration was for response bias, which refers to survey respondents' tendencies to respond inaccurately (Barratt et al., 2015; Donaldson & Grant-Vallone, 2002). There are many forms of response bias and though complete prevention may not be feasible, it is important to mitigate it as much as possible. Negatively worded statements were used to reduce the likelihood of agreement bias, and a six-point Likert scale was chosen to reduce the likelihood of automatically opting for neutral responses. The third consideration was for the BRUSO criteria for effective items (Peterson, 2000). According to this set of criteria, survey items must be Brief, Relevant to the research, Unambiguous, Specific (e.g. no double-barreled questions), and Objective (e.g., no leading questions).

The validity of the EMI items was first assessed in terms of content / face validity in particular. To ensure that the items satisfied the BRUSO criteria, the supervision team and research advisory panel reviewed the nine perception targets, operational definitions, and items. Following which, the revised items were built in an online survey platform and tested with a group of colleagues at the university to gage the appropriateness of the survey format and to fix any technical issues.

1.2 *Stage 2: Scale Development*

As guided by empiricism, the development of the EMI scale, done during Study 2, needed to go beyond the generation of items. That is, it was important to determine the perception targets through empirical data in order to consider the scale developed enough for refinement and validation in Stage 3. It must be reiterated, at this point, that although the development of the EMI scale was primarily empirical, the theoretical guidance provided in Stage 1 was invaluable in generating hypotheses for factor correlations and meaningful data interpretation. A detailed account of the methods for Study 2 is provided later in Section 3 of this chapter.

1.3 *Stage Three: Scale Refinement and Validation*

Based on the results of the first survey, the EMI scale was refined and administered to another group of respondents in Study 3 (Survey 2). Data analyses were used to further refine the scale, ascertain construct validity and estimates of reliability for internal consistency, and estimate the generalizability of the findings. A detailed account of the methods for Study 3 is provided in Section 4 in this chapter.

Relevant to Stages 2 and 3, a balance was needed between parsimony and plausibility. Scale development and refinement often require item reduction (parsimony) in order to remove any superfluous items or factors. Nevertheless, to sufficiently account for data patterns in a meaningful way, considerations of plausibility were just as important for this research. As such, item reduction was based on both the results of the statistical analyses and an examination of conceptual congruity.

2. Study One: Scoping Literature Review

2.1 *Study Purpose*

The purpose of this study was to generate the items for the EMI scale, specifically, it aimed to find relevant scale items for five of the nine perception targets, that can be used in the EMI scale, and generate new items should no such scale items exist. An initial search was carried out to identify systematic reviews, scoping reviews and meta-analyses that may answer the study's question: 'What quantitatively validated scale items exist to measure five of the performance management perception targets identified in the literature review?'

Since no relevant literature reviews were found, it was evident that a scoping literature review study was needed to generate items for the EMI scale. The scoping review method was chosen instead of a systematic review or meta-analysis as it was the most appropriate method for synthesizing literature that has yet to be comprehensively reviewed (Arksey & O'Malley, 2005; Khalil et al., 2016; Levac et al., 2010; Wilson, 2014). The scoping review followed the Joanna Briggs Institute method as described in the JBI Reviewer's manual in which the following had to be explicated; inclusion criteria, search terms and databases, search strategy, publication selection, and results (Aromataris & Munn, 2019). The results of this study are presented in Chapter 4; the remainder of this section outlines the remaining four requirements of the JBI method.

2.2 *Inclusion Criteria*

The following criteria for publication inclusion were used to ensure relevance of the search results and to make the search feasible for one reviewer. Consistent with scoping studies, the search terms and inclusion criteria were broader than those of a systematic review (Pham et al., 2014). Though as the study progressed the terms and criteria were refined.

Type of Participants. Though the overall research focused on the performance management perceptions held by the implementers, the type of participants in the literature was restricted to employees in general to avoid missing scale items that can be adapted to this thesis.

Concept and Context. The review only considered publications that described validated scales for performance management most relevant to the measurement of the perception targets. Specifically, the application of the desired scales had to be in the context of organizational performance, employee performance, and vocational learning and development.

Publication Characteristics. The initial search included only full-text literature, literature in peer-reviewed sources published between 2004 (the year of Bowen and Ostroff's seminal paper) and 2017 (the year this study was conducted). Both electronic and hard copy formats were accepted but only in English. Moreover, only empirical measurement studies with explicit quantitative data collection and analysis were included. Quantitatively reliable, valid and conceptually sound scales were included to help generate the items of the EMI scale for testing in the first survey study.

2.3 *Search Terms and Databases*

The terms were chosen based on the narrative literature review results and thesauri-based expansion. The databases used were well known to be user friendly, accessible, and collectively provided a comprehensive pool of evidence.

Search Terms. In title or abstract: Scale* OR instrument* OR surv* OR tool* OR questionnaire* OR metric* OR measure* OR assess* OR evaluat* OR test* OR examin* OR review* OR audit AND valid*OR reliab* OR sensitiv* AND Personnel OR staff OR work* OR job OR employ* OR role* OR team OR HR OR human resource*AND evidence-based HR* OR EBHRM OR EBMgt OR evidence-based manage* OR organizational communication OR

organizational communication OR perce* OR perceived organizational support OR perceived organizational support OR engage* OR behavi* OR perform* OR produ*.

Search Databases. Only multi-disciplinary, peer-reviewed, and conceptually relevant databases were used. They were: (1) University of Melbourne Library catalogue, (2) Science Direct, (3) Directory of Open Access Journals, (4) JSTOR Journals, (5) WILEY Online Library, (6) SAGE Journals, (7) Journals @ OVID, (8) ERIC, (9) Business Source Complete, and (10) SocINDEX with full text.

2.4 *Search Strategy*

An initial limited search of ERIC full text and Business Source Complete using the previously mentioned search terms was done. This search was followed by text analysis of the titles and abstracts to gain an understanding of amount of duplication and appropriateness of the search terms and date range. Afterwards, a second search was conducted within the same date range but across all the previously stated databases, and with more sophisticated search terms. The results of the second search were subjected to a title then abstract review based on which duplicates and publications not fitting the purpose of the review were discarded. The remainder was used for a full-text review during which a second round of exclusion was done to discard any publications not fitting the review purpose. A third search was done by examining the reference list of each included publication, which was not restricted by a date range. There was only one iteration of the whole search strategy.

2.5 *Publication Selection and Data Charting*

All publications were manually examined to extract data on scale development, scale items, and scale validation. The data extracted were entered into a Microsoft Excel database for coding and analysis. The extracted text was coded by; author information, location of the study, sample size, main analysis, and the related performance management perception target. To be selected in this study, in addition to focusing on one or more of the perception targets, the publications had to: (a) include scale items, be it the full list or sample items, (2) provide a clear outline of study methodology, (3) include scale items designed for staff participants rather than only for patients, clients, customers, students or parents.

3. Study Two: First Online Survey

3.1 *Study Purpose*

The primary aim of this study was to empirically determine the performance management perception targets by exploring the underlying structure of the survey data in order to infer latent factors. Second, the study aimed to reduce the large item set to a more manageable and defensible one to increase the feasibility of using the EMI scale by the intended users (performance managers). The third purpose was to explore antecedent factors that can influence the performance management perception targets, i.e. characteristics of the perceiver and the situation.

3.2 *Study Participants*

This study used purposive sampling to produce a sample that can be logically presumed, as best as possible, to be representative of the target population. Because the research was concerned with a specific population, random sampling methods would have been less

appropriate. As the focus of this research was performance management, the target population consisted of performance managers. There was no restriction on the country or industry in which participants worked, only that they spoke English and were employed in a performance management role at the time of the survey.

Sampling. An adequate sample size was of primary concern for this thesis because sufficient statistical power is essential for observing true data relationships (Kyriazos, 2018; Tabachnick & Fidell, 2014; Wolf et al., 2013). There was no shortage of sample size recommendations so, to meet the purposes of this thesis, a target sample was chosen based on two considerations; the size of the target population, and the type of analysis.

Population Size. According to the Australian Bureau of Statistics, *Specialist Managers*[15] are those who plan and coordinate organizational functions including human resources (Australian Bureau of Statistics, 2013). As of 2017, over 700,000 people were employed as specialist managers in Australia, 51,300 of whom were HR managers (Department of Employment, 2017; Department of Employment Skills Small and Family Business, 2018). Moreover, there were 1,673,900 administrative workers in Australia. However, given that the survey was open to an international participant pool, the size of the population was considered infinite, so, recommendations based on the analysis type were considered.

Analysis Type. As discussed later in this chapter, Study 2 data were analyzed using exploratory factor analysis. Guidelines for sample size in factor analysis are generally stated in terms of the minimum sample size, *N*, or the minimum ratio of participants to the number of variables *p*. (McCallum et al., 1999). Cut off criteria for sample sizes vary considerably from at

[15] Australian and New Zealand Standard Classification of Occupations, 2013, Version 1.2

least 100 to at least 500 participants (Osborne & Costello, 2004, 2005; Tabachnick & Fidell, 2014). As do the recommended ratios ranging from five to twenty participants per survey item (Boateng et al., 2018; Comrey, 1988).

Furthermore, Comrey and Lee (1992) provided a general scale for sample sizes in factor analysis in which 100 participants = poor, 200 = fair, 300 = good, 500 = very good and, 1000 or more = excellent. Given these recommendations and that the number of items from Stage 1 was 130, the target sample was estimated to be between 300 and 1,300 participants.

Recruitment. The AHRI[16] was contacted and asked to distribute the survey to its members in exchange for a report of the findings, however, the request was denied. Therefore, online recruitment strategies, described next, were used instead.

3.3 *Data Collection*

Software. The survey was created using the University of Melbourne's access to Qualtrics Research Core and its links were distributed through Prolific Academic (ProA) and LinkedIn[17]. Qualtrics Research Core was chosen as it was a prevailing survey engine that allowed users to create and distribute surveys for a variety of research and teaching purposes. LinkedIn, an employment-focused online platform was chosen as it operated via websites and mobile applications with a primary purpose to build career paths, networking and sharing professional and academic news. A link to the survey was advertised through the researcher's LinkedIn profile in early 2018. ProA, a participant sourcing service was also used to distribute

[16] The Australian Human Resource Institute was chosen because it was a membership-based organization for HR practitioners and students. It provided learning and development resources to over 20,000 members with whom it also conducted research to inform HR best practices.
[17] As approved by the University of Melbourne Human Research Ethics Committee

the survey. ProA has been shown to be the superior option when choosing a crowd sourcing platform (Palan & Schitter, 2018). Compared to other platforms, like Mechanical Turk, ProA provided higher levels of transparency to participants in terms of survey purposes and tasks, as well as to researchers in terms of the subject pool (Peer et al., 2017).

Launched in 2014 by graduate students from Oxford and Sheffield universities, ProA was specifically geared towards academic research. It provided a participant screening process allowing the researcher to invite only those who fit the characteristics of the target population (Palan & Schitter, 2018). ProA required participants be reimbursed for their time, the amount for which was calculated by the ProA software based on the time needed to complete the survey, in this case 20 minutes. Importantly, the reimbursement calculation was provided to the researcher prior to survey distribution. Moreover, any reimbursement could be only given after certain requirements of the research were met, like completion of the survey, accurate response format, and time taken to complete the survey.

The platform allowed the researcher to view the data prior to validating the response and releasing the reimbursement to the respondent. The ProA platform linked to Qualtrics seamlessly, and directed each respondent to and from the survey with minimal hassle. To protect the interests of the participants, if survey requirements were met but a submission was not validated by the researcher, the platform would automatically validate the submission after a two week period. This stipulation was stated clearly to the researcher as was the requirement to provide a justification for rejecting a submission, should that occur.

3.4 *Data Analysis*

This study used exploratory factor analysis, Cronbach's alpha calculations, and Multivariate Analysis of Variance (MANOVA).

Missing data. Due to the large number of scale items, the few demographic items were placed at the end of the survey, and were designed as optional, i.e. did not force a response by not progressing the survey until a response is given. Therefore, any missing data were only for the demographic items. However, the demographic items were set to request a response with an error message indicating missing responses in case this was accidental, but the participants could still progress through the survey if they wished without responding to the demographic items.

Exploratory Factor Analysis. An exploratory factor analysis is a statistical method used in the early stages of scale development (Netemeyer et al., 2003; Tabachnick & Fidell, 2014; Yu & Richardson, 2015). It was chosen over principal component analysis because, unlike principal component analysis, factor analysis only analyzes shared variance, thereby yielding the same solution while avoiding the inflation of estimates of variance (DeCoster, 1998; Woldearegay & Woldearegay, 2013). The analysis was carried out to empirically examine the relationships between the items to identify item clusters that share enough variation and justify the presence of a factor to be measured by the scale, i.e. to determine the perception targets empirically. The analysis was done using the Statistical Package for the Social Sciences (SPSS) version 24.

Moreover, the analysis was used to reduce the 130 items to a smaller but defensible set of items that can best replicate the correlations observed and explain the most variance (Field et al., 2012). Prior to performing the factor analysis, soundness of the data for further analysis had to be established (Conway & Huffcutt, 2003; Woldearegay & Woldearegay, 2013).

First, measurement suitability for the EMI items was evaluated through an examination of the descriptive statistics. The mean for all responses and the standard deviation of each response were calculated. The Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) test was used to verify sampling adequacy for the exploratory factor analysis and Bartlett's Test of Sphericity was used to determine the size sufficiency for item correlations. The threshold for the KMO test was the recommended minimum value of .60, and for Bartlett's Test of Sphericity a statistical significance of less than .05 was required to conduct an exploratory factor analysis (Osborne & Costello, 2004, 2005).

To identify the optimal number of factors, expected number of factors, eigenvalues and scree plots were used to reach a decision. Factors that had an eigenvalue of greater than one were initially retained for further examination. The scree test involved examining the graph of the eigenvalues produced by SPSS, specifically, the natural bend or break point in the data where the curve flattens out (Tabachnick & Fidell, 2014). Moreover, the literature suggested that nine constructs should be expected, and this educated expectation was used to guide the interpretation of the factor patterns. The analysis was done with maximum likelihood extraction and an oblique promax rotation. Maximum likelihood extraction was used to test the statistical significance of factor loadings, calculate factor correlations and confidence intervals. In the social sciences, it is generally expected that some correlation exists among factors, because behavior in real-world settings is complex and cannot be divided into units that function independently of each another. Therefore, using oblique promax rotation should lead to a fairly accurate factor solution (Osborne & Costello, 2004, 2005).

This study used the following recommendations to help determine the EMI scale's structure; first, as Tabachnick and Fidell (2014) suggested, a .30 loading was used as a minimum loading for an item, which is equivalent to 10% overlap of variance with the other items within the relevant factor. Second, cross-loading patterns were examined; a *cross-loading* item was an item that loaded at .30 or higher on two or more factors. As it has traditionally been up to the researcher to decide whether a cross-loading item be dropped or not, an item was dropped only if it were conceptually redundant and there was at least one other conceptually important and statistically adequate (≥ .50) item loading on the respective factor. Third, no factors with less than three items were accepted as this could compromise their stability, and four or more sufficiently loading items (≥ .30) were desirable (Osborne & Costello, 2004, 2005).

Reliability. Given that this stage of the research entailed empirically developing the EMI scale, estimates of reliability for internal consistency were prioritized over others like test / re-test reliability. Internal consistency was calculated using Cronbach's alpha. An alpha co-efficient greater than .70 provided confidence to progress in further analysis (Biddix, 2009; Field et al., 2012; Switzer et al., 1999; Tabachnick & Fidell, 2014; Wetzel, 2012). The factor scores were calculated as the mean of the salient items belonging to each factor.

Multivariate Analysis of Variance. A MANOVA, conducted in SPSS version 24, explored the antecedents of performance management perception targets by examining response differences among participant groups. It was chosen over Analysis of Variance (ANOVA) to eliminate Type 1 error rate as well as to sufficiently account for the effects of independent variables on the latent factors. The independent variables (antecedents) used in this study included participant location, work role, industry sector, organization size, and employment status.

4. Study Three: Second Online Survey

4.1 *Study Purpose*

The aims of this study were to; (a) refine and validate the EMI scale, (b) verify the performance management perception targets identified in the previous study using new survey data, and (c) determine the statistically significant antecedents of the perception targets.

4.2 *Participants*

Target Population and Sample. Similar to the first survey, the target population included people who worked in roles with performance management. There was no restriction on the country or industry in which participants worked, but they were required to speak English and work in roles in which they were required to manage the performance of others. The calculation of the target sample size used a calculator for structural equation modelling. The calculator was based on three equations, shown next, to compute a priori sample sizes (Tabachnick & Fidell, 2014). Equation 1 shows the calculation of the error function.

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad \textbf{(1)}$$

Equation 2 calculates the lower bound sample size.

$$n = \max(n_1, n_2) \quad \textbf{(2)}$$

Where:

$$n_1 = \left\lceil 50\left(\frac{j}{k}\right)^2 - 450\left(\frac{j}{k}\right) + 1100 \right\rceil$$

$$n = \left\lceil \frac{1}{2H}\left(A\left(\frac{\pi}{6} - B + D\right) + H + \sqrt{A\left(\frac{\pi}{6} - B + D\right) + H)^2 + 4AH\left(\frac{\pi}{6} + \sqrt{A} + 2B - C - 2D\right)}\right\rceil$$

$$A = 1 - p^2$$

$$B = p \arcsin\left(\frac{p}{2}\right)$$

$$C = p \arcsin(p)$$

$$D = \frac{A}{\sqrt{3 - A}}$$

$$H = \left(\frac{\delta}{z_{1-\alpha/2} - z_{1-\beta}}\right)^2$$

Where:

j is the number of observed variables
k is the number of latent variables
ρ is the estimated Gini correlation for a bivariate normal random vector
δ is the anticipated effect size
α is the Sidak-corrected Type I error rate
β is the Type II error rate
z is a standard normal score.

Equation 3 calculates the upper bound sample size

$$F(x;\ \mu, \sigma^2) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{x - \mu}{\sigma\sqrt{2}}\right)\right] \quad \textbf{(3)}$$

Where:

μ is the mean
σ is the standard deviation
erf is the error function

The anticipated effect size was .10, desired statistical power level was .80, probability level was .05, and, based on the results of the first survey, the anticipated number of latent factors was eight and number of observed variables (items) was 55. The calculation resulted in a desired minimum sample of 370 participants.

4.3 *Data Collection*

The survey consisted of 55 items, requiring ten minutes to complete, measuring eight perception targets. It was created using the University of Melbourne's access to Qualtrics Research Core and its links were distributed through ProA and LinkedIn. Unlike the previous survey, all items, including demographic items were compulsory resulting in no missing data.

4.4 *Data Analysis*

Confirmatory Factor Analysis. A confirmatory factor analysis was used as it is a form of structural equation modeling used in theory testing and construct validation (Ullman, 2006). The analysis was used to improve the psychometric quality of the EMI scale and determine the factor model that best fits the data. Further, the analysis was conducted, to explicitly test a priori assumptions, from the first survey results, about the relationships between the items and the factors (Brown, 2015; Jackson et al., 2009). Specifically, it restricted item-factor loadings to zero on all factors except the one to which each set of items belong. Maximum likelihood estimation was used and several indices were used to determine model fit, i.e. how well the suggested model attained the covariance between the items. The analysis was conducted using Analysis of Moment Structures (AMOS) software version 24. The factor scores were calculated as the mean of the salient items belonging to each factor.

Model Fit. It was highly recommended to use a combination of fit indices to ascertain model fit. The analysis first used the Root Mean Square Error of Approximation (RMSEA), which is an estimation of a model's lack of fit compared to a perfect model (Jackson et al., 2009; Tabachnick & Fidell, 2014). The RMSEA determined how well model parameters reproduced population co-variances with unknown but ideally chosen parameter estimates (Thompson, 2004). A maximum RMSEA value of .08 was accepted to show good model fit, and since it favored parsimony and was not sensitive to large sample sizes, it was used as the main fit index to follow (Hooper et al., 2008). Second was the Standardized Root Mean Square Residual (SRMR), which is the standardized difference between predicted and observed correlation. An SRMR value of less than .08 was considered to indicate good model fit. Third, the Comparative Fit Index (CFI) which is a measure of model misspecification, for which, as a rule of thumb, any value greater than .90 indicates good model fit (Kline, 2011; Sivo et al., 2006). Furthermore, Chi-square Mean/Degree of Freedom (CMIN/DF) was used with a threshold of between 1 and 3 as acceptable (Bentler, 1990; Hooper et al., 2008).

Reliability. Estimates of reliability (internal consistency) of each factor were calculated using Cronbach's alpha, in SPSS 24. Given the scope of the thesis and that this stage of the research entailed validating a new scale, estimates of reliability for internal consistency were prioritized over others as it allows for a single administration of the scale by one researcher. An alpha co-efficient greater than .70 provided confidence to progress in further analysis.

Multivariate Analysis of Variance. For this study, MANOVA, done in SPSS 24, was used to ascertain the effect of *perceiver* and *situation* moderators on the verified performance management perception targets. Based on the recommendations from the first survey findings, the independent variables used in this study included gender, age, work role, number of years

with current employer (tenure), organizational size, industry sector, employment status, number of years in the workforce, and work location.

5. Ethics and Funding

As this research thesis involved participation of consenting adults in a non-sensitive research topic, it was categorized as a minimal-risk project under the regulations of the University of Melbourne Human Research Ethics Committee and received approval from the Melbourne Graduate School of Education Human Ethics Advisory Group (Ethics ID/HREC:1648539). The previously outlined studies were fully confidential and participants were invited to provide secure de-identified contact details where appropriate. This was a self-funded research project, and used resources, like software and desk space at the Melbourne Graduate School of Education. A supervisory team and an advisory committee guided the research at every stage.

6. Chapter Summary

This chapter presented the three research stages, summarized in Figure 11, namely; item generation, scale development and, scale refinement and validation. Three empirical studies were presented in terms of their purposes, data collection and analytical approaches. The next chapter discusses the results of the first empirical study and presents the item pool for the EMI scale tested in the first online survey.

Figure 11

Overview of Research Stages and Empirical Studies

	Stage One – Qualitative	**Stage Two – Quantitative**	**Stage Three – Quantitative**
Name	Item Generation	Scale Development	Scale Refinement & Validation
Aims	• Identify perception targets* • Generate items	• Determine perception targets • Refine items • Explore significant antecedents	• Verify perception targets • Refine items • Determine significant antecedents
Data Collection	Scoping literature review	Online survey 1	Online survey 2
Data Analysis	Deductive thematic analysis	Exploratory factor analysis Estimates of reliability Multivariate analysis of variance	Confirmatory factor analysis Estimates of reliability Multivariate analysis of variance
Software	Online databases MS Word MS Excel	Prolific Academic LinkedIn Qualtrics MS Excel SPSS 24	Prolific Academic LinkedIn Qualtrics MS Excel SPSS 24 AMOS 24

Note. * The purpose of stage one included identifying perception targets and generating items for them. However, it was through the narrative literature review that the perception targets were conceptualized, and through scoping study that the items were generated.

Chapter 4: Scoping Literature Review

This chapter provides a recap of the scoping literature review study purpose and presents the results of the search strategy. It follows with a discussion of the publications and the scales used to inform the item pool for the EMI scale. The chapter then provides a detailed description of the perception targets from Chapter 2 in terms of their operational definitions and respective items generated for the EMI scale. To conclude, the chapter presents a summary of the study's key findings and specific recommendations for the first online survey.

1. Purpose

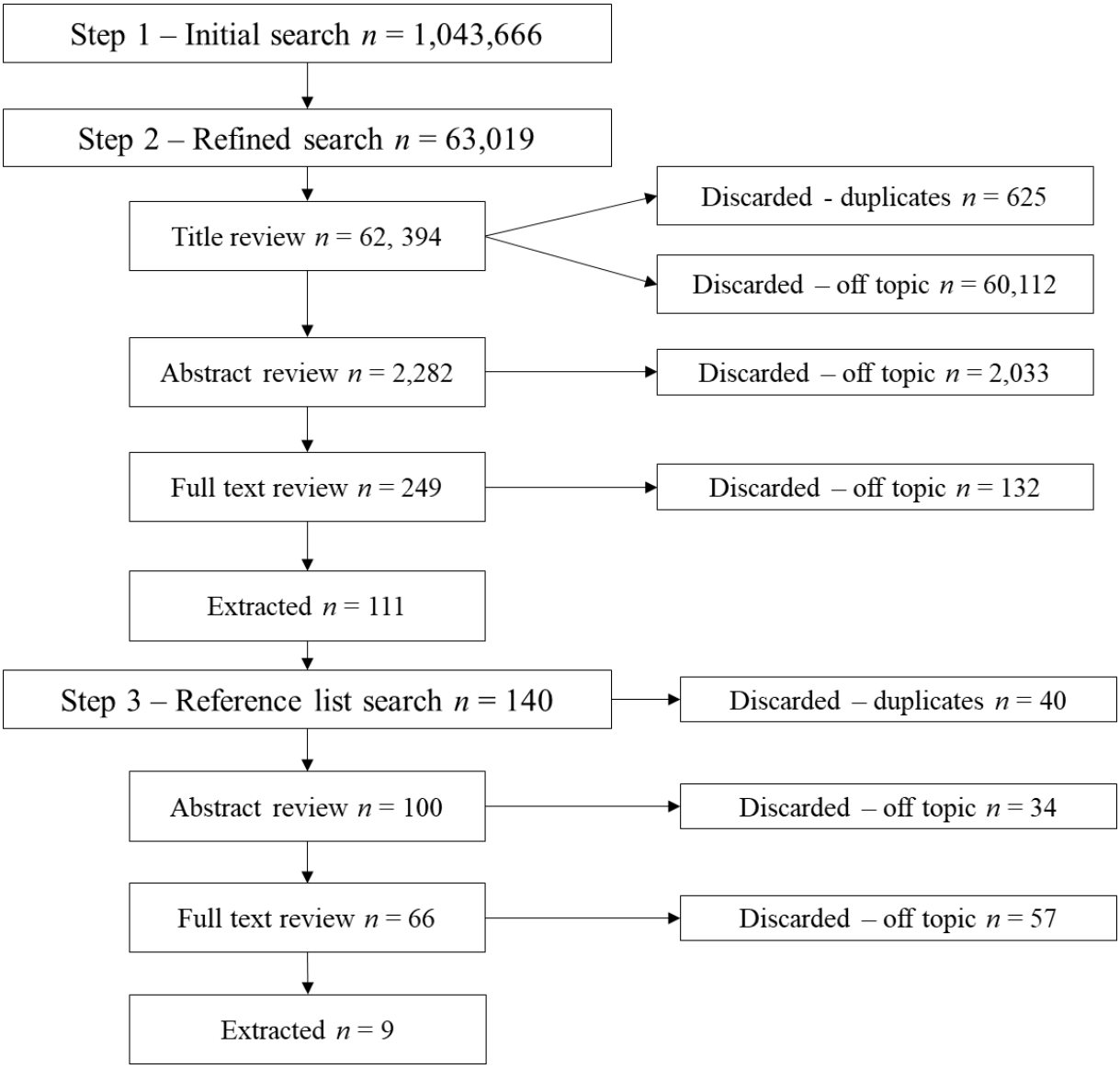
Based on the narrative literature review and advice from the advisory committee, it was decided that it would be prudent to learn from existing scales prior to creating the initial set of items for the EMI scale for the first survey. Therefore, the purpose of this scoping literature review study was to contribute to the generation of the items for the EMI scale by finding scale items relevant to the following performance management perception targets; evidence-based practice, organizational communication, organizational support (including supervisory support), work engagement and, others' behavior. In early 2017, an initial search was carried out to identify systematic reviews, scoping reviews and meta-analyses that may answer the study's question: 'What quantitatively validated scale items exist to measure five of the performance management perception targets identified in the literature review?' Since no relevant literature reviews were found, further confirming the need for this scoping study, the search strategy outlined in Chapter 3 Section 10 was implemented.

2. Search Strategy Results

The initial search resulted in more than a million records so, the search parameters were refined based on a text analysis of the relevant records' subject terms, titles, and abstracts to narrow the scope of the search to find a more manageable set of records. The second search found 63,019 publications for title review from which 625 duplicates and 60,112 off-topic publications were removed. After the abstract review, 2,033 publication were deemed to be off-topic and discarded. After the full text review, a further 132 publications were discarded, leaving 111 records for data extraction. The last search step was done using the reference list within each of the 111 publications resulting in 140 records for analysis. Forty publications were duplicates, leaving 100 for abstract review, based on which 66 publications were included in the full text review from which only nine were extracted. The entire process resulted in 120 publications for data extraction, shown in Appendix A. Figure 12 illustrates the search strategy steps and respective results.

Figure 12

Results of the Scoping Literature Review Search Strategy



Note. One iteration of the whole search strategy was done.

3. Results

3.1 *Publication Description and Characteristics*

As shown in Appendix A, the extracted 120 studies were from a variety of countries across the globe, several of which were relevant to more than one of the performance management perception targets. The most prominent perception target was HRM (47 studies), followed by organizational support (30 studies), organizational communication (26 studies), work engagement (25 studies), then supervisory support (six studies). No publications were found relevant to evidence-based performance management nor perception of others' behaviour, which presented a gap in the literature. The review resulted in many publications fitting the criteria to answer the review question. However, many were excluded from the item generation process as they were inferior by comparison to the others for several reasons. The following discusses those reasons using specific publications as examples.

Conceptual Rigor. Several studies used scales that were, arguably, conceptually flawed. For example, Malach-Pines (2005) used the shortened version of the Burnout Measure, the items for which can reasonably be related to non-work factors, such as employee sleeping habits or mood disorders. As such, their relationship to work engagement was confounded by factors outside the scope of the EMI scale and the thesis overall.

Scale Clarity and Item Rigor. Some of the studies were not clear about the categories used in their Likert scales, and while one can usually make an educated assumption given at least a numerical code, it would have been negligible to do so with, for example, the paper by Postmes et al. (2001). The authors indicated a five point scale on items that measure the quantity of communication, but without the text categories for each point, it was unclear if 1= rarely, or 1= never, which is conceptually confusing and potentially misleading.

Other papers used scales that were, possibly, far too open for interpretation. For example, a popular measure for engagement from Rich et al. (2010), based on Kahn (1990) and Rothbard (2001), was categorized to reflect the physical, emotional and cognitive effort an employee invests in their work. Aside from the items relating to emotional effort, the items can be related to negative concepts such as stress or workaholism, for example "At work, I am absorbed by my job". Moreover, the authors suggested that in an engaged employee the three effort categories would correlate simultaneously, however, their paper did not offer guidance on interpreting this correlation. Therefore, it was unclear if all three needed to be correlated for an employee to be deemed engaged or if at least two would suffice. Furthermore, a threshold for correlation was not provided. Likewise, other studies like Boon et al. (2011), used items that were too vague like, "The organisation offers (me) work that gives me the opportunity to express myself". It is not clear what self-expression means in this context. Instead, the items could have been clearer had it provided context for self-expression like expressing creativity in problem solving. Also, the item "The organisation offers (me) possibilities to present my opinion on matters" needs to clarify the 'matters' in question.

By comparison, Though the Utrecht Work Engagement Scale (UWES), created by Schaufeli et al. (2002), contains items that can be indicative of negative constructs, like, "To me, my work is challenging", which could be related to a state of feeling overwhelmed rather than dedication, the authors' work has certain advantages. The saving grace of Schaufeli et al. (2002) and Schaufeli and Bakker (2010) is in their clarification of the purpose of their items and inquiry lens. For instance, the authors stated that it would be inappropriate to measure engagement exclusively in terms of observed voluntary extra effort because both stressed and engaged employees may display similar behavior, like working longer hours (Schaufeli & Bakker, 2010).

Such clarity in the interpretation and implications of their items made their work far more helpful for the EMI scale items than others. Moreover, other items in the UWES, which has been widely used and translated to 21 languages, were very helpful for the EMI scale. For example, the item "When I get up in the morning, I feel like going to work" was especially helpful as it indicate a state of vigor and offers a good balance of perceptual and observable qualities leaving little room for misinterpretation (Balducci et al., 2010; Petrović et al., 2017; Xanthopoulou et al., 2012).

Some studies used items that were too restrictive in a conceptual sense. For example, Sun et al. (2007) used a measure of high performance HR practices in customer service, which included items like "employees in customer contact jobs will normally go through training programs every few years". The main issue lies in the fact that such items do not reflect the characteristics of effective HRM, like responsiveness to individual employee development needs. Similarly, the Communication Satisfaction Questionnaire by Downs and Hazen (1977) contains items limited in terms of observability. For example, respondents are asked to rate their level of satisfaction with the amount and/or quality of information related to their job, like "information about my progress in my job" (Clampitt & Downs, 2004). Such items offer only perceptual data and do not indicate whether the respondent is satisfied with the quality or quantity of the information. However, the Downs-Hazen questionnaire also gives items that are more balanced on the perceptual to observable spectrum, like item 26 which asks the respondent to rate their satisfaction with the "extent to which I receive in time the information needed to do my job".

Analysis Suitability. Castanheira and Story (2016) conducted an exploratory factor analysis for HRM perceptions but their sample size (117 respondents from one organization) was far too small and their factor analysis was not truly exploratory (Boateng et al., 2018; Bujang et al., 2012). To elaborate, the authors stated that, since previous studies had demonstrated HRM as a unidimensional construct, they proceeded with their factor analysis by imposing a single factor solution. This approach is more suitable for a confirmatory factor analysis than an exploratory one (Brown, 2015; DeVellis, 2017). Moreover, to determine mediation effects, Castanheira and Story (2016) measured the relationship between HRM perceptions and affective commitment, however, given that their study was observational and cross-sectional, their strong claims for mediation effects are limited. For conclusive mediation effects one needs to meet three requirements of causality; correlation, removal of confounding factors and temporal precedence (Brown, 2015; de Brito & de Oliveira, 2016; Walker, 2005). However, observational cross-sectional studies can only demonstrate correlation (DeVellis, 2017; Mercer et al., 2007; Tabachnick & Fidell, 2014).

Some publications omitted important information, perhaps due to the word limits imposed by their respective journals. For example, Bednall et al. (2013) included a table showing the standardized factor loadings for their measurement model but they did not show the items for each of the HRM strength subscales they used (distinctiveness, consensus and, consistency), nor how the items loaded onto the HRM strength factor, instead they displayed how each subscale loaded onto the HRM Strength factor.

Ang et al. (2013) also had a small sample for factor analysis (193 employees linked to 58 managers in one regional Victorian hospital in Australia). Moreover, their publication presented the means, standard deviations and coefficient alpha ranges, but not the factor loadings. Otherwise, their items were helpful from a conceptual perspective, though they were very similar to the Joint Committee's list of key questions (2008). The latter was preferred for this thesis because even though the key questions had not been tested psychometrically, they were developed by the top experts in evaluation and have been applied in evaluation projects and educational organizations for decades and were based on operational definitions of each standard. Thus, they presented an opportunity to narrow the gap in evaluation research and in practice. Therefore, the scales used significantly for this thesis from the 120 studies were; The UWES, The Downs-Hazen Communication Satisfaction Questionnaire, and the Survey of Perceived Organizational Support. A discussion of these scales and their validity is provided next.

3.2 *Relevant Scales*

Organizational Communication. Many measurement tools developed in the 1970s are currently used to assess organizational communication, like the Communication Satisfaction Questionnaire (Downs & Hazen, 1977). Developed by Downs and Hazen in 1977, the Communication Satisfaction Questionnaire is still one of the most comprehensive tools as it assesses the direction of communication flow, the forms of communication and the relationships between staff (Gray & Laidlaw, 2004). It is also relatively easy to administer and understand; 34 items on a seven point Likert scale, one open text item and the remainder in a multiple choice format.

After much psychometric testing, the developers isolated 44 items within an eight factor solution; communication climate, media quality, organizational integration, supervisory relationship, organizational perspective, personal feedback and relationship with subordinates (Clampitt & Downs, 2004). More recently, the Communication Satisfaction Questionnaire has shown a test-retest reliability of .94 (Zwijze-Koning & de Jong, 2007). Cronbach's alpha calculations showed good estimates of reliability for an adapted version of the Communication Satisfaction Questionnaire (horizontal communication; .77, relationship with subordinates; .91, communication climate; .84, relationship with superiors; .93, personal feedback; .83, organizational integration; .79, media quality; .90, corporate perspective; .82) (Meintjes & Steyn, 2006). The eight factor solution was confirmed by Crino and White (1981), but disputed by Gray and Laidlaw (2004) who suggested a two-factor solution; relational communication and informational communication.

Organizational Support. Eisenberger et al. (1986) developed a 36-item survey instrument to measure perceived organizational support using a seven point Likert scale (1 = strongly disagree, 7 = strongly agree) called the Survey of Perceived Organizational Support. The authors ran the survey with employees across nine diverse organizations. Similar subsequent studies used the survey in an assortment of organizational settings (Erdogan et al., 2015; Gyekye & Salminen, 2009; Joo et al., 2015; Kottke & Sharafinski, 1988; Maden & Kabasakal, 2014; Neves & Eisenberger, 2012; Ngo et al., 2014). The survey is based on the hypothesis that an employee's perception of organizational support will grow as they perceive their performance will be proportionally repaid with rewards from the company in not only task-related benefits, like role advancement or increased autonomy, but also in the form of investment in their wellbeing (Eisenberger et al., 1986; Worley, 2006). The factor analysis of the 36-item scale

produced one main factor, which accounted for 48% of the total variance. A reliability analysis of internal consistency using Cronbach's alpha, resulted in an alpha of .97 with item-total correlations ranging from .42 to .83 (Eisenberger et al., 1986). Two years later, the Survey of Perceived Supervisory Support was created by Kottke and Sharafinski (1988) heavily based on the Survey of Perceived Organizational Support.

Kottke and Sharafinski (1988) administered both surveys to 216 employees and found that the items from the Survey of Perceived Organizational Support loaded on one factor (63.8% of variance). Likewise, items from their Survey of Perceived Supervisory Support loaded on one factor (74.1% of variance). Their reliability analysis nearly duplicated that of Eisenberger et al. (1986) showing a coefficient alpha for the original survey of .96, and for their survey of .98. Item-total correlations for the Survey of Perceived Organizational Support ranged from .48 to .88, while those for the Survey of Perceived Supervisory Support the ranged from .58 to .92 (Kottke & Sharafinski, 1988).

The Survey of Perceived Organizational Support contains many items that are imbalanced in terms their perceptual versus objective wording, but they do contain important concepts. For example, when examining the item "the organization is unconcerned about paying me what I deserve", one can see that the word *unconcerned* is hard to quantify, however the concept of fair reward is clear. Moreover, since the survey aims to assess perception specifically, overly objective items may not be appropriate. For instance, statements about the frequency with which an organization advertises an opportunity for promotion do not indicate how staff perceive the fairness of this opportunity. Instead it would be far more telling to ask staff if they feel that career opportunities at their organization are influenced by organizational politics or favoritism,

because the mere perception of injustice can lead to disengagement and / or withdrawal behavior (Chan & Lai, 2017; Hung et al., 2009; Lavelle et al., 2009).

Work Engagement. Various instruments of assessing work engagement were developed with varying levels of validity (Schaufeli & Bakker, 2010). Most studies, to date, have looked at differences between employees with regards to work engagement and the possible causes of these differences using the UWES (Bakker, 2011; Schaufeli & Bakker, 2010). This scale is based on the work engagement components of vigor, dedication, and absorption. The scale items are scored on a seven-point frequency scale extending from "never" to "always", and the scale has been validated in a variety of ways such as tests for internal consistency, and discriminant validity (Bakker, 2011; Schaufeli & Bakker, 2010).

The two versions of the UWES (UWES 17 and UWES 9), have been extensively adapted and tested (Balducci et al., 2010; de Bruin & Henn, 2013; Tomás et al., 2018; Xanthopoulou et al., 2012). For this thesis, the UWES 17 was used to allow for the examination of a larger item pool. Theoretically, six items measure absorption, five measure dedication and the remaining six items measure vigor. All three were found to be somewhat strongly correlated (mean $r > .60$) however, Schaufeli et al. (2002) concluded that the three constructs did not load on one higher order factor for engagement, and that a three-factor solution yielded best fit. Nevertheless, it is essential to note that the scales for vigor and absorption were highly correlated ($r > .90$ for the latent variables; and $r > .70$ for the observed variables). This high correlation should be expected at least from a psychological point of view, because a person who is happily absorbed by their work can be enthusiastic and energetic (Schaufeli et al., 2002).

Perception of Others' Behavior. Measurement of in-role behavior can be different both between and within organizations, as it is dependent on each position description. Measurement of extra-role behavior tends to be less concrete and requires metrics of a more qualitative nature, as well as the merger of self-rating, supervisor rating, and co-worker rating. Most importantly, as Rich et al. (2010) noted, the categorization of performance can be tricky given that what constitutes extra-role behavior in one organization can be considered as in-role behavior in another. Measures of withdrawal behavior are, arguably, more direct than extra-role behavior. Measurement of unplanned lateness and absenteeism includes frequency counts and duration (Smith et al., 2016; Youngcourt et al., 2007). Such data can be collected systematically using dashboard systems fed by automatic key card logs or manual notation by the supervisor. Turnover can be calculated automatically as the process of exiting a role is recorded in an HRM system. Moreover, several studies have used intention to quit as a proxy for voluntary turnover in several studies (Boswell et al., 2004).

Items from the aforementioned scales were added the item pool for the EMI scale in addition to newly created items, and the key questions set by the Joint Committee for the evaluation standards. The next section provides a discussion of the items for the EMI scale.

4. Items for the EMI Scale

The intended purpose of the EMI scale is the measurement of performance management perception targets, held by performance managers, to help organizations evaluate their performance management. That is, to inform the organizations of priority areas for which further inquiry is needed. As such, in line with utilization-focused evaluation, discussed in Chapter 2 Section 4.2, the items needed to be written with the intended users in mind. Namely, staff who design, implement or use performance management practices like HR executives, HR staff and

team-leaders / supervisors. The wording of some of the items found in this study was changed to be more observable, and, given the potential for survey fatigue, item selection had to be prioritized and restricted. It was also important to have a balance of items in terms of their orientation (i.e. descriptive, evaluative, observational, or experiential).

Descriptive items measures aim to capture the observed reality of practices or interactions in as quantitative a manner as possible. Therefore, they usually require the respondents to recall frequency of an event, locations or number of people (Edgar & Geare, 2005; Kehoe & Wright, 2013; Wang et al., 2020). They are, hence, inextricably linked to memory which can lead to recall bias, however, they are less prone to subjectivity from mood fluctuations like evaluative items.

Evaluative items require the rater to provide their own impression or judgement of the subject matter. For example, items may require the rater to provide their opinion on the relevance of their supervisor's feedback rather than the frequency of interaction with their supervisor (Alfes, Shantz, et al., 2013; Jensen et al., 2013). Although evaluative items are prone to subjectivity from cultural influences, experiences and mood fluctuations, they are suited for assessments of perception, which is subjective by nature.

Observational items require raters to place themselves as the third person, and provide information purely on what they have seen or heard happen to or by others (Wang et al., 2020). For example, items could ask raters about the communication channels available to staff or the level of gossip in the workplace. Such items can be too presumptuous of the reality of a situation or others' behaviour, but they are suited for understanding the rater's perceived environment.

By contrast, experiential items require the rater to provide information solely based on what they have experienced firsthand (Conway et al., 2016; Wang et al., 2020). For example, items about workplace communication could ask about the rater's preferred communication medium, or how often they engage in workplace gossip. Such items are subject to social desirability bias especially if the subject matter is socially sensitive. Though that bias may never be fully eliminated, it can be mitigated by rater anonymity and careful item wording.

Many studies have connected perceptions of HRM content to worker outcomes like turnover intentions (Kuvaas & Dysvik, 2009), early retirement (Herrbach et al., 2009), organizational commitment (Conway & Monks, 2009; Edgar & Geare, 2005; Macky & Boxall, 2007), performance (Liao et al., 2009), organizational citizenship behavior (Alfes, Shantz, et al., 2013) and, job satisfaction (Macky & Boxall, 2007). Moreover, the effect of some moderators like characteristics of the perceiver or the situation (organization), on the connection between HRM content perceptions and worker outcomes was explored (Li et al., 2017; Liu et al., 2017). Other studies considered such characteristics as control or independent variables (Bednall et al., 2013; Biswas et al., 2013; Boon et al., 2011; Harris et al., 2013). In this thesis, characteristics of the perceiver and the situation were treated as antecedent variables.

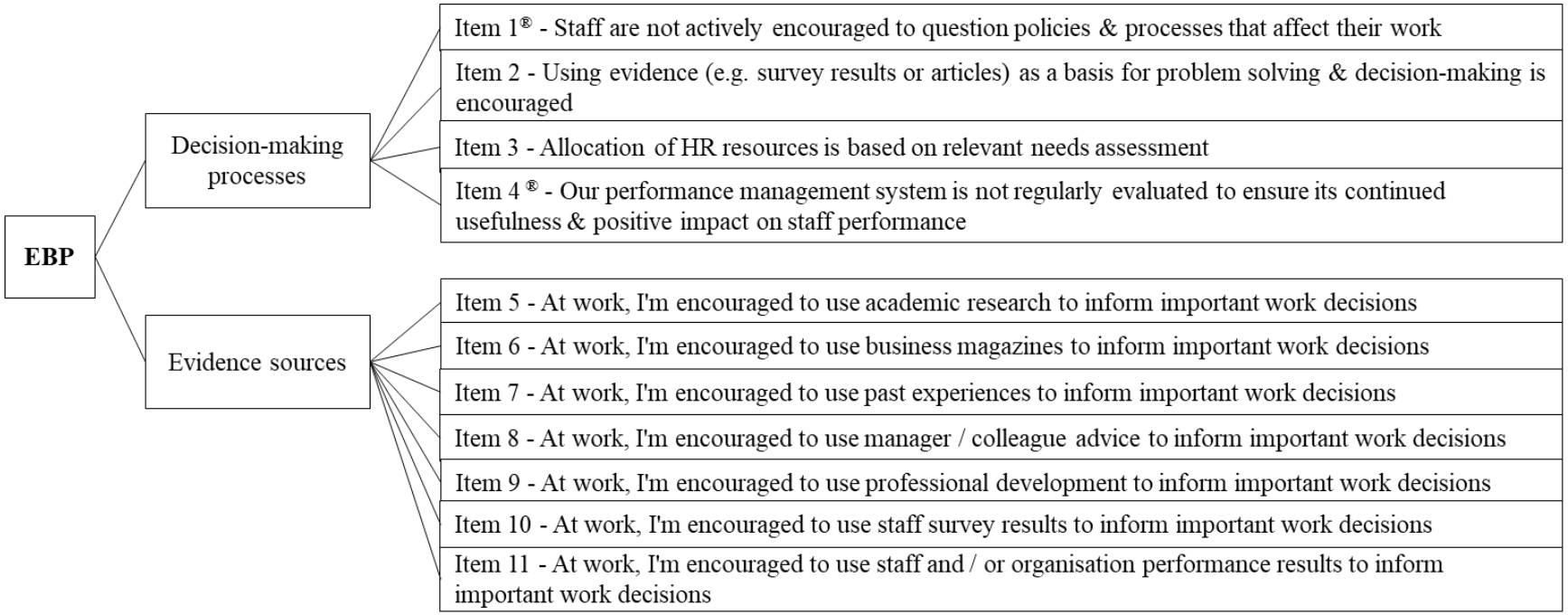
As previously stated, items for five of the performance management perception targets were generated based on their operational definitions. Developing these definitions required some of the steps recommended by Gable and Wolf (1993), previously discussed in Chapter 3. The first perception target, called evidence-based practice had two sub-constructs, shown in Figure 13. The items for this target were heavily guided by the work of Rousseau (2012) and Barends et al. (2014) who stated that management decisions must be based on evidence gathered from a variety of reliable sources. Namely from scientific research, experiential evidence,

organizational evidence, and stakeholder evidence like data collected from staff affected by performance management decisions.

The first sub-construct was called decisions (four items) with a focus on *processes*; collectively, they referred to the evidence basis of important decision-making processes in performance management like resource allocation, problem solving, and evaluation. The second sub-construct was called evidence sources (seven items) and its focal point was *variety*. Essentially, this sub-construct stressed the importance of having a variety of sources of reliable evidence, as recommended by Rousseau (2012) and Barends et al. (2014). The resultant operational definition stated that evidence-based practices are those in which performance management decisions are informed by reliable evidence from multiple sources and open to continuous improvement.

Figure 13

Items for Perception of Evidence-Based Practice

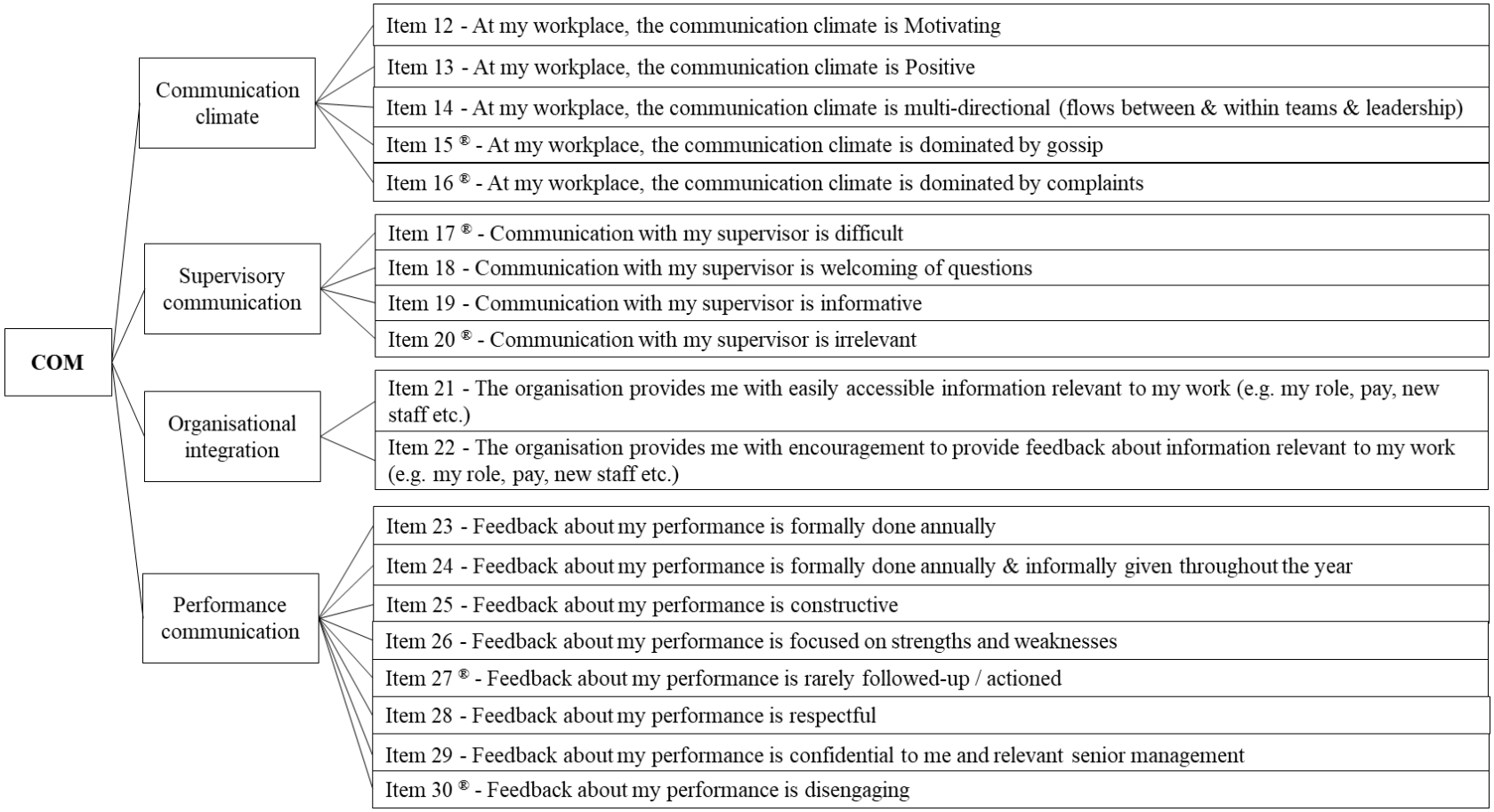


Note. Reverse coding is denoted ®. EBP = evidence-based practice.

The second target, called perception of organizational communication had four sub-constructs, shown in Figure 14. The first, called communication climate (six items) focused on the effect of different types of communication on the work environment. For example, communication which is engaging or constructive, a workplace dominated by gossip or complaints, and having communication channels which allow for multi-directional communication between and within organizational departments. The second sub-construct called supervisory communication (four items) specifically focused on the nature of information from and to one's supervisor, for example having a supervisor who is approachable, as well getting information from one's supervisor, which is relevant to one's work. The third sub-construct was called organizational integration (two items) and it focused on the accessibility and nature of information to and from one's organization. The fourth sub-construct was called performance communication (eight items) as it specifically focused on the information about one's performance and its measurement. For example, the formality of procedures, the accuracy of measurement, and utility of the information to improve one's future performance. The resultant operational definition stated that perception of organizational communication refers to perceptions of the content and delivery of communication at work, both in terms of information specific to performance and the daily communication within the proximal work environment.

Figure 14

Items for Perception of Organizational Communication

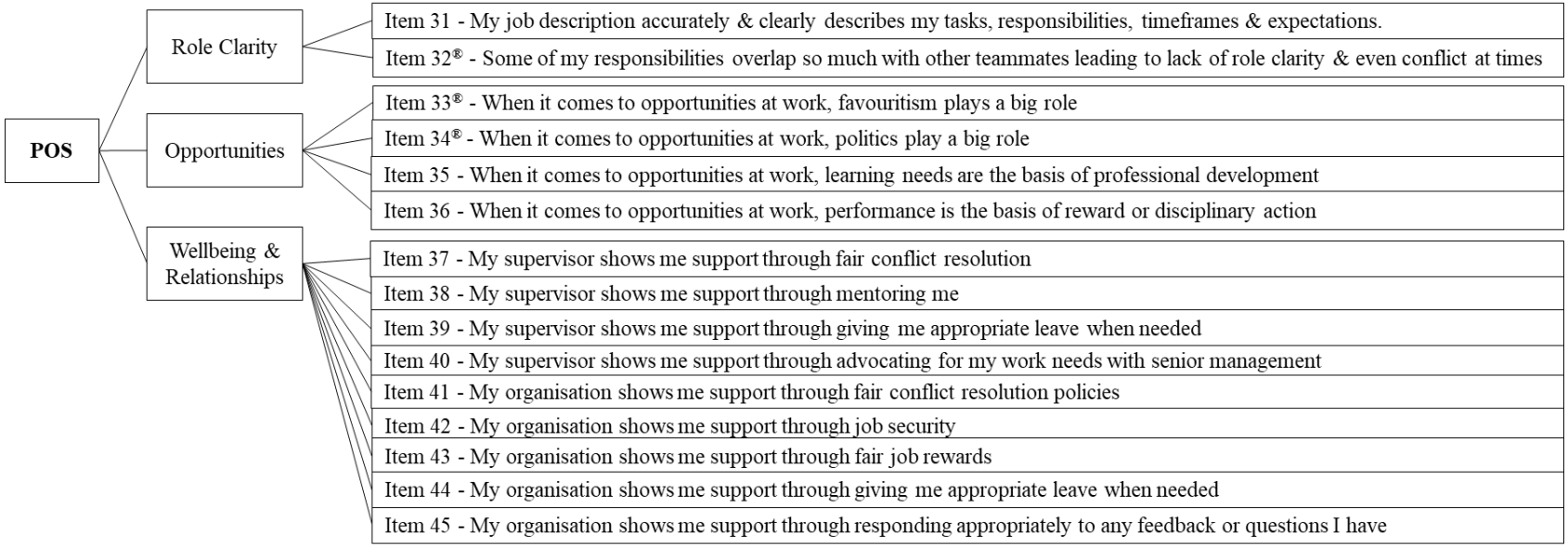


Note. Reverse coding is denoted ®. COM = perception of organizational communication.

The third target, called perception of organizational support had three sub-constructs, shown in Figure 15. The first, called role clarity (two items) as it focused on the specific tasks to be performed in terms of the accuracy of one's job descriptions and the distinctiveness of one's responsibilities. The second, called opportunities (four items) as it targeted staff development opportunities in terms of their legitimacy and fairness or, conversely, the existence of favoritism. The third sub-construct, called wellbeing and relationships (nine items), targeted the investments made by the organization and the supervisors to the staff wellbeing. For example, fair conflict resolution, mentorship, fair leave when needed, and job security. The resultant operational definition for perception of organizational support was that it is the perception of the degree to which an organization and its agents support staff career development and welfare.

Figure 15

Items for Perception of Organizational Support

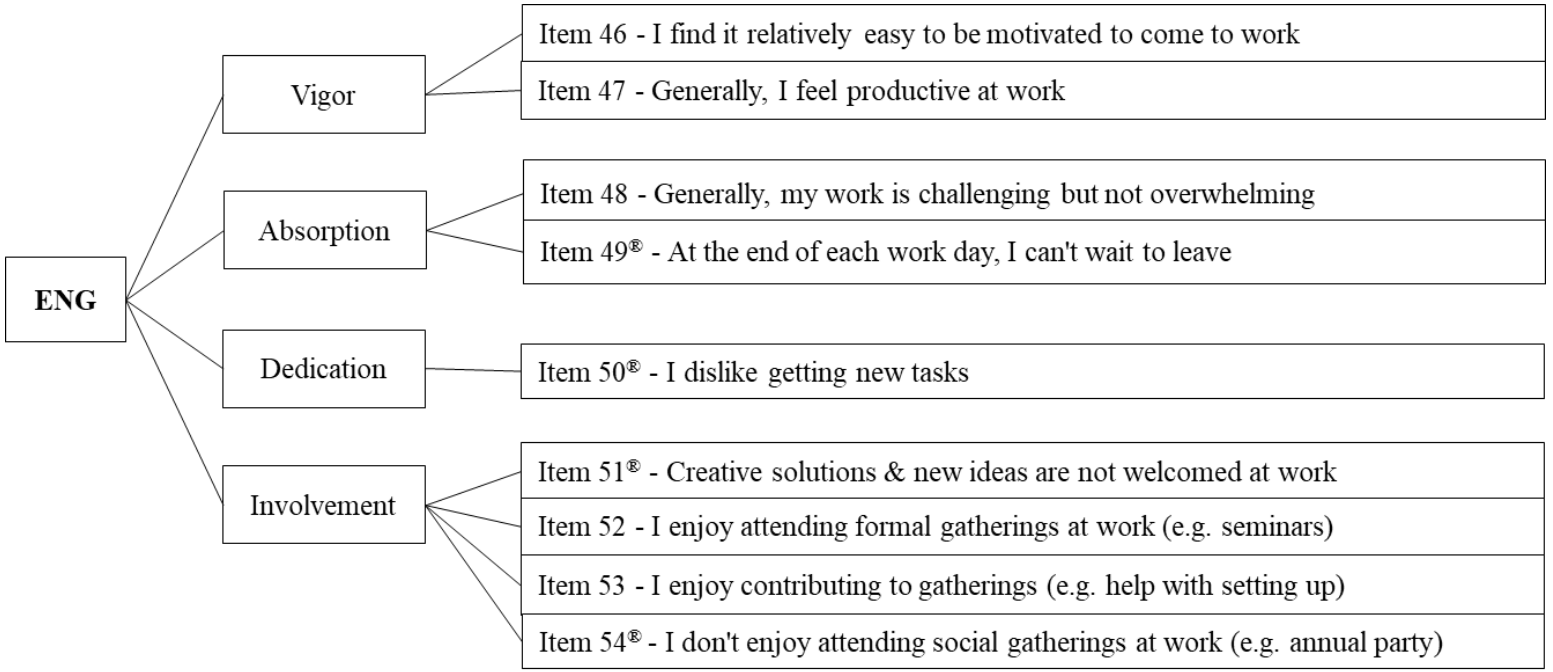


Note. Reverse coding is denoted ®. POS = perception of organizational support

The fourth target, called work engagement had four sub-constructs, shown in Figure 16. The first, called involvement (three items), referred to the voluntary participation of staff in organizational activities in terms of the enjoyment one feels in attending or contributing to formal and informal activities. The second sub-construct, called vigor (two items), referred to one's energy levels at work, for example feeling motivated to come to work or feeling productive at work. The third sub-construct, called absorption (two items), referred to the levels of one's ability to focus on their tasks, for example, having a sense of time flying by while working on interesting tasks. The fourth sub-construct, called dedication (one item), referred to one's levels of enthusiasm with their work tasks, for example a state of looking forward to getting new tasks at work. The resultant operational definition of work engagement stated that it referred to one's keenness to contribute to and participate in organizational events, as well as their energy and motivation levels in task performance.

Figure 16

Items for Perception of Work Engagement

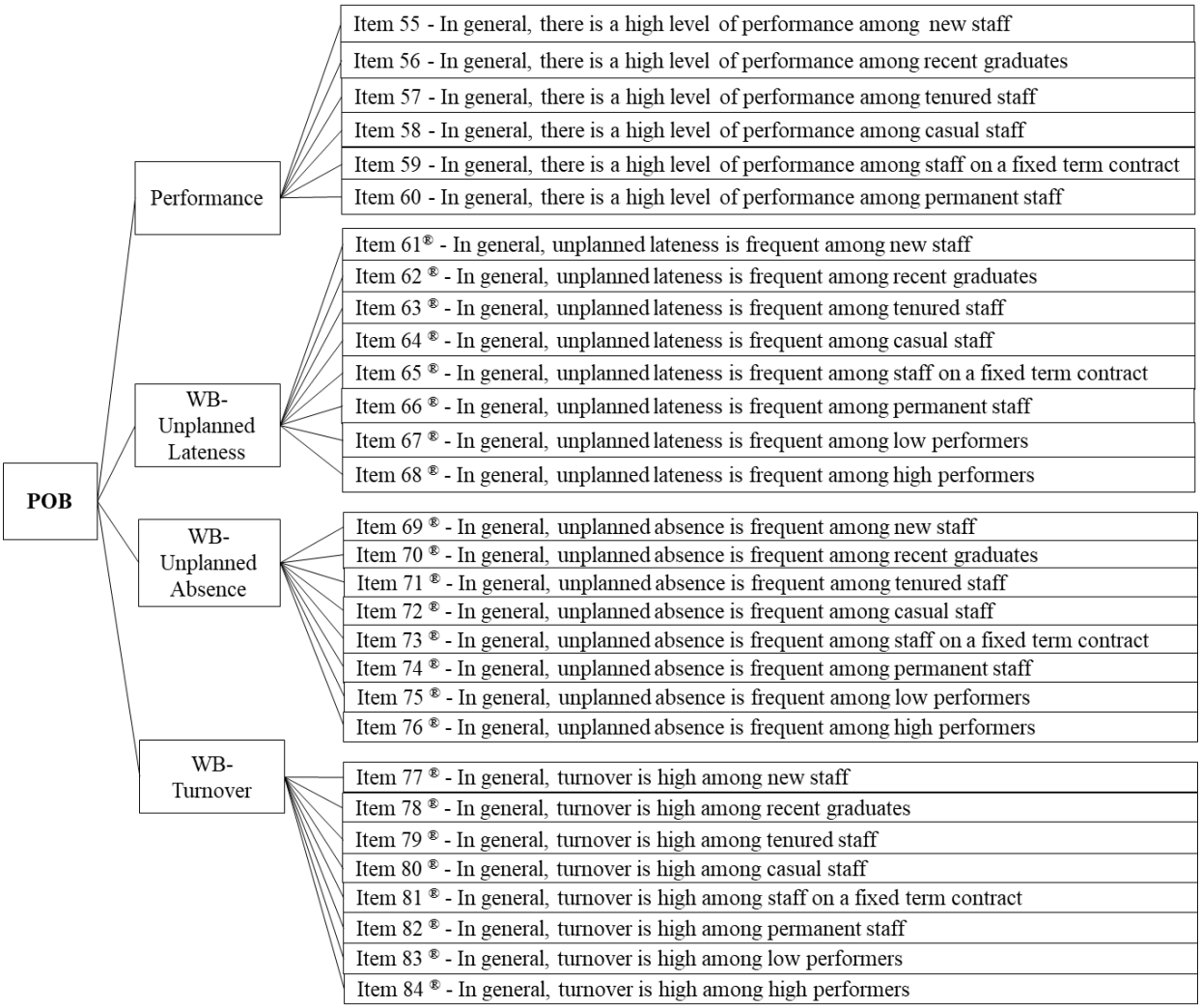


Note. Reverse coding is denoted ®. ENG = work engagement.

The fifth target, called perception of others' behavior, had two sub-constructs, shown in Figure 17. The first was called performance (six items) referred to the rates of high performance of a variety of staff like staff new to the organization, recent graduates, tenured staff, and those on a casual, permanent, or fixed-term contract. The second sub-construct was called withdrawal which had three foci; unplanned lateness (eight items), unplanned absence (eight items) and turnover (eight items). All three referred to behavior among staff who are usually high performers or those who are usually low performers, staff new to the organization, recent graduates, tenured staff, and those on a casual, permanent, or fixed-term contract. The resultant operational definition for perception of others' behavior, stated that it refers to one's perception of their team or colleagues in terms of their behavior at work.

Figure 17

Items for Perception of Others' Behavior



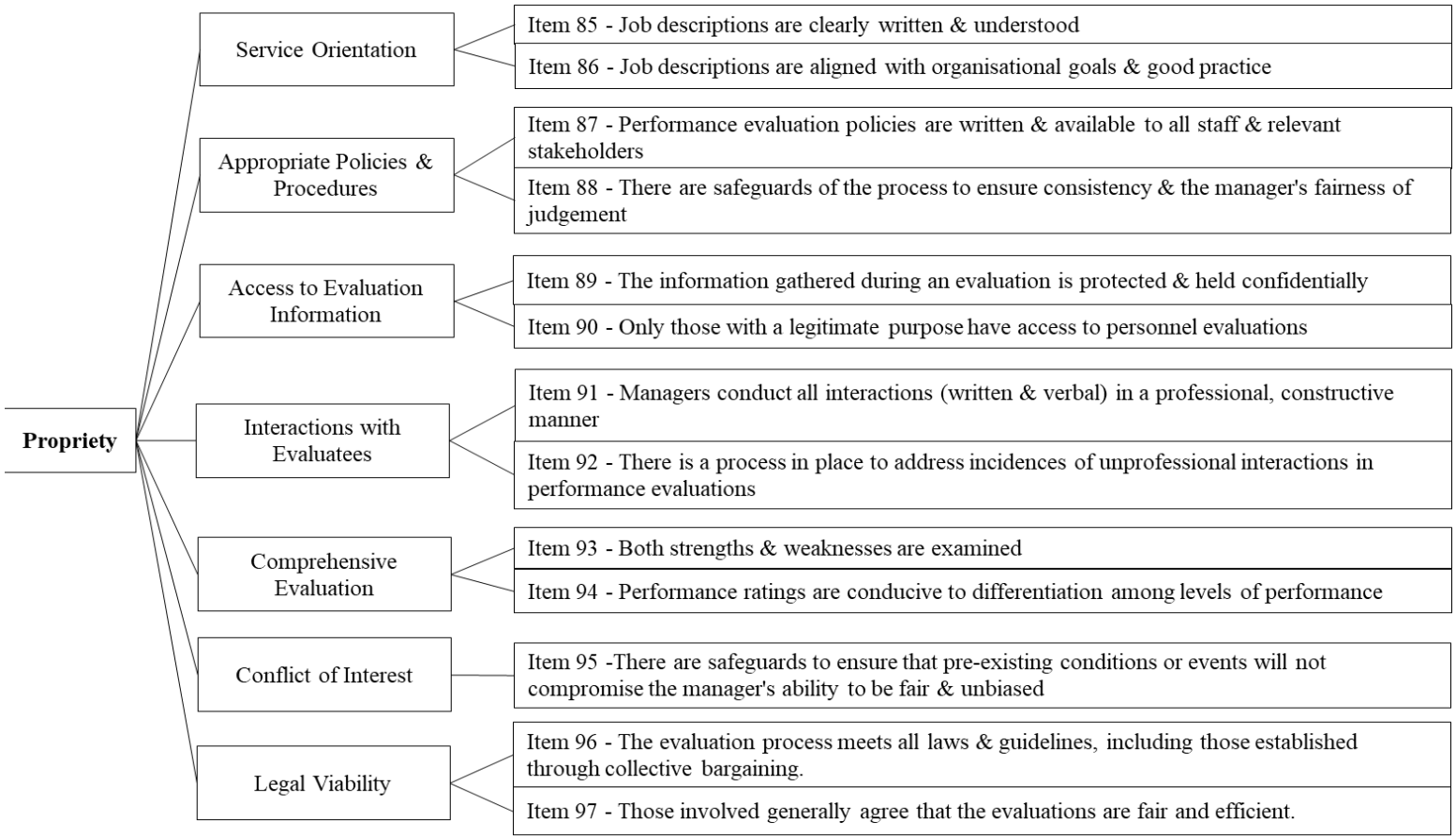
Note. Reverse coding is denoted ®. POB = perception of others' behavior. Withdrawal behavior (WB) was divided into its three focal areas to show the specific origins of the items.

Items for the four evaluation standards were adopted from the key question list provided by the Joint Committee (2008). Forty-six items were taken from the Joint Committee standards, none of which were reverse coded to keep item integrity as much as possible as this was the first time their psychometric properties were to be tested. However, to reduce response fatigue some terms were re-worded by their synonyms as recommended by the feedback from the initial survey test with colleagues and the advisory committee.

Thirteen items belonged to the seven propriety standards, shown in Figure 18, eleven items belonged to the six utility standards, shown in Figure 19, four items belonged to the three feasibility standards, shown in Figure 20, and the final 18 items belonged to the eleven accuracy standards shown in Figure 21.

Figure 18

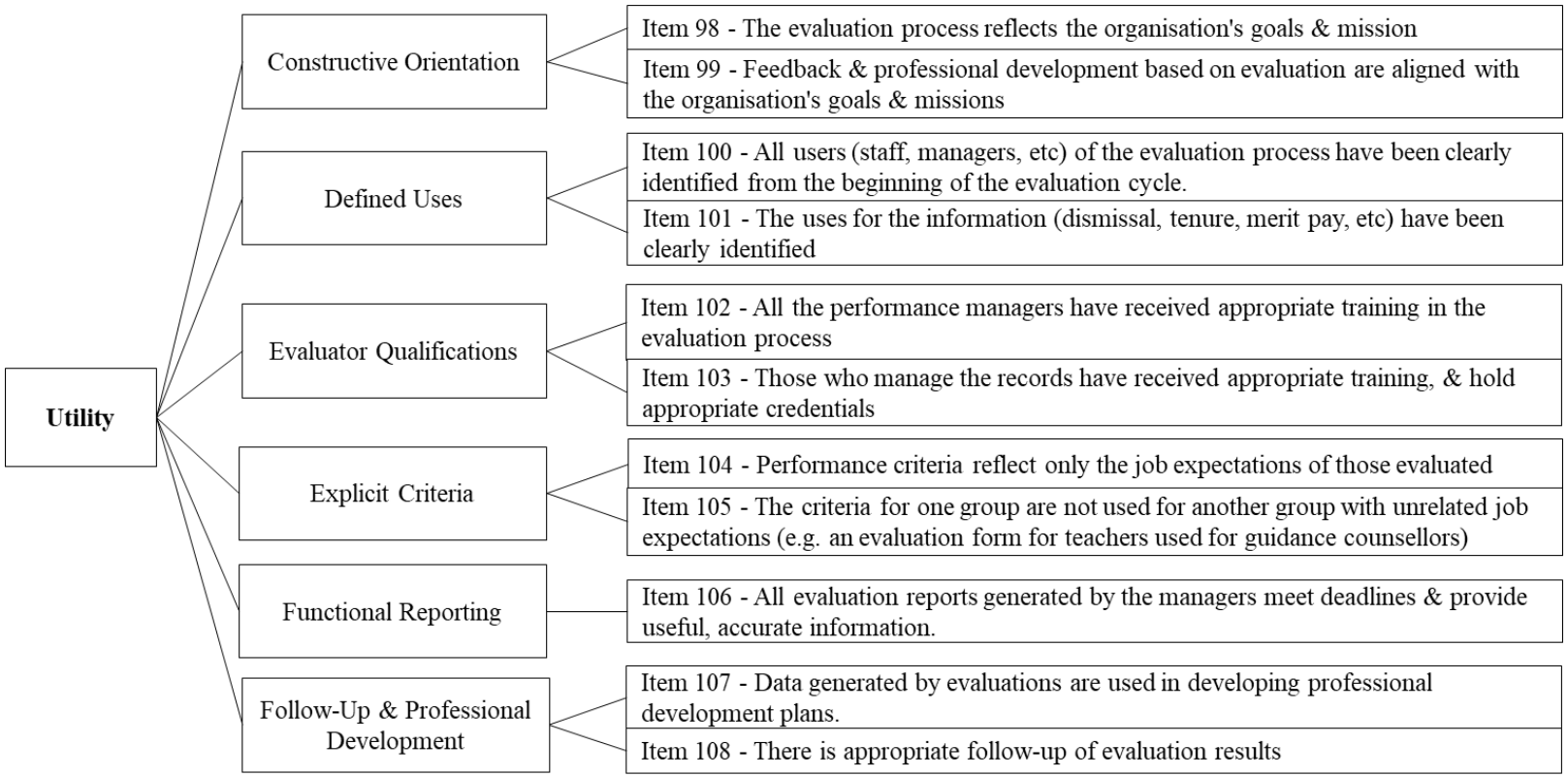
Items for Perception of Propriety



Note. Perception of propriety, the targeted standards and related items. Adapted from the Joint Committee (2008) list of key questions.

Figure 19

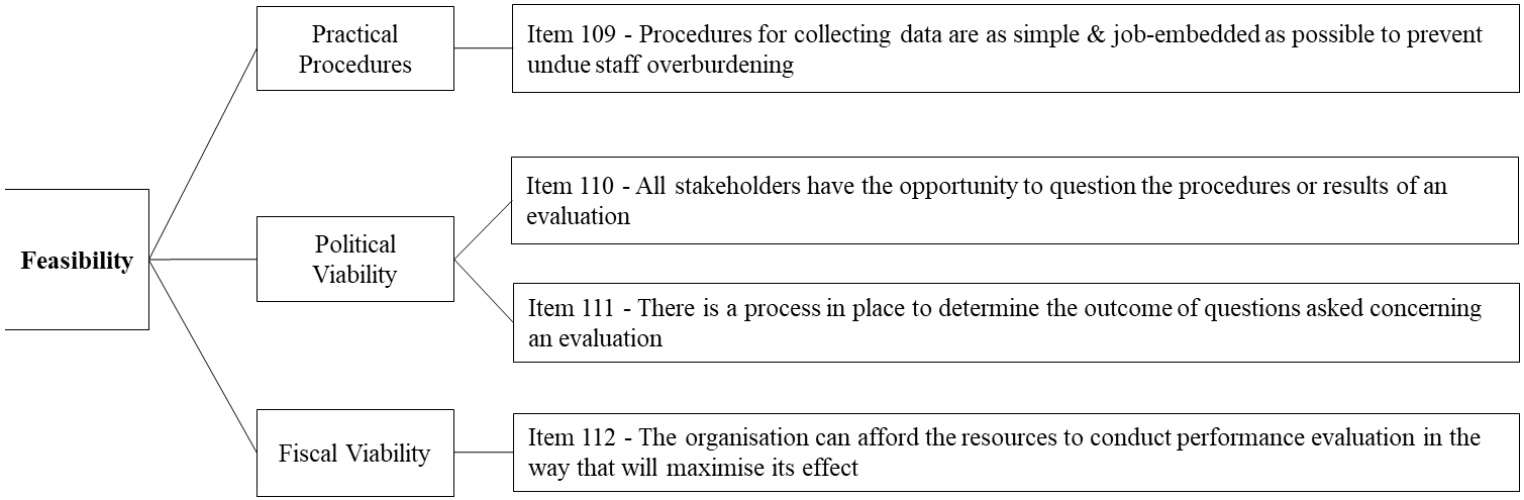
Items for Perception of Utility



Note. Perception of utility, the targeted standards and related items, adapted from the Joint Committee (2008) list of key questions.

Figure 20

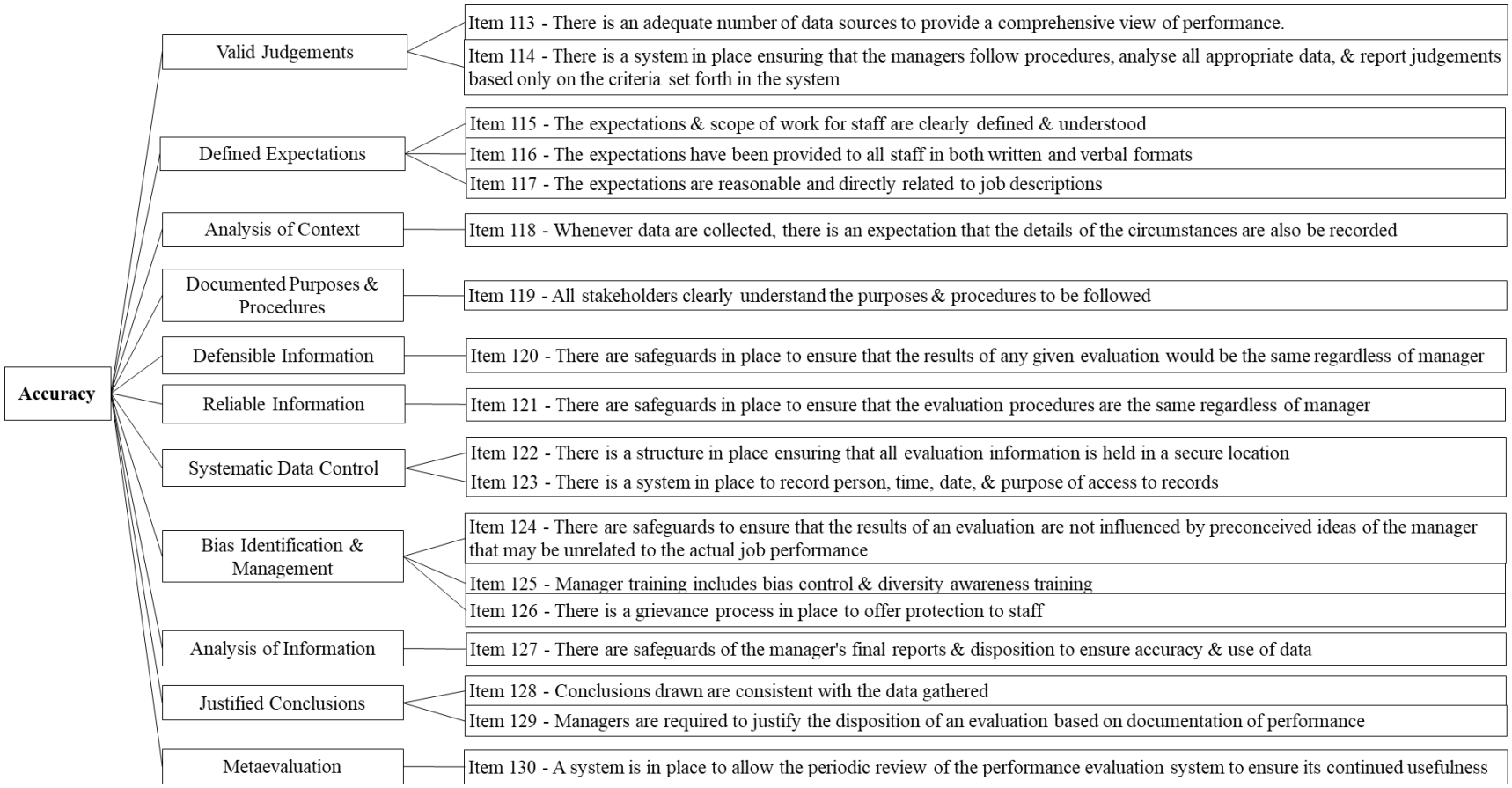
Items for Perception of Feasibility



Note. Perception of feasibility, the targeted standards and related items, adapted from the Joint Committee (2008) list of key questions.

Figure 21

Items for Perception of Accuracy



Note. Perception of accuracy, the targeted standards and related items, adapted from the Joint Committee (2008) list of key questions.

5. Chapter Summary

This chapter outlined the results of the scoping literature review and discussed the scales from which items were used for the EMI scale. Items for perception of propriety, accuracy, feasibility and utility were adopted from the Joint Committee's key questions list (2008). Items for perception of work engagement, organizational communication and organizational support were created based on several existing scale items found in this study as well as the operational definitions created for each of those perception targets. The items were designed to measure self-perceptions rather than perceptions of others' work engagement, satisfaction with organizational communication or organizational support. Whereas items for perception of others' behavior require raters to reflect on other staff. This is because behavior is more observable than perception. For example, it would be difficult to assess others' work engagement based on the number of hours they work because that can be indicative of workaholism. Further, as the literature noted, the perception of others' behavior of can also influence one's own behavior through its effect of their perception of the results of performance management. Since no scale items were found for perception of evidence-based practice nor perception of others' behavior, new items were generated following the steps proposed by Gable and Wolf (1993), discussed in Chapter 3, Section 1.1.

To reduce response bias like acquiescence bias, items were re-worded to include positive and negative items. All the items were placed on a six-point Likert scale "strongly agree", "somewhat agree", "agree", "somewhat disagree", "agree", "strongly disagree" (1= strongly disagree, to 6 = strongly agree. Due to the large number of items, it was recommended that the

first survey study have few demographic items, placed at the end of the survey and designed to be optional.

Of note is the importance of role clarity that was found in the Joint Committee standards and the literature regarding organizational communication and organizational support. Employment contracts and job descriptions tend to increase in ambiguity as the complexity of the work role increases (McGovern et al., 2007). Vague or incomplete job descriptions are susceptible to a divergence in performance expectations and subsequent outcomes (McGovern et al., 2007). Therefore, the clarity and relevance of tasks, expected outcomes, and role boundaries are especially necessary for complex roles. It was thus expected, that the items regarding job descriptions and role clarity would withstand the factor analyses in the survey studies.

As previously mentioned, the scoping literature review was conducted in 2017 hence its search parameters included studies published up to 2017. Since then however, a seminal article was published in which several relevant studies were reviewed that would have been included in this chapter had they not been published after 2017. Nonetheless, the article from Wang et al. (2020) further cemented the contribution of this research as discussed in Chapter 7. The next chapter presents the results of the first online survey in which the 130 items were administered to another group of international participants in early 2018.

Chapter 5: First Survey Results

This chapter presents the first survey study, starting with a recap of the study purpose, data collection and coding. It follows with the results of the exploratory factor analysis, estimates of reliability and MANOVA. Then, the chapter provides a discussion of a comparison between the nine conceptual perception targets from the literature review and the latent factors from the survey data. Last, a summary of the findings and a set of recommendations for the second survey study are provided.

1. Purpose

Drawing on the results of the scoping literature review (Chapter 4), the goal of this study was to empirically develop the EMI scale specifically by determining the performance management perception targets, refining the scale items to a more manageable set, and exploring the antecedents of the perception targets.

2. Data Collection and Coding

One hundred and thirty items on a six-point Likert scale were built into the Qualtrics platform. The Likert scale ranged from one to six; 6 “strongly agree”, 5 “somewhat agree”, 4 “agree”, 3 “somewhat disagree”, 2 “disagree”, 1 “strongly disagree” (reverse coding was applied for negative items). There were five demographic questions, which were optional and placed at the end of the survey, shown in Appendix B. The demographic questions collected data on participant work role, location, employment status, industry sector, and organizational size. Participant work role had two categories namely HR staff and team leaders. Location was coded by continent, while employment status was categorized into secure or precarious employment. Organizational size was determined by number of staff and categorized into five groups; very small (1-50 staff), small (51 to 200 staff), medium (201 to 2000 staff), large (2001 to 4000 staff)

and very large (over 4000 staff). Responses to industry sector were coded according to their respective economic sectors (Kellerman, 1985). Primary and secondary industries related to mining or farming, and manufacturing respectively; tertiary industries related to services like finance or retail; quaternary industries included information services like research or technology; and quinary industries encompassed public services like government.

To adequately represent the target population, people working in performance management were invited to participate. The selection criteria included HR staff and team leaders from any industry worldwide, and they were recruited through LinkedIn and ProA. Considering the lengthy nature of the survey, participants were offered a modest reimbursement (through ProA and direct email of book vouchers) to thank them for their time. Three hundred and one participants responded to the survey, only five of whom were recruited from LinkedIn. Thirty-three participants dropped out of the survey before completing all the items, while the remaining 268 participants responded to the compulsory 130 items, but only 147 of whom also responded to all the optional demographic items (54.9%).

2.1 *Demographics*

Current work role was the only demographic item answered by all 268 respondents, the majority of whom were team leaders (n = 221, 82%) and the remaining 47 (21%) were HR staff. Of the 162 respondents who indicated their location, 25% were in Australia (n = 40), 41% were in Europe (n = 66) and 35% were in North America (n = 56). Only 147 respondents indicated their organization's size; 36 (24%) respondents worked in very small organizations, 29 (20%) worked in small organizations, 36 (24%) worked in medium-sized organizations, 16 (11%) worked in large organizations, and the remaining 30 (20%) worked in very large organizations. Of the 147 participants who indicated the industry in which they worked; 19 (13%) respondents

worked in primary and secondary industries; 84 (57%) worked in tertiary industries, 14 (10%) worked in quaternary industries, and the remaining 30 (20%) worked in quinary industries. Of the 147 respondents who indicated their employment status, 24 (9%) were in precarious employment and the remaining 123 (84%) were in secure employment.

3. Data Analysis

Three analyses were used to address the study's purposes; an exploratory factor analysis, estimates of reliability using Cronbach's alpha and a MANOVA. All of which were calculated in SPSS version 24.

3.1 *Exploratory Factor Analysis*

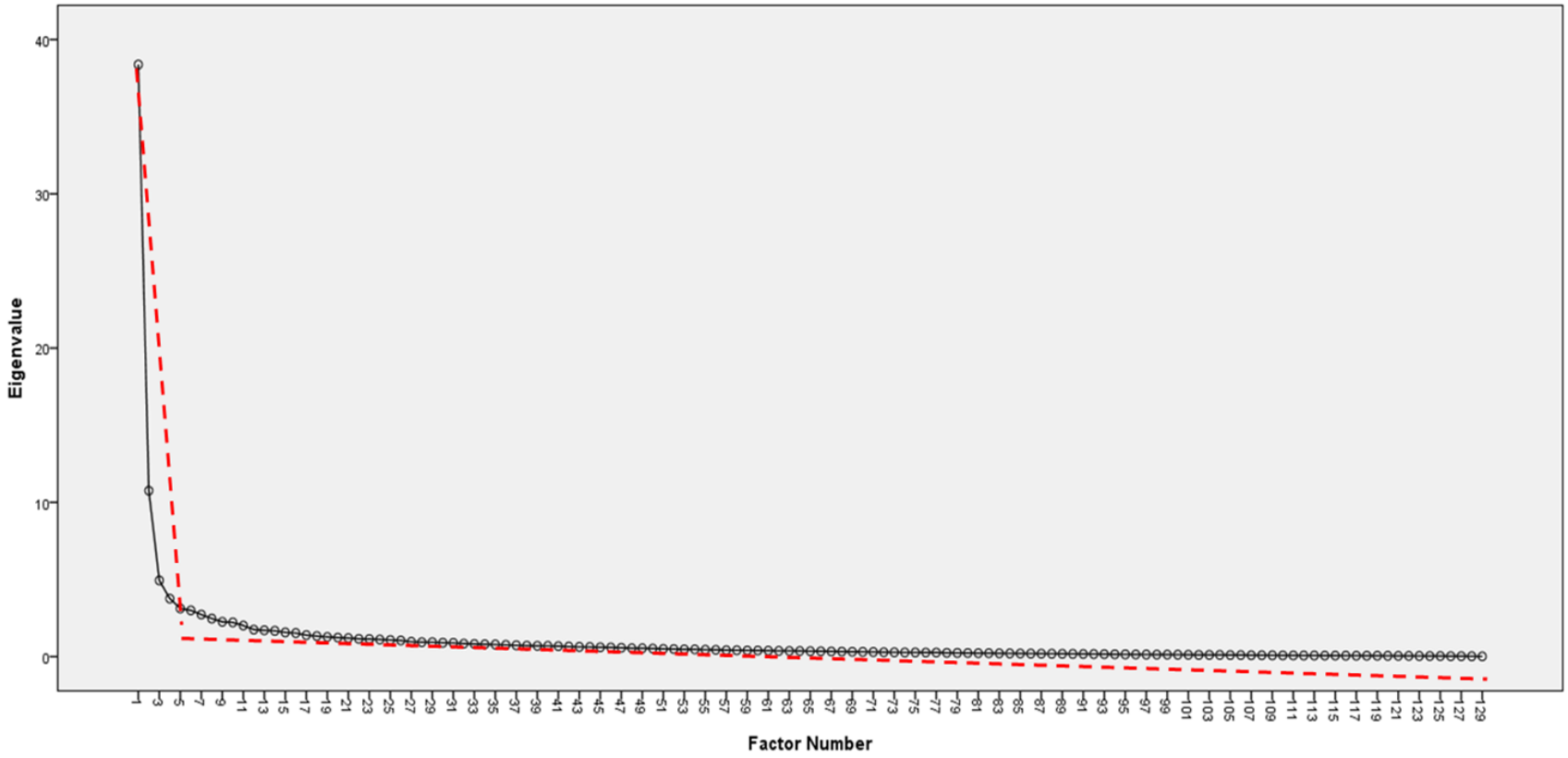
An exploratory factor analysis with maximum likelihood extraction and an oblique (promax) rotation was the first analysis conducted. Given that the achieved sample was just shy of the desired sample size, data suitability for factor analysis was judged based on three criteria; a visual examination of the data matrix for any correlations over .30, the KMO measure of sampling adequacy and, Bartlett's test of Sphericity. Several conditions were taken into consideration to ascertain the final number of factors and their names. First, using the eigenvalue rule, 19 factors with a value of 1.00 or more were noted which accounted for 62% of variance. Because this method tends to retain a more than defensible number of factors, Cattell's scree test was also used. The results of this test showed five factors above the 'elbow' of the line curve on the scree plot shown in Figure 22.

The results at this point indicated that the number of factors was between five and eleven so, a few more iterations of the exploratory factor analysis were done but with a different specified number of factors for extraction rather than by eigenvalues. At each iteration, the factor loadings of each item were examined to identify the final number of factors. Only items that

showed no cross-loading, high factor loading and clear conceptual interpretability were examined to identify each factor. An eight-factor solution was reached and considered superior to others; the KMO value obtained for this solution was .91, over and above the recommended minimum value of .60, while Bartlett's test of Sphericity reached statistical significance, $p < .001$, indicating suitability of the data for further analysis. The eight-factor solution ($\chi^2 = 2071.92$, $df = 1073$, $p < .001$) contained 55 items and explained 62% of the total variance.

Figure 22

Survey 1 Exploratory Factor Analysis Scree Plot



Note. The plot shows factors by descending order of eigenvalues.

Item Deletion. It was essential to have a balance between statistical and conceptual rigor therefore, only items with weak conceptual contribution and poor factor loading (below .30 onto a factor or loaded above .30 on multiple factors) were deleted. For example, Item 4®[18] from perception of evidence-based practice loaded poorly across all factors and was conceptually repetitive with Item 130, from perception of Accuracy[19] which loaded highly (.83) on Factor 1. However, Item 87, from perception of propriety[20], which loaded on Factor 1 (.34) and Factor 7 (.51) was kept due to its conceptual importance. Moreover, it was not surprising to see such cross loading considering the conceptual similarities between Factors 1 and 7.

Factor Naming. The items were inspected repeatedly for conceptual trends, based on which the eight factors were named. A comparison between these factors and the nine perception targets identified in Chapter 2 is presented in Section 4 of this chapter. Factor 1 was named Performance Evaluation because items loading on this factor described perceptions of the formal structures and procedures for evaluating staff performance. Performance management rated highly on this factor would be perceived as having accurate, proper, feasible, and useful performance evaluation policies and practices. The items defining this factor stemmed from the Joint Committee standards.

---

[18] 'Our performance management system is not regularly evaluated to ensure its continued usefulness & positive impact on staff performance'
[19] 'A system is in place to allow the periodic review of the performance evaluation system to ensure its continued usefulness'
[20] 'Performance evaluation policies are written & available to all staff & relevant stakeholders'

Factor 2 was named Critical staff Withdrawal because its items described the perceptions of unplanned absence and unplanned lateness of critical staff like tenured, permanent, and high performing employees. All items were reverse coded, therefore, performance management rated high on this factor would be perceived as successful in limiting undesirable and costly withdrawal of critical staff.

Factor 3 was called Supervisory Support because its items described perceptions of the types of support received from the respondent's supervisor such as fair conflict resolution, mentorship, and feedback. Performance management rated high on this factor would be perceived as having positive supervision practices. Factor 4 was named Organizational Support because its items described perceptions of the support received from the organization in terms of providing appropriate leave, job security, challenging tasks, and feedback. Performance management rated high on this factor would be perceived to deliver or facilitate positive support from the organization to the employees.

Factor 5 was called Turnover as its items described perceptions of the loss of human capital. All items were reverse coded, therefore, performance management rated high on this factor would be perceived as successful in limiting loss of human capital. Factor 6 was called New staff Withdrawal because its items described unplanned absence and unplanned lateness of staff new to the organization and recent graduates. All items were reverse coded, therefore, performance management rated high on this factor would be perceived as successful in limiting withdrawal of new staff and the associated costs.

Factor 7 was named Documentation because its items described perceptions of the clarity, accessibility, and strategic alignment of performance management documentation like job descriptions. Performance management rated high on this factor would be perceived as having strategic, clear, and accessible information crucial to performance achievement. Factor 8 was named Climate as its items described the perceptions of the day-to-day work atmosphere influenced by general communication and organizational politics. All items were reverse coded, therefore performance management rated high on this factor would be perceived as having work environments in which new ideas are welcomed, work opportunities are fair, and communication is constructive. Table 1 shows the pattern matrix of the items and their factor loadings.

Table 1

Survey 1 Factor Pattern Matrix

Item	Factor							
	1	2	3	4	5	6	7	8
Factor 1: Performance Evaluation								
1. There are safeguards to ensure that the results of an evaluation are not influenced by preconceived ideas of the manager that may be unrelated to the actual job performance	**.92**	.01	-.06	.03	.06	-.01	-.16	-.01
2. There are safeguards in place to ensure that the results of any given evaluation would be the same regardless of manager	**.89**	.02	-.01	.01	.16	-.13	-.14	-.06
3. There are safeguards in place to ensure that the evaluation procedures are the same regardless of manager	**.88**	.03	-.03	.07	.10	-.05	-.07	-.09
4. There are safeguards of the manager's final reports & disposition to ensure accuracy & use of data	**.84**	-.02	.05	.04	.03	.07	-.12	-.05
5. A system is in place to allow the periodic review of the performance evaluation system to ensure its continued usefulness	**.83**	-.05	.13	-.11	-.05	.04	-.07	.04
6. There is a system in place ensuring that the managers follow procedures, analyze all appropriate data, & report judgements based only on the criteria set in the system	**.82**	.00	-.03	-.01	.00	-.03	-.10	.04
7. Data generated by evaluations are used in developing professional development plans	**.76**	-.05	-.03	-.05	-.10	.06	.11	.12
8. Conclusions drawn are consistent with the data gathered	**.73**	.06	.06	.06	-.06	.05	.10	-.12
9. There is an adequate number of data sources to provide a comprehensive view of performance	**.70**	.02	-.02	.02	-.05	-.07	.10	-.03
10. There is appropriate follow-up of evaluation results	**.65**	-.04	.05	-.05	-.02	-.07	.08	.16

Item	Factor							
	1	2	3	4	5	6	7	8
11. All the performance managers have received appropriate training in the evaluation process	**.62**	-.02	.02	-.05	-.05	-.04	.13	.12
12. The expectations & scope of work for staff are clearly defined & understood	**.60**	.06	-.07	.01	.04	.05	.24	-.01
13. Feedback & professional development based on evaluation are aligned with the organisation's goals & missions	**.60**	.06	.05	-.04	-.04	.02	.26	-.07
14. Procedures for collecting data are as simple & job-embedded as possible to prevent undue staff overburdening	**.57**	-.08	-.10	.05	-.06	.08	.17	.02
15. All stakeholders have the opportunity to question the procedures or results of an evaluation	**.56**	-.15	-.08	.01	-.07	.08	.07	.19
Factor 2: Critical staff Withdrawal								
16. In general, unplanned absence is frequent among permanent staff®	-.02	**1**	.02	-.01	-.02	-.23	-.13	.07
17. In general, unplanned absence is frequent among high performers®	-.03	**.81**	.02	-.01	-.02	.11	-.07	-.04
18. In general, unplanned lateness is frequent among permanent staff®	-.09	**.81**	-.08	-.08	-.11	-.18	.11	.22
19. In general, unplanned absence is frequent among tenured staff®	-.02	**.69**	.01	.09	.11	.10	.04	-.25
20. In general, unplanned lateness is frequent among tenured staff®	.06	**.65**	-.14	.02	-.01	.08	.08	-.06
21. In general, unplanned absence is frequent among staff on a fixed-term contract®	.10	**.65**	.08	.02	.13	.17	-.17	-.10
22. In general, unplanned lateness is frequent among high performers®	-.08	**.63**	-.01	-.08	-.12	.17	.07	.15

Item	Factor							
	1	2	3	4	5	6	7	8
Factor 3: Supervisory Support								
23. My supervisor shows me support through giving me appropriate leave when needed	.00	-.05	**.97**	.05	.05	.04	-.08	-.12
24. My supervisor shows me support through advocating for my work needs with senior management	-.01	-.06	**.94**	.05	.01	.00	.02	-.06
25. My supervisor shows me support through fair conflict resolution	.02	.02	**.91**	.06	-.04	-.09	-.04	-.01
26. My supervisor shows me support through mentoring me	.04	-.07	**.82**	.05	-.06	-.02	.08	.03
27. Communication with my supervisor is difficult®	-.01	.08	**.61**	-.13	.10	.00	.00	.22
28. Feedback about my performance is respectful	.24	.06	**.32**	.02	.01	.07	.13	-.02
Factor 4: Organizational Support								
29. My organization shows me support through giving me appropriate leave when needed	.00	-.03	.04	**.96**	.04	.09	-.06	-.06
30. My organization shows me support through job security	.03	.00	.00	**.88**	-.06	.00	-.02	.11
31. My organization shows me support through fair conflict resolution policies	.10	.04	-.02	**.87**	-.04	-.03	.00	.06
32. My organization shows me support through responding appropriately to any feedback or questions I have	.08	-.04	.03	**.85**	-.01	-.01	.05	.03
33. My organization shows me support through fair job rewards	.11	.04	.05	**.74**	-.04	-.08	.02	.10
34. Generally, my work is challenging but not overwhelming	-.09	-.10	.18	**.40**	-.03	-.05	.12	.10
Factor 5: Turnover								
35. In general, turnover is high among casual staff®	.09	-.02	.03	-.08	**.72**	-.04	-.02	.04
36. In general, turnover is high among recent graduates®	.04	-.02	-.16	.05	**.71**	-.01	.02	.15
37. In general, turnover is high among staff on a fixed-term contract®	-.05	.20	.02	.00	**.69**	-.16	.02	-.04
38. In general, turnover is high among low performers®	-.05	-.22	.07	-.09	**.67**	-.01	-.09	.12

Item	Factor							
	1	2	3	4	5	6	7	8
39. In general, turnover is high among tenured staff®	-.11	.03	.07	.01	**.57**	.10	.22	-.02
40. In general, turnover is high among new staff®	-.01	-.07	.00	.05	**.57**	.12	.01	.24
41. In general, turnover is high among permanent staff®	-.10	.21	.04	-.01	**.42**	.08	.22	-.01
Factor 6: New staff Withdrawal								
42. In general, unplanned lateness is frequent among recent graduates®	.05	.08	-.07	.07	-.10	**.86**	-.02	-.04
43. In general, unplanned lateness is frequent among new staff®	.00	-.03	-.03	-.07	-.06	**.86**	.00	.12
44. In general, unplanned absence is frequent among new staff®	-.01	-.04	.07	-.07	.08	**.75**	-.12	.07
45. In general, unplanned absence is frequent among recent graduates®	-.06	.11	-.03	.07	.12	**.68**	.04	-.07
Factor 7: Documentation								
46. Job descriptions are clearly written & understood	.15	-.07	-.06	-.03	.06	-.01	**.85**	-.02
47. Job descriptions are aligned with organizational goals & good practice.	.13	.06	-.05	.09	.04	-.06	**.77**	-.04
48. My job description accurately & clearly describes my tasks, responsibilities, timeframes & expectations	.12	.02	.12	.01	-.05	-.08	**.53**	.08
49. Performance evaluation policies are written & available to all staff & relevant stakeholders.	.34	-.02	.08	-.06	.02	.02	**.51**	-.20
Factor 8: Climate								
50. When it comes to opportunities at work favoritism plays a big role®	.00	.07	-.02	.17	.07	.02	-.09	**.69**
51. When it comes to opportunities at work organizational politics play a big role®	-.01	-.16	-.09	.08	.20	.01	-.05	**.68**
52. Feedback about my performance is rarely followed-up / actioned®	.30	.19	-.01	-.05	-.03	.05	-.04	**.48**

Item	Factor							
	1	2	3	4	5	6	7	8
53. At my workplace, the communication climate is dominated by complaints®	.10	.09	-.03	.01	.22	-.04	.10	**.41**
54. Creative solutions & new ideas are not welcomed at work®	.05	.15	.21	.00	.03	.04	-.04	**.40**
55. Feedback about my performance is disengaging®	.05	.21	.18	.01	-.12	.16	.03	**.38**

Note. N = 286. Maximum likelihood extraction with oblique (promax with Kaiser normalization) rotation.
Reverse-coded items are denoted with ®.
Factor loadings greater than .30 are in bold

Table 2 shows the correlations between the eight factors, several of which were positively correlated to a moderate degree (r = .50 – .69). The correlations indicate the potential for the presence of higher order factors. Very weak negative correlations were present between New staff Withdrawal and three other factors; Performance Evaluation, Organizational Support and Documentation.

Table 2

Survey 1 Factor Correlations

Factor	PE	CW	SS	OS	TU	NW	DO	CL
PE	—							
CW	.06	—						
SS	**.55****	.20**	—					
OS	**.65****	.04	**.66****	—				
TU	.18**	.45**	.29**	.17**	—			
NW	-.09	**.54****	.10	-.05	.46**	—		
DO	**.65****	.15*	.45**	.49**	.19**	-.02	—	
CL	**.51****	.41**	**.55****	**.52****	.42**	.27**	.37**	—

Note. N = 286. $^{*}p < 0.05$. $^{**}p < .01$. Correlations greater than .50 are in bold.
CL = Climate.
CW = Critical staff Withdrawal.
DO = Documentation.
NW = New staff Withdrawal.
OS = Organizational Support.
PE = Performance Evaluation.
SS = Supervisory Support.
TU = Turnover.

Table 3 shows the factor means, standard deviations, and estimates of reliability. The factor means were adjusted for the number of items and the highest means related to perception of Documentation, New staff Withdrawal, Supervisory Support, and Critical staff Withdrawal. The lowest means related to perception of Climate, Turnover, Organizational Support, and Performance Evaluation. All the factors showed excellent estimates of reliability with Cronbach's alpha equal to .80 or higher. Much confidence, therefore, was placed on the total scores of these factors in subsequent analyses.

Table 3

Survey 1 Factor Means, Standard Deviations and Estimates of Reliability

Factor	Number of Items	Cronbach's α	*M*	*SD*
Performance Evaluation	15	.95	4.04	.99
Critical staff Withdrawal	7	.90	4.22	1.10
Supervisory Support	6	.92	4.09	1.36
Organizational Support	6	.95	3.79	1.57
Turnover	7	.85	3.73	1.06
New staff Withdrawal	4	.88	4.26	1.18
Documentation	4	.85	4.29	1.07
Climate	6	.83	3.58	1.03

Note. $N = 286$.

3.2 *Multivariate Analysis of Variance*

Two of the components of perception, the perceiver and the situation, were used as independent variables to examine the extent to which they influence the eight perception targets. The independent variables for this analysis were participant work role, employment status, location, organizational size and industry sector. Prior to running the MANOVA, checks for the assumptions was undertaken. The variables were measured at the interval level, the groups were independent, each survey was completed independently by the participants, the sample size was sufficiently large to detect meaningful differences, there were no outliers on any of the scales, and there was no multicollinearity among the factors. Given that the assumptions were met, the MANOVA was undertaken with confidence.

Work Role. Of the 268 respondents, 47 (18%) worked as HR staff and 221 (82%) worked as team leaders at the time of the survey. Multivariate tests, $\Lambda = .88$, *Mult. F* (8, 259) = 4.46, $p < .001$, showed that a respondent's role had a statistically significant impact on their responses. The univariate ANOVAs, shown in Table 4, indicated that these overall differences were primarily related to the responses regarding six of the factors; Performance Evaluation, Documentation, Organizational Support, Turnover, New staff Withdrawal and Critical staff Withdrawal.

Table 4

Survey 1 Statistically Significant Effects of Work Role

Factor	*F* ratio	*df*
Performance Evaluation	7.85**	1
Documentation	5.22*	1
Organizational Support	9.34**	1
Turnover	8.82**	1
New staff Withdrawal	12.31**	1
Critical staff Withdrawal	12.85***	1

Note. $N = 303$. Team leader $n = 221$, HR staff $n = 47$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

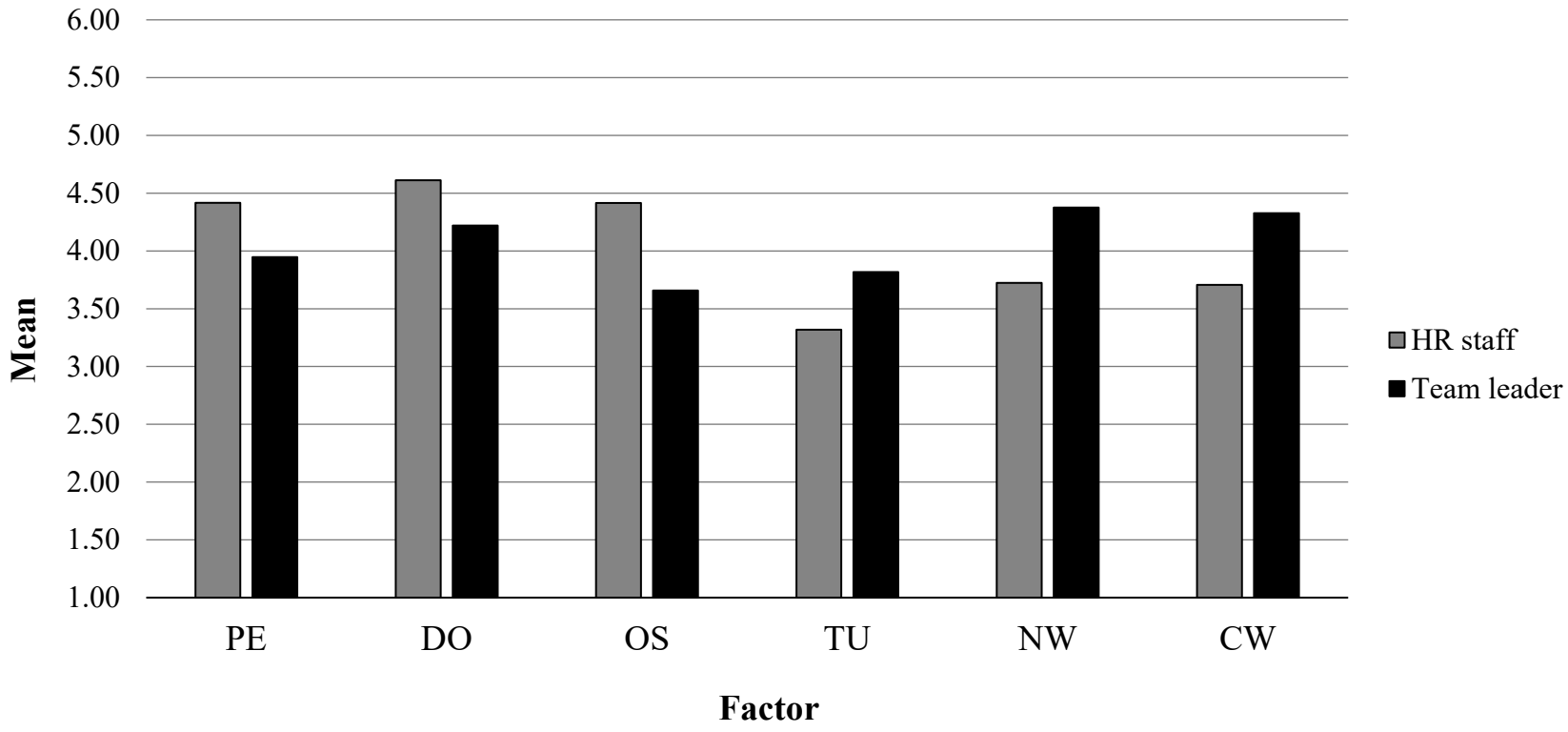
Within the respondent cohort, HR staff had more favorable perceptions of Performance Evaluation at their organization than team leaders, $\overline{X}_{(1)} = 4.42$, $SD = .90$; $\overline{X}_{(2)} = 3.95$, $SD = 1.07$, $d = .48$. They also had more favorable perceptions of performance related documentation, like their job descriptions or performance policies, than team leaders ($\overline{X}_{(1)} = 4.61$, $SD = .87$; $\overline{X}_{(2)} = 4.22$, $SD = 1.10$, $d = .40$). Additionally, HR staff had more favorable perceptions of the support they obtain from their organization than team leaders ($\overline{X}_{(1)} = 4.41$, $SD = 1.04$; $\overline{X}_{(2)} = 3.66$, $SD = 1.63$, $d = .57$). Such support included job security, fair rewards, and receiving appropriate leave when needed.

A reversed pattern was noted in perceptions of Turnover, New staff Withdrawal and Critical staff Withdrawal. Team leaders had more favorable perceptions of turnover rates in their organization than HR staff ($\overline{X}_{(1)} = 3.82$, $SD = 1.03$; $\overline{X}_{(2)} = 3.32$, $SD = 1.10$, $d = .47$). They also had more favorable perceptions of the rates of unplanned lateness or absence of critical staff, like those on a permanent contract, than HR staff ($\overline{X}_{(1)} = 4.33$, $SD = 1.03$; $\overline{X}_{(2)} = 3.71$, $SD = 1.27$, $d = .54$). Moreover, they had more favorable perceptions of the rates of unplanned lateness or absence of new staff than HR staff ($\overline{X}_{(1)} = 4.38$, $SD = 1.10$; $\overline{X}_{(2)} = 3.72$, $SD = 1.41$, $d = .52$).

Figure 23 shows the effect of role on responses to the six aforementioned factors. In particular, it illustrates that the most notable effect was seen in perceptions of Organizational Support (moderate effect size according to Cohen (1988)).

Figure 23

Survey 1 Statistically Significant Effects of Work Role



Note. $N = 303$. HR staff $n = 47$. Team leaders $n = 221$. PE = Performance Evaluation, DO = Documentation, OS = Organizational Support, TU = Turnover, NW = New staff Withdrawal, CW = Critical staff Withdrawal.

Location. Of the 268 respondents, only 162 indicated their work location, of whom 40 (25%) were in Australia, 66 (41%) were in Europe, and 56 (45%) were in North America at the time of the survey. Multivariate tests, $\Lambda = .82$, *Mult. F* (16, 304) = 1.94, $p = .02$, showed that a participant's work location had a statistically significant effect on their responses. The univariate ANOVAs, shown in Table 5, showed that these overall differences were primarily related to responses regarding five of the factors; Performance Evaluation, Documentation, Supervisory Support, Organizational Support, and Climate.

Table 5

Survey 1 Statistically Significant Effects of Location

Factor	*F* ratio	*df*
Performance Evaluation	3.25*	2
Documentation	5.65**	2
Supervisory Support	6.44**	2
Organizational Support	7.11**	2
Climate	3.53*	2

Note. $N = 303$. Australia $n = 40$, Europe $n = 66$, North America $n = 56$. * $p < .05$. ** $p < .01$. *** $p < .001$

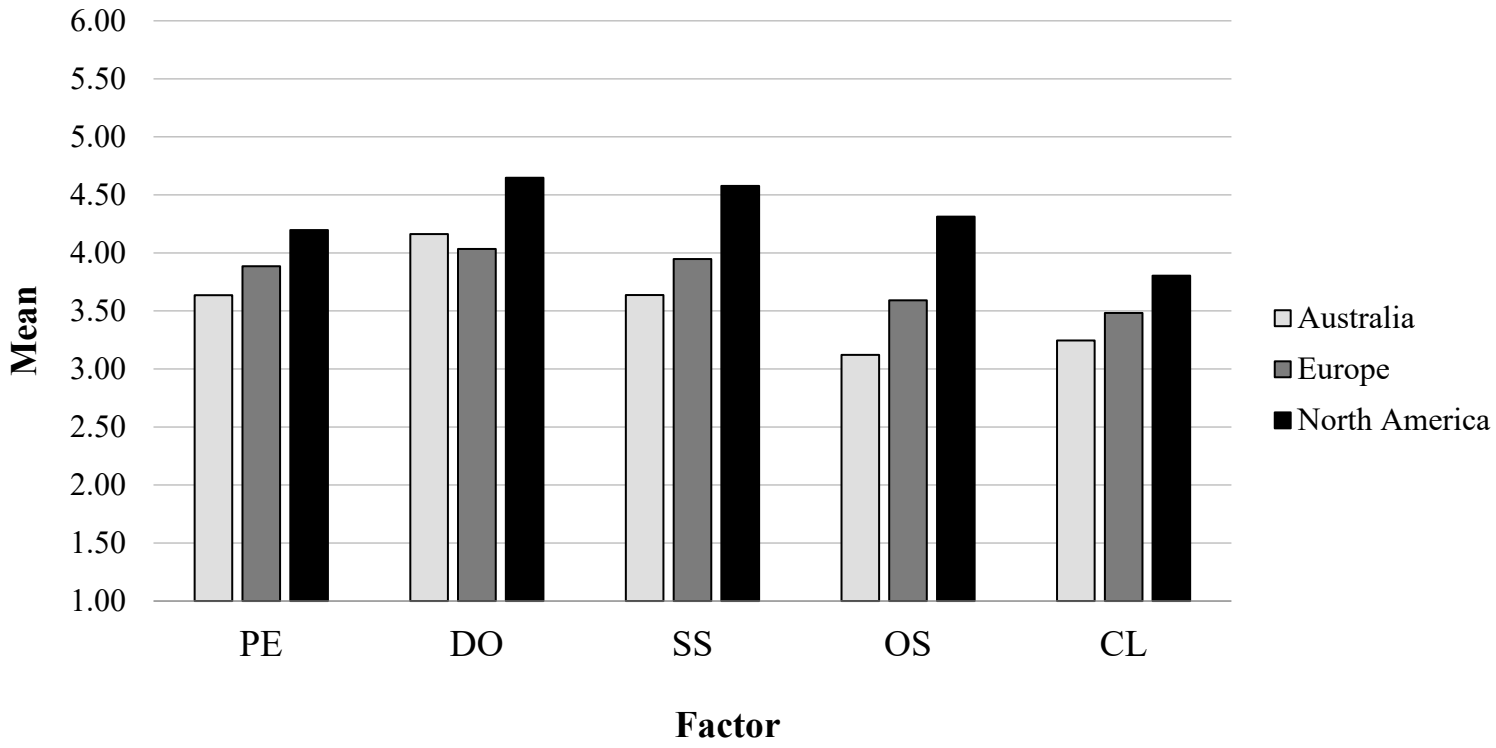
Within the participant cohort, those who were in North America had the most favorable perceptions of Performance Evaluation at their organization, while those who were in Australia had the least favorable perceptions ($\overline{X}_{(1)} = 4.20$, $SD = 1.12$; $\overline{X}_{(2)} = 3.64$, $SD = 1.07$, $d = .52$). A similar pattern was noted for Supervisory Support, Organizational Support, and Climate. Participants in North America had the most favorable perceptions of the support they receive from their supervisors, like mentorship, while those in Australia had the least favorable perceptions ($\overline{X}_{(1)} = 4.58$, $SD = 1.29$; $\overline{X}_{(2)} = 3.64$, $SD = 1.42$, $d = .70$).

Likewise, participants in North America had the most favorable perceptions of the support they receive from their organization than those who were in Australia ($\overline{X}_{(1)}$= 4.31, *SD* = 1.56; $\overline{X}_{(2)}$= 3.12, *SD* = 1.61, *d* = .76). Participants in North America also had the most favorable perceptions of their work climate, which included aspects of organizational politics or day-to-day communication, while those who were in Australia at the time of the survey had the least favorable perceptions ($\overline{X}_{(1)}$= 3.80, *SD* = 1.12; $\overline{X}_{(2)}$= 3.25, *SD* = .95, *d* = .54).

Respondents in North America had most favorable perceptions of performance related documentation, however, unlike the other factors, it was those who were in Europe who had the least favorable perceptions ($\overline{X}_{(1)}$= 4.65, *SD* = .85; $\overline{X}_{(2)}$= 4.03, *SD* = 1.07, *d* = .59). Figure 24 shows the effect of participant location on responses to the five aforementioned factors. Above all, it illustrates that the most notable effect was seen in perceptions of Organizational Support (moderate effect size according to Cohen (1988)), as was seen in the previously presented section on the effect of participant work role.

Figure 24

Survey 1 Statistically Significant Effect of Participant Location



Note. $N = 303$. Australia $n = 40$, Europe $n = 66$, North America $n = 56$. PE = Performance Evaluation, DO = Documentation, SS = Supervisory Support, OS = Organizational Support, CL = Climate

Organization Size. Of the 268 respondents, only 147 indicated their size of their organization, of whom 36 (24%) worked in small organizations, 65 (44%) worked in medium sized organizations, 16 (11%) worked in large organizations, and the remaining 30 (20%) worked in very large organizations. Multivariate test results, $\Lambda = .77$, *Mult. F* (24, 395) = 1.54, $p = .05$, showed that the size of the organization in which a participant worked at the time of the survey had a statistically significant effect on their responses. The univariate ANOVAs, shown in Table 6, showed that these overall differences were primarily related to responses regarding perception of Turnover, Climate and New staff Withdrawal.

Table 6

Survey 1 Statistically Significant Effects of Organization Size

Factor	*F*	*df*	*p*
Climate	2.94	3	.04
Turnover	2.74	3	.05
New staff Withdrawal	2.84	3	.04

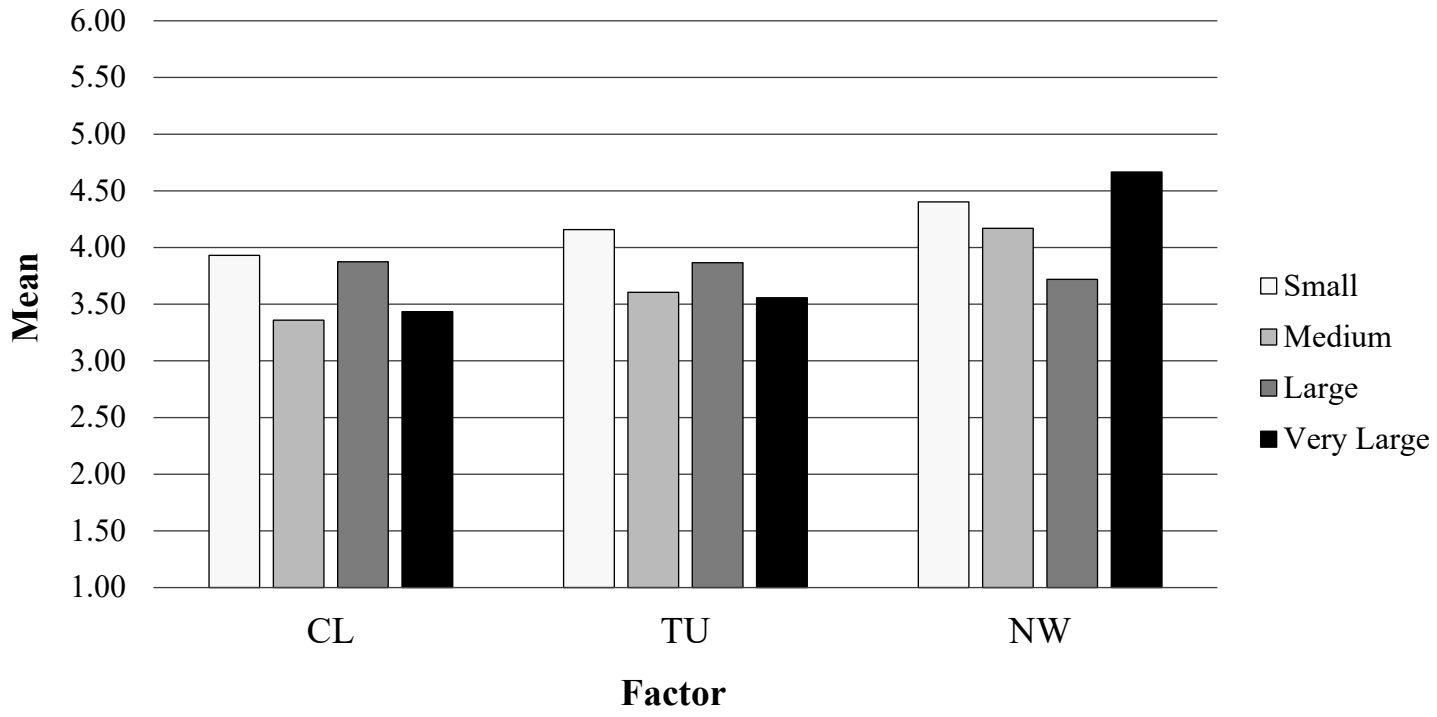
Note. N = 303. Small organization *n* = 36*,* medium organization *n* = 65*,* large organization *n* = 16*,* very large organization *n* = 30.

Participants in small organizations had the most favorable perceptions of their work climate, while those in medium sized organizations had the least favorable perceptions ($\overline{X}_{(1)}$= 3.93, *SD* = .95; $\overline{X}_{(2)}$= 3.36, *SD* = 1.05, *d* = .54). Similarly, those in small organizations had the most favorable perceptions of turnover rates in their organization; however, it was those in very large organizations who had the least favorable perceptions ($\overline{X}_{(1)}$= 4.16, *SD* = .99; $\overline{X}_{(2)}$= 3.56, *SD* = 1.13, *d* = .58).

Conversely, participants who worked in very large organizations had the most favorable perceptions of unplanned lateness or absence of new staff, while those in large organizations had the least favorable perceptions ($\overline{X}_{(1)}$= 4.67, *SD* = 1.05; $\overline{X}_{(2)}$= 3.72, *SD* = 1.22, *d* = .85). Figure 25 shows the effect of the organization's size on participant responses to the three aforementioned factors. In particular, it illustrates that the most notable effect was seen in New staff Withdrawal behavior (large effect size according to Cohen (1988)).

Figure 25

Survey 1 Statistically Significant Effect of Organization Size



Note. $N = 303$. Small $n = 36$, Medium $n = 65$, Large $n = 16$, Very Large $n = 30$. CL = Climate, TU = Turnover, NW = New staff Withdrawal

Non-Statistically Significant Effects. Industry sector and employment status showed no statistically significant effects on the factors, as shown in Table 7.

Table 7

Survey 1 Multivariate Results of Non-Statistically Significant Effects

Independent Variable	Λ	*F* (df1, df2)	p
Industry sector	.84	1 (24, 395)	.46
Employment status	.96	.76 (8, 138)	.64

Note. Λ = Wilk's multivariate test criterion, *p* = exact *p* value, *F* (df1, df2) = F distribution with hypothesis degrees of freedom and error degrees of freedom.

4. Construct Comparison

The eight factors were compared to the nine conceptual performance management perception targets from the literature review, Chapter 2. To recap, the nine targets included perception of four evaluation standards (propriety, feasibility, utility and accuracy) and five performance management constructs; evidence-based practice, organizational communication, organizational support, work engagement and others' behavior. The comparison of the original nine targets and the eight factors from this study showed some noteworthy changes. As discussed in Chapter 4, Section 4, the original perception of others' behavior target included items that described perceptions of others' performance and three withdrawal behaviors, namely turnover, unplanned lateness, and unplanned absence. This target changed in three ways; first, all the performance items were dropped due to their unacceptably poor factor loading. This change was somewhat expected as the literature showed that perception of others' behavior is affected more by negative experiences than positive ones (Kim et al., 2015; Op de Beeck et al., 2016).

Second, Turnover emerged as a stand-alone factor, which was only somewhat unexpected but not baffling because even though much of the literature included turnover as one of three main withdrawal behaviors, it has been a focal subject of a large body of research (Allen et al., 2003; Bothma & Roodt, 2013; Branham, 2012; Eley et al., 2007; Holtom et al., 2008; Yamamoto, 2013; Yousaf et al., 2018). Third, unplanned lateness and unplanned absence emerged as two latent factors but with respect to specific types of staff, accordingly named Critical staff Withdrawal, and New staff Withdrawal. This change was unexpected for two reasons; first, despite stemming from the same conceptual target, the two factors were unlike Turnover as the latter did not relate to a specific type of staff. Second, the literature reviewed only somewhat indicated the need to discriminate between staff in relation to their withdrawal behavior (Farrell & Petersen, 1984; Kahn, 1990; Nagadevara et al., 2008; Rhoades & Eisenberger, 2002).

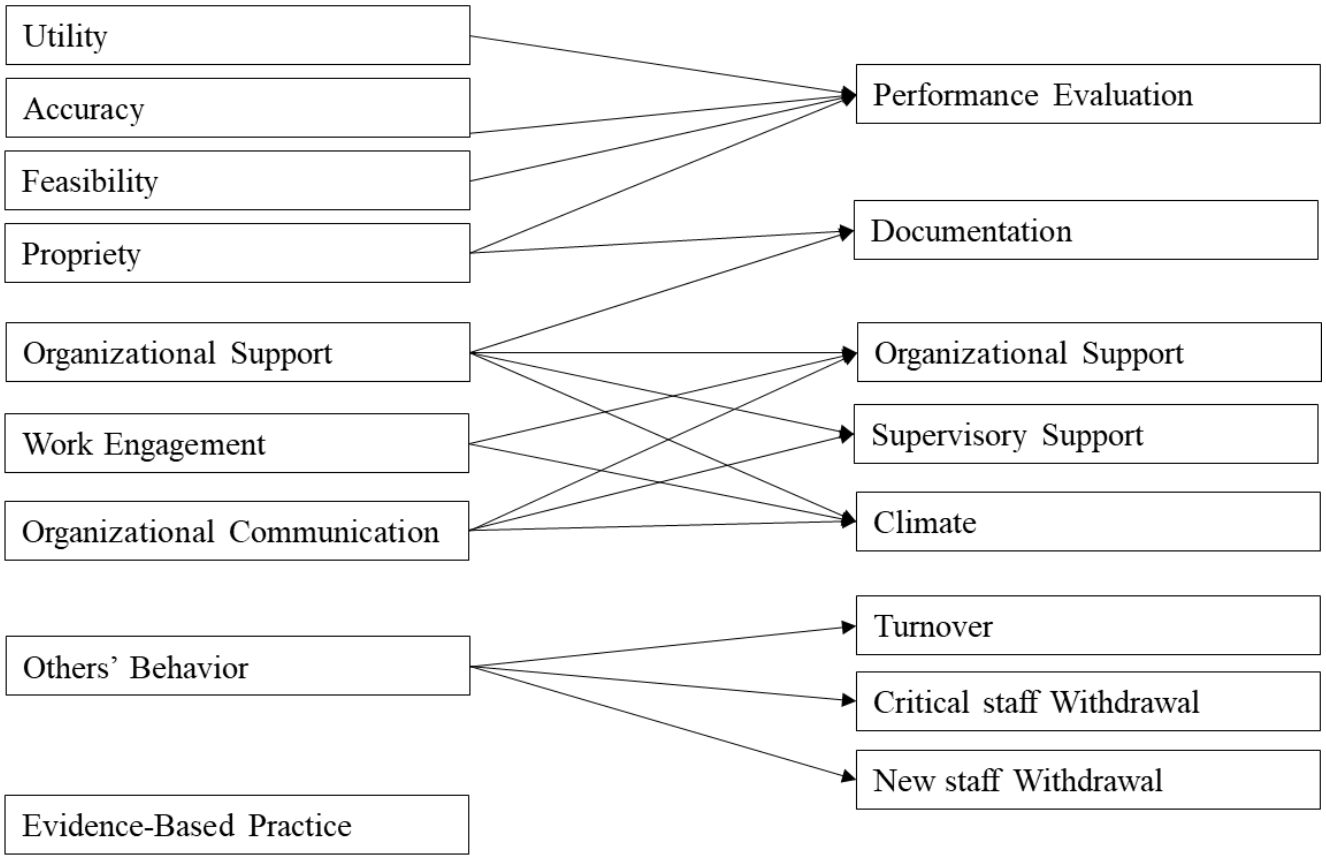
Items from three original perception targets (organizational support, work engagement, and organizational communication) loaded onto four latent factors, namely; perception of Documentation, Supervisory Support, Organizational Support, and Climate. In particular, the separation of the original organizational support target into two factors was not surprising given the debate within the literature on the difference between support from one's organization and one's direct supervisor (Kottke & Sharafinski, 1988; McCarthy et al., 2013; Rhoades & Eisenberger, 2002).

Items from the four evaluation standards combined to form the strongest factor; Performance Evaluation. This indicates that the previously untested list of key questions from the Joint Committee has sufficient psychometric qualities that merit further investigation. Lastly, the original evidence-based practice target was dropped for two reasons; first, some of its items had

extremely poor factor loading. Second, the remainder of its items had somewhat poor factor loading but only multiple factors, and were conceptually much weaker than other similar items. A discussion of this finding is provided in Chapter 7. Figure 26 illustrates the nine conceptual performance management perception targets and the latent factors.

Figure 26

Relationship between the Hypothesized Conceptual Performance Management Perception Targets and the Eight Latent Factors



Note. The arrows illustrate where the items from a conceptual perception target loaded on a latent factor.

5. Chapter Summary

This chapter described the results of the first survey study in which the items generated from the first stage of research were tested. The study determined eight perception targets for performance management from the exploratory factor analysis, and reduced the 130 items to a parsimonious set of 55 items, which accounted for the most variance. The eight perception targets (factors) were perception of; Performance Evaluation, Documentation, Supervisory Support, Organizational Support, Climate, Turnover, Critical staff Withdrawal, and New staff Withdrawal. Estimates of reliability using Cronbach's alpha calculations were high enough for all eight factors, indicating good internal consistency, to confidently proceed with further analysis. The MANOVA showed that a participant's work role, location and the size of their organization had statistically significant effects on the eight perception targets.

Several factors were statistically correlated to a degree indicating that there may be at least one higher order factor. Therefore, it was recommended that the confirmatory factor analysis in Survey 2 should examine both the eight factor model and a higher order factor model. The survey also revealed that some improvements were needed to avoid missing data. For example, because the item pool was considerably reduced, survey fatigue was not as much of a concern for the second survey so, all demographic items required a response. In addition, more demographic items were included, like participant gender and age to determine the effect of perceiver characteristics, as noted in Chapter 2 Section 5.1, on participant responses. The next chapter provides the results of the second online survey.

Chapter 6: Second Survey Results

This chapter begins with a recap of the second survey's purpose, data collection and coding. It follows with a presentation of the results of the confirmatory factor analysis, estimates of reliability, and MANOVA. The chapter concludes with a summary of the findings to be discussed in the final chapter of this thesis.

1. Purpose

The purpose of the second online survey was to verify the eight factor structure of the EMI scale, developed in the first online survey. Moreover it aimed to ascertain the statistical significance of several antecedent factors relating to the characteristics of the perceiver and the situation on the eight performance management perception targets measured by the EMI scale. Lastly, it aimed to explore the possibility of a higher order factor model as recommended by the first survey's findings.

2. Data Collection

The 55 items from survey one were built in the Qualtrics platform on a six-point Likert scale. In addition to the original five demographic items from the first survey, four demographic items were included in this survey. The items (shown in Appendix C) related to respondent age, gender, work role, organizational size, workforce experience, tenure, industry, location, and employment status. The Likert scale was numerically coded as in survey one (strongly disagree = 1, disagree = 2, somewhat disagree = 3, somewhat agree = 4, agree = 5, strongly agree = 6), and negatively worded items were inversely coded accordingly. Similar to Survey 1, Qualtrics was used to administer the survey, which was distributed via the ProA crowdsourcing platform as well a survey link advertised on LinkedIn.

The survey was open to an international audience from any industry but only to those working in performance management. This survey was significantly shorter that the first survey and required a maximum of ten minutes to complete. All items were compulsory

2.1 *Demographics*

Three hundred and three people responded to the survey, only three of whom were recruited from the LinkedIn link and the remaining 300 from ProA. Most respondents were female (62%, *n* = 189), 38% were male (*n* = 114). This is similar to the target population demographic data which showed, in Australia, at least 57% of managers in HR were female (Department of Employment Skills Small and Family Business 2018). The least number of respondents were aged between 18 and 25 years (14%, *n* = 41), while 40% were 26-35 years of age (*n* = 120), 28% were 36-45 years of age (*n* = 84), and the remaining 19% (*n* = 58) were over 45 years old. Gender distribution across all age groups was consistent with more female than male respondents (between 61% - 66% female).

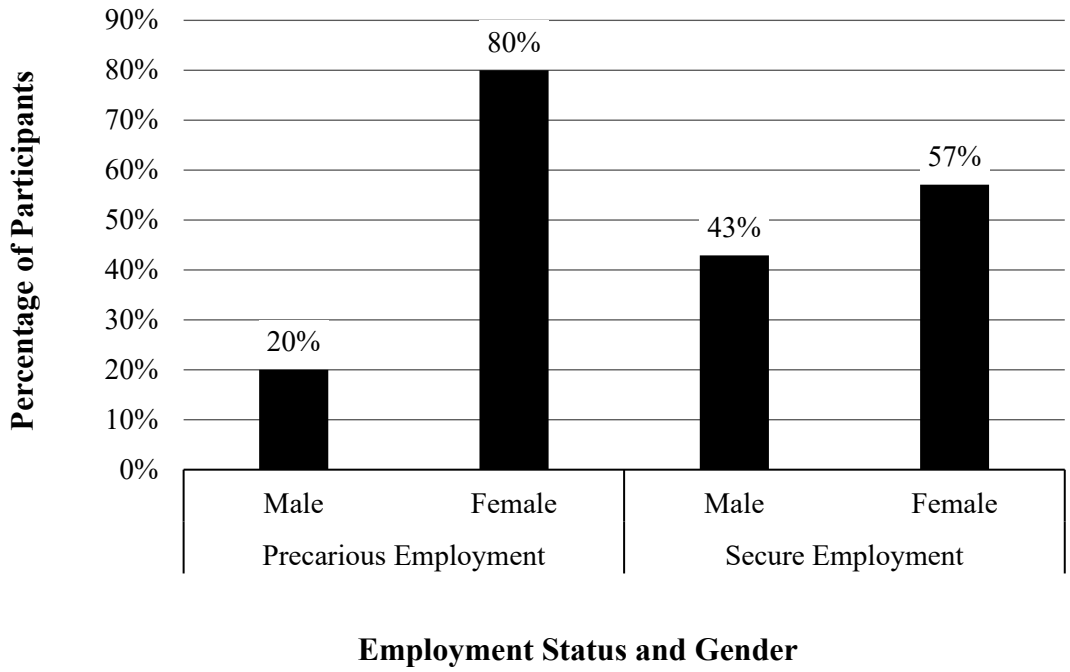
Work roles were categorized based on organizational hierarchy resulting in four groups: administrative staff (40%, *n* = 122), junior management (20%, *n* = 60), middle management (30%, *n* = 90) and, senior management (10%, *n* = 31). The majority of administrative staff and junior managers were female (81% and 67% respectively), while the majority of middle managers and senior managers were male (52% and 77% respectively). This is consistent with gender and role distribution trends in the target population which has fewer females in senior management[21] (Department of Jobs and Small Business, 2019).

[21] As of November 2018, in Australia, only 30% of Chief Executives and General Managers were female

Most respondents were in secure employment like having a permanent contract (77%, *n* = 233), and only 23% were in precarious employment, like casual workers (*n* = 70). Although the distribution of men and women in secure employment was somewhat comparable (males *n* = 100, 43%, females *n* = 133, 57%), this was not the case in precarious employment. In fact, a staggering 80% of those in precarious employment were female (*n* = 56), and only 20% were male (*n* = 14). This is consistent with the literature showing that women occupy more precarious employment than men (Vosko et al., 2009). Figure 27 illustrates the distribution of men and women within the employment status categories.

Figure 27

Survey 2 Respondent Demographics; Gender and Employment Status

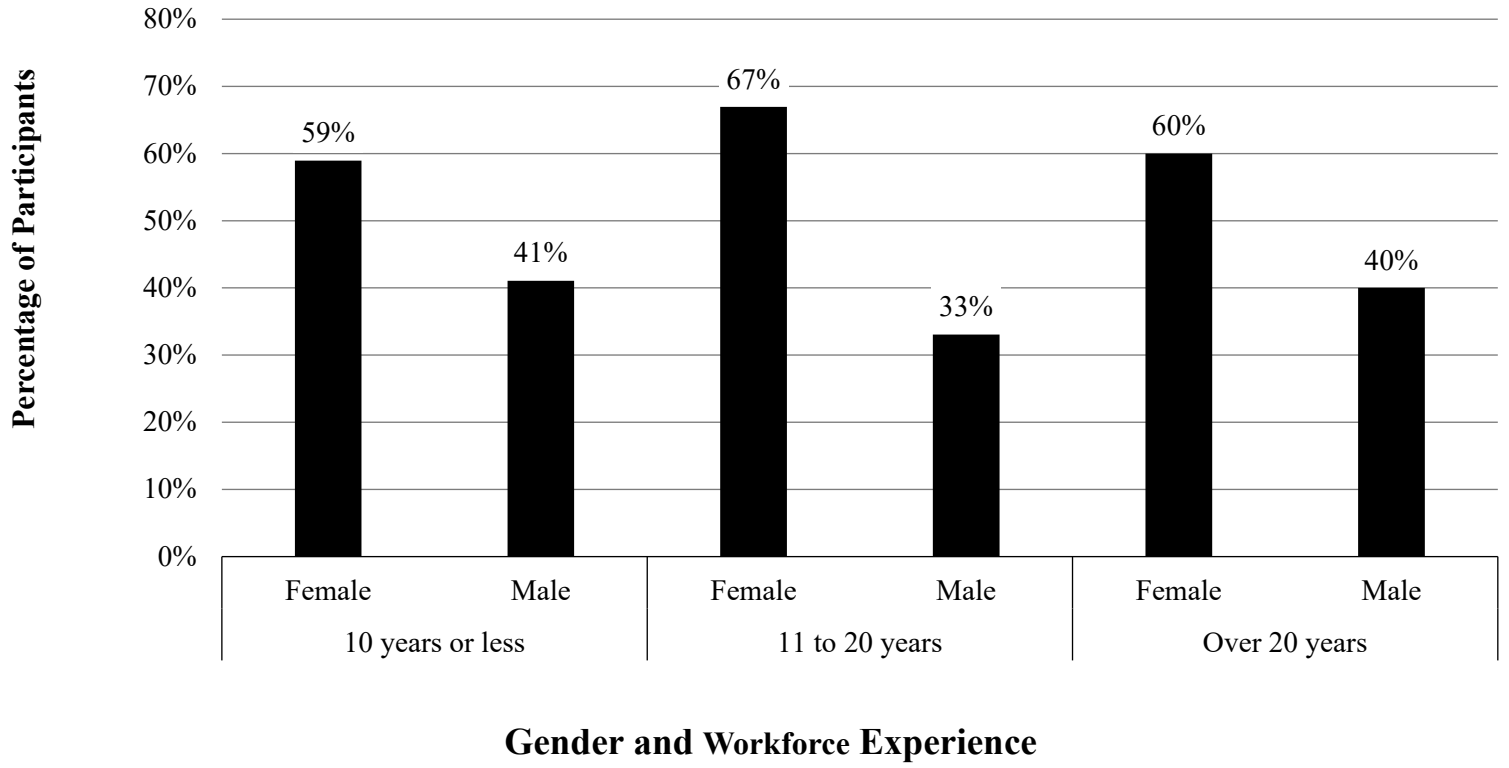


Note. N = 303. Precarious employment (males *n* = 14, females *n* = 56), secure employment (males *n* = 100, females *n* = 133).

The majority of respondents indicated that they had been with their employer for ten years or less (78%, *n* = 235), while the remaining 22% (*n* = 68) had been with their employer for over ten years. The size of the respondent's organization was divided into four groups according to the estimated number of staff: small (1-50), medium (51-1000), large (1001 – 4000), and very large (over 4000). Twenty-one per cent of respondents worked in small organizations (*n* = 64), 38% worked in medium organizations (*n* = 114), 17% worked in large organizations (*n* = 50), and the final 25% worked in very large organizations (*n* = 75). Thirty-one percent of respondents had been in the workforce for ten years or less (*n* = 95), 39% had been working for 11-20 years (*n* = 118), and the remaining 30% have been in the workforce for over 20 years (*n* = 90). As shown in Figure 28 the vast majority of those with 11 – 20 years' worth of workforce experience were female.

Figure 28

Survey 2 Respondent Demographics; Gender and Workforce Experience



Note. $N = 303$. 10 years or less (males $n = 39$, females $n = 56$), 11 – 20 years (males $n = 39$, females $n = 79$), over 20 years (males $n = 36$, females $n = 54$)

Similar to the first survey study, the industry in which respondents worked at the time of the survey was categorized into four groups based on their economic sector. Very few respondents worked in primary and secondary industries (n = 38, 13%), 119 respondents worked in tertiary industries (39%), 88 respondents worked in quaternary industries (29%), and the remaining 58 respondents worked in quinary industries (19%). The majority of respondents were in Europe at the time of the survey (n = 255, 84%) and the remaining 48 were in North America (16%).

3. Data Analysis

A confirmatory factor analysis was carried out in AMOS software version 24 to test the construct validity of the eight factors proposed by Survey 1. Additionally, estimates of reliability were calculated using Cronbach's alpha, and the effects of various independent variables were analyzed using MANOVA in SPSS version 24.

3.1 *Confirmatory Factor Analysis*

Maximum likelihood estimation methods were used, and the model's goodness-of- fit was assessed using absolute and relative indices. The absolute goodness-of-fit indices were; 1) the Chi-square (χ^2) goodness-of-fit; 2) the Root Mean Square Error of Approximation (RMSEA); 3) the Standardized Root Mean Square Residual (SRMR), and 4) the Goodness-of-Fit Index (GFI). The relative goodness-of-fit index calculated was the Comparative Fit Index (CFI).

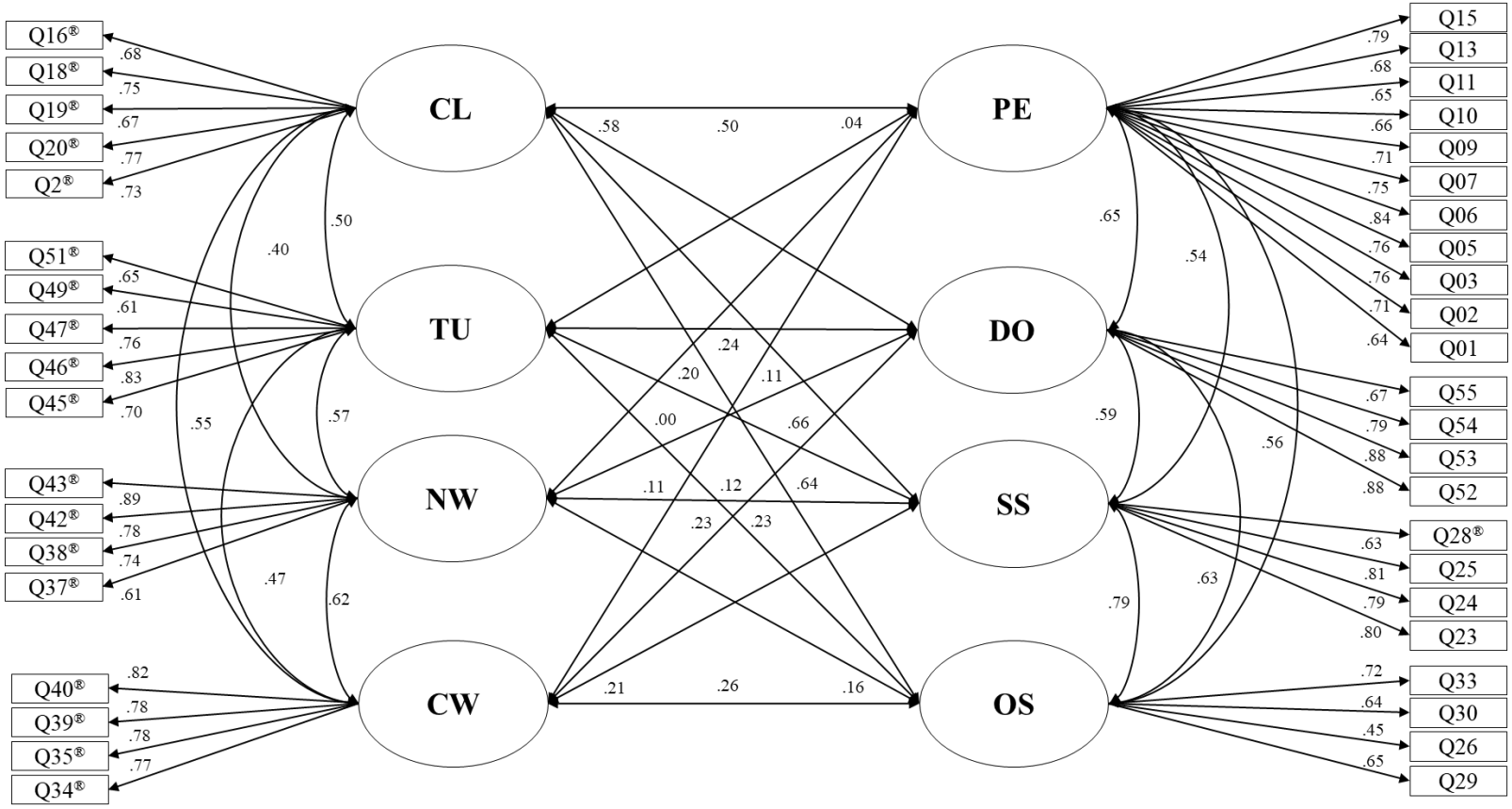
The χ^2 result of 3661.03 was statistically significant (df = 1402, $p < .001$) suggesting a rejection of the model however, this index is sensitive to large sample sizes. Overall, the model was considered to fit somewhat well (RMSEA = .07, SRMR = .07, GFI = .69, CFI = .80, CMIN/DF = 2.61), but there was much room for improvement to this model.

The analysis identified several items that loaded below .45 onto a factor, however, because it was important to strike a balance between statistical and conceptual rigor only items that had low factor loading and were conceptually weaker than others were removed. As a result, 14 items were deleted, and the resultant model provided enough items to clearly identify the desired factor interpretation. The eight-factor solution contained 41 items and its goodness of fit measures indicated better model fit ($\chi 2$ = 1862.41, df = 751, p < .001, RMSEA = .07, SRMR = .07, GFI = .77, CFI = .85, CMIN/DF = 2.48). Figure 29 shows the standardized paths for this model illustrating all remaining items had sufficient loadings onto the appropriate factors. Unlike the correlations found in Survey 1, all factors in this survey were positively correlated.

In Survey 1, Turnover and Organizational Support had a very week positive correlation (r = .17), whereas in Survey 2, no linear relationship was found between the two factors (r = .00). Moreover, in Survey 1, New staff Withdrawal and Documentation had a very weak negative correlation (r = -.02), whereas in this survey they had a very week positive correlation (r = .11). Also, in Survey 1, Performance Evaluation and Turnover had a very weak positive correlation (r = .18), however in this survey they had a weaker positive relationship (r = .04). Finally, in Survey 1, the strongest positive correlation was between Supervisory Support and Organizational Support (r = .66) and though that remained the case, the correlation was stronger in this survey (r = .79), further indicating the possible presence of a higher order factor.

Figure 29

Survey 2 Standardized Paths of the Eight-Factor Model



Note. CL = Climate, TU = Turnover, NW = New staff Withdrawal, CW = Critical staff Withdrawal, PE = Performance Evaluation, DO = Documentation, SS = Supervisory Support, POS = Organizational Support. Reversed coding is denoted with ®.

Table 8 shows the number of items per factor, estimates of reliability, factor means, and standard deviations. As this table shows, the EMI scale indicated acceptable to excellent reliability. The factor means were adjusted for the number of items (by calculating the mean item value across the items within each factor) and similar to Survey 1, New staff Withdrawal, Documentation, and Supervisory Support had higher means than Climate, Performance Evaluation, and Turnover.

Table 8

Survey 2 Factor Means, Standard Deviations and Estimates of Reliability

Factor	Number of Items	α	*M*	*SD*
Performance Evaluation	11	.92	4.14	.89
Documentation	4	.88	4.32	1.08
Organizational Support	4	.71	4.36	.85
Supervisory Support	4	.84	4.26	1.02
Climate	5	.85	3.68	1.03
Turnover	5	.84	3.91	1.09
New staff Withdrawal	4	.85	4.39	1.03
Critical staff Withdrawal	4	.87	4.23	1.12

Note: $N = 303$

Table 9 shows the factor correlations and highlights a similar pattern to the one noted in the first survey study. Specifically, that several factors were positively correlated to a moderate to high degree ($r \geq .50$ but less than .80), and some had much lower relationships indicating unique contributions of these factors to the overall model.

Table 9

Survey 2 Factor Correlations

	CW	NW	TU	CL	SS	OS	DO	PE
CW	—							
NW	**.62**	—						
TU	.47	**.57**	—					
CL	**.55**	.40	**.50**	—				
SS	.21	.12	.20	**.66**	—			
OS	.26	.16	.23	**.64**	**.79**	—		
DO	.23	.11	.25	**.58**	**.49**	**.63**	—	
PE	.11	.00	.04	**.50**	**.54**	**.56**	**.65**	—

Note.
$N = 286$.
$^{*}p < 0.05$. $^{**}p < .01$.
Correlations ≥ .50 are in bold.
CL = Climate.
CW = Critical staff Withdrawal.
DO = Documentation.
NW = New staff Withdrawal.
OS = Organizational Support.
PE = Performance Evaluation.
SS = Supervisory Support.
TU = Turnover.

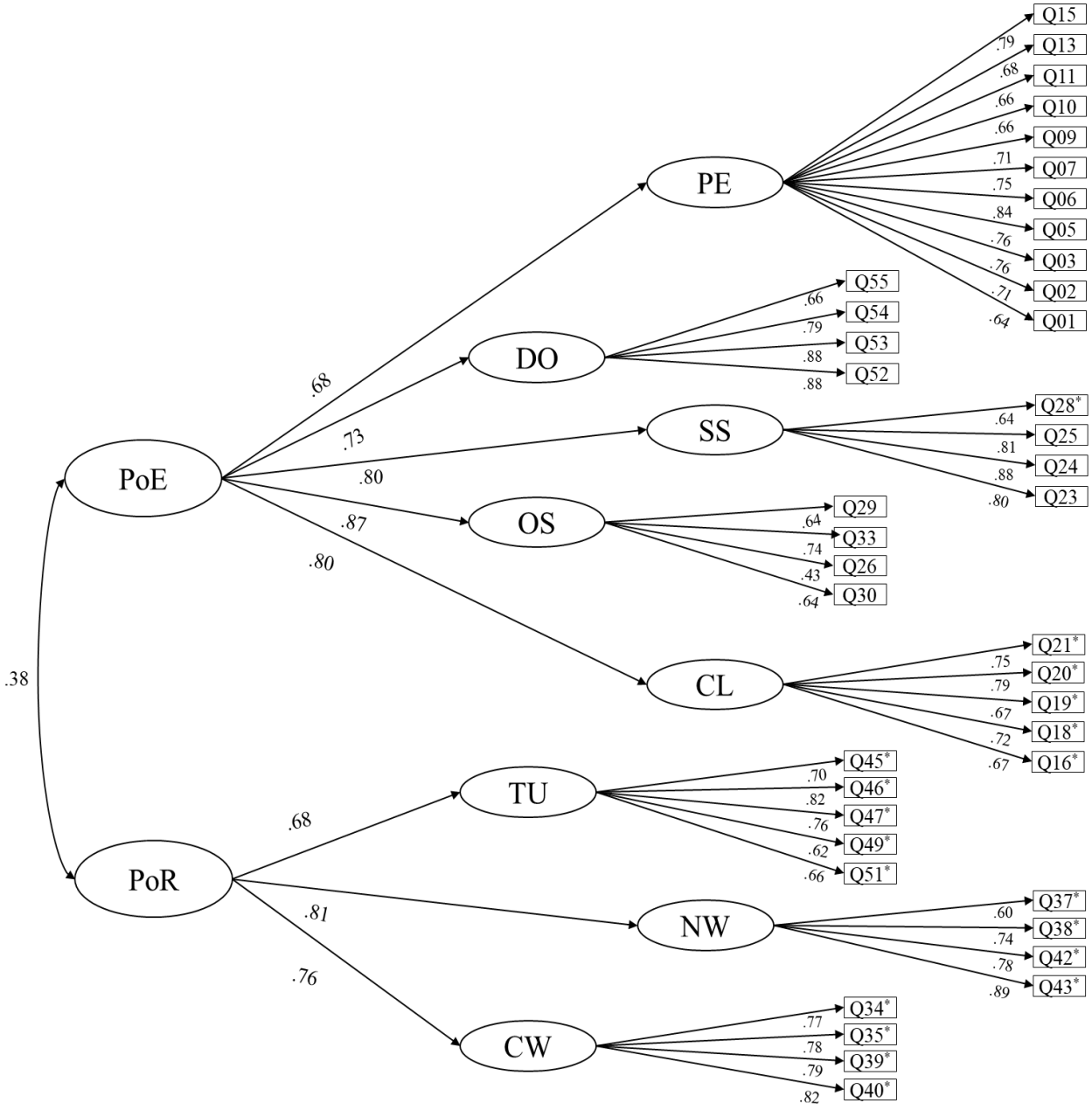
Higher Order Factors. Upon inspecting the correlations of the eight factors and the results of the first survey study, the model was tested for the possibility of a higher order structure. Figure 30 shows the standardized paths for this model. Overall, the model was considered to fit well (Chi-square = 2004.25, df = 770, RMSEA = .07, SRMR = .09, GFI = .75, CFI = .83, CMIN/DF = 2.60). As shown in Figure 30, the lower order factors loaded well onto the two higher order factors. Of note was Item Q26 (Generally, my work is challenging but not overwhelming) because it was kept despite its .43 loading due to its conceptual importance.

In particular, it provides insight into the balance between interest in work tasks and stress. Moreover, the item had a sufficient loading (.45) in the lower order model. Aside from item Q26, all items had sufficiently high loadings onto the appropriate factors, and the higher order factors had a small positive correlation (r = .38). The latter indicates that the higher order factors are positively correlated but not enough to be one and the same.

Factor Naming. The first higher order factor was called Perception of Efforts because its lower order factors related to perceptions of resources and activities an organization invests in performance management. Performance management rated high on this factor would be perceived as having good performance evaluation, clear and strategic performance management documentation, activities that support staff performance, and a positive work climate. The second higher order factor was called Perception of Results because its lower order factors related to the withdrawal behavior and turnover among workers other than the respondent. Performance management rated high on this factor would be perceived as successful in limiting undesirable and costly staff withdrawal behavior.

Figure 30

Survey 2 Standardized Paths for the Higher Order Factor Model



Note. PoE = Perception of Effort, PoR = Perception of Results, CL = Climate, TU = Turnover, NW = New staff Withdrawal, CW = Critical staff Withdrawal, PE = Performance Evaluation, DO = Documentation, SS = Supervisory Support, POS = Organizational Support. Reversed coding is denoted with ®.

To decide which of the two models fit the data best, a Chi-square difference test was done and revealed that the lower order model fit slightly better than the higher order model. However since Chi-square difference tests have the same strengths and weaknesses as Chi-square tests (applied to a single model), they too are affected by large samples and can show significance in potentially inconsequential differences (Schermelleh-Engel et al., 2003). A comparison of the RMSEA showed no difference between the two models, though the correlation matrices (see Table 2 and Table 9) indicated the likely presence of a higher order structure. Therefore, while both models were generally similar in terms of their statistical ability to explain the data, the higher order model was deemed to fit better than the lower order model as it provided more information about the data and plausible explanation of the lower order factor correlations.

3.2 *Multivariate Analysis of Variance.*

MANOVA was carried out with several independent variables on the eight lower order factors and the higher order factors. The independent variables included participant gender, age, current role, tenure, workforce experience, employment status, location, organization size, and industry type. Prior to running the MANOVA, checks for the assumptions was undertaken. The variables were measured at the interval level, the groups were independent, each survey was completed independently by the participants, the sample size was sufficiently large to detect meaningful differences, there were no outliers on any of the scales, and there was no multicollinearity among the factors. Given that the assumptions were met, the MANOVA was undertaken with confidence

Work Role. A participant's work role was categorized into four groups all of which work in performance management; 122 (40%) participants worked as administrative staff, 60 (20%) worked in junior management, 90 (30%) worked in middle management and the remaining 31 (10%) worked in senior management. Multivariate tests of the eight lower order factors, $\Lambda = .88$, *Mult. F* (24, 847.49) = 1.63, $p = .03$, showed that a respondent's role had a statistically significant effect on their responses. The univariate ANOVAs, shown in Table 10, indicated that these overall differences were primarily related to the responses regarding perception of withdrawal behavior of critical and new staff.

A similar pattern was noted in the multivariate tests of the two higher order factors ($\Lambda = .95$, *Mult. F* (6, 596) = 2.37, $p = .03$) as well as the univariate ANOVAs which showed that the overall differences were mainly related to Perception of Results ($F = 2.95$, $df = 3$, $p = .03$). Despite the difference in work role categories between the first survey and this survey, work role had a statistically significant effect on perceptions; in particular, perceptions of critical and new staff withdrawal were consistently affected by respondent work role.

Table 10

Survey 2 Statistically Significant Effects of Work Role

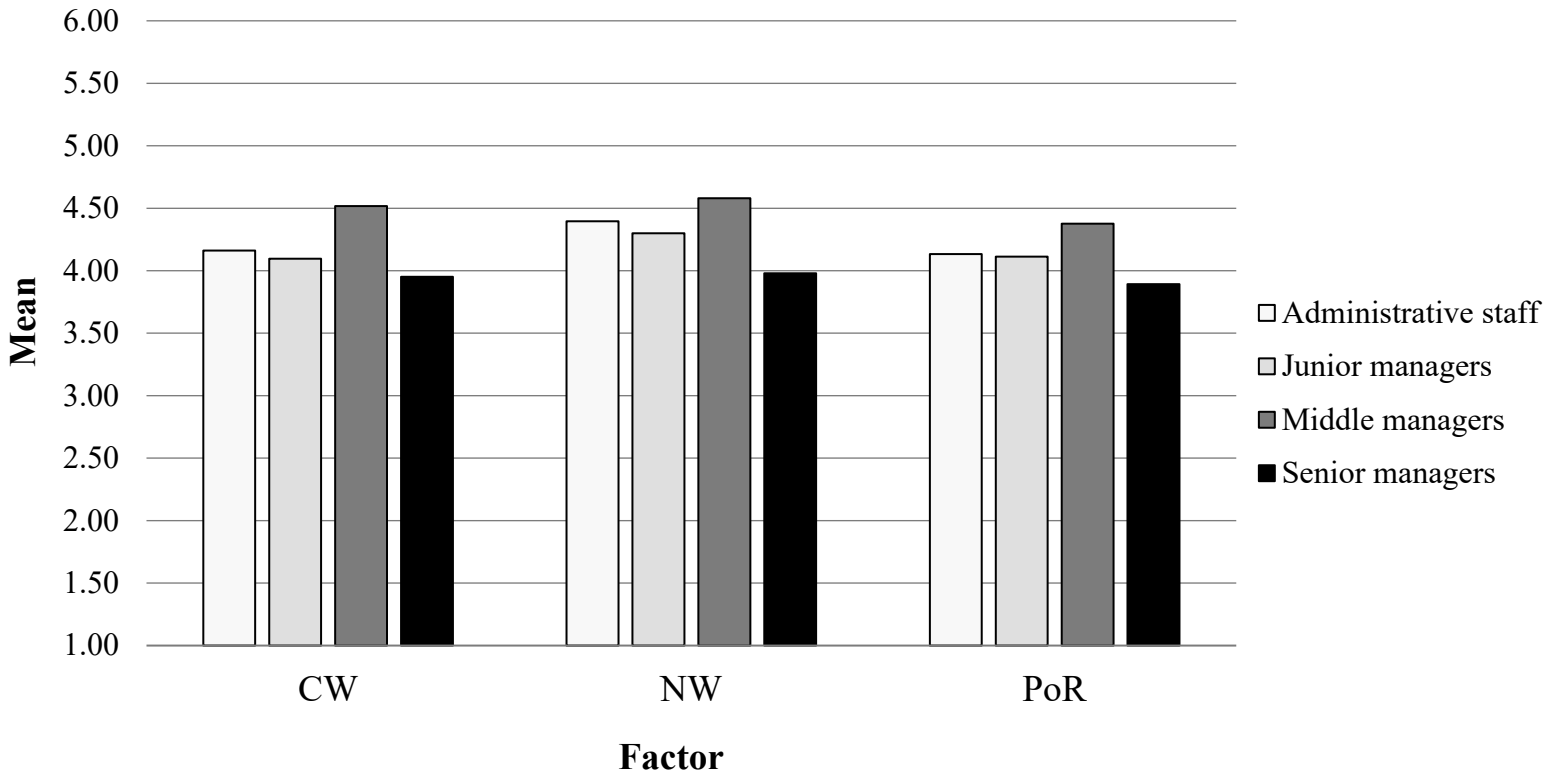
Factor	*F* ratio	*df*
Critical staff Withdrawal	3.11*	3
New staff Withdrawal	2.82*	3
Perception of Results	2.95*	3

Note. *$p < .05$. Administrative staff $n = 122$, junior managers $n = 60$, middle managers $n = 90$, senior managers $n = 31$.

Participants in middle management had the most favorable perceptions of performance management results, while those in senior management had the least favorable perceptions ($\overline{X}_{(1)}$= 4.38, SD = .84; $\overline{X}_{(2)}$= 3.89, SD = 1.19, *d* = .53). This was particularly the case with perceptions of unplanned lateness or absences of critical staff like those on a permanent contract ($\overline{X}_{(1)}$= 4.52, SD = 1.10; $\overline{X}_{(2)}$= 3.95, SD = 1.33, *d* = .49) and new staff ($\overline{X}_{(1)}$= 4.58, SD = .94; $\overline{X}_{(2)}$= 3.98, SD = 1.40, *d* = .55). Figure 31 shows the effect of work role, illustrating the larger effect of role on New staff Withdrawal.

Figure 31

Survey 2 Statistically Significant Effects of Work Role



Note. N = 303. Administrative staff *n* = 122, junior managers *n* = 60, middle managers *n* = 90, senior managers *n* = 31.
CW = Critical staff Withdrawal, NW = New staff Withdrawal, PoR = Perception of Results.

Participant Age. Age was categorized into four age ranges; 41 (14%) participants were aged between 18 and 25 years, 120 (40%) were aged between 26 and 35 years, 84 (28%) were aged between 36 to 45 years and the remaining 58 (19%) were over 45 years of age. Multivariate tests of the eight lower order factors, $\Lambda = .87$, *Mult. F* (24, 847.49) = 1.69, $p = .02$, showed that participant age had a significant impact on their responses. The univariate ANOVAs, shown in Table 11, indicated that these overall differences were primarily related to the responses regarding perceptions of performance evaluation and the withdrawal behavior of critical staff. This pattern was not noted in the multivariate test of the two higher order factors ($\Lambda = .96$, *Mult. F* (6, 596) = 1.85, $p = .09$).

Table 11

Survey 2 Statistically Significant Effects of Participant Age

Factor	*F* ratio	*df*
Performance Evaluation	3.90*	3
Critical staff Withdrawal	2.76*	3

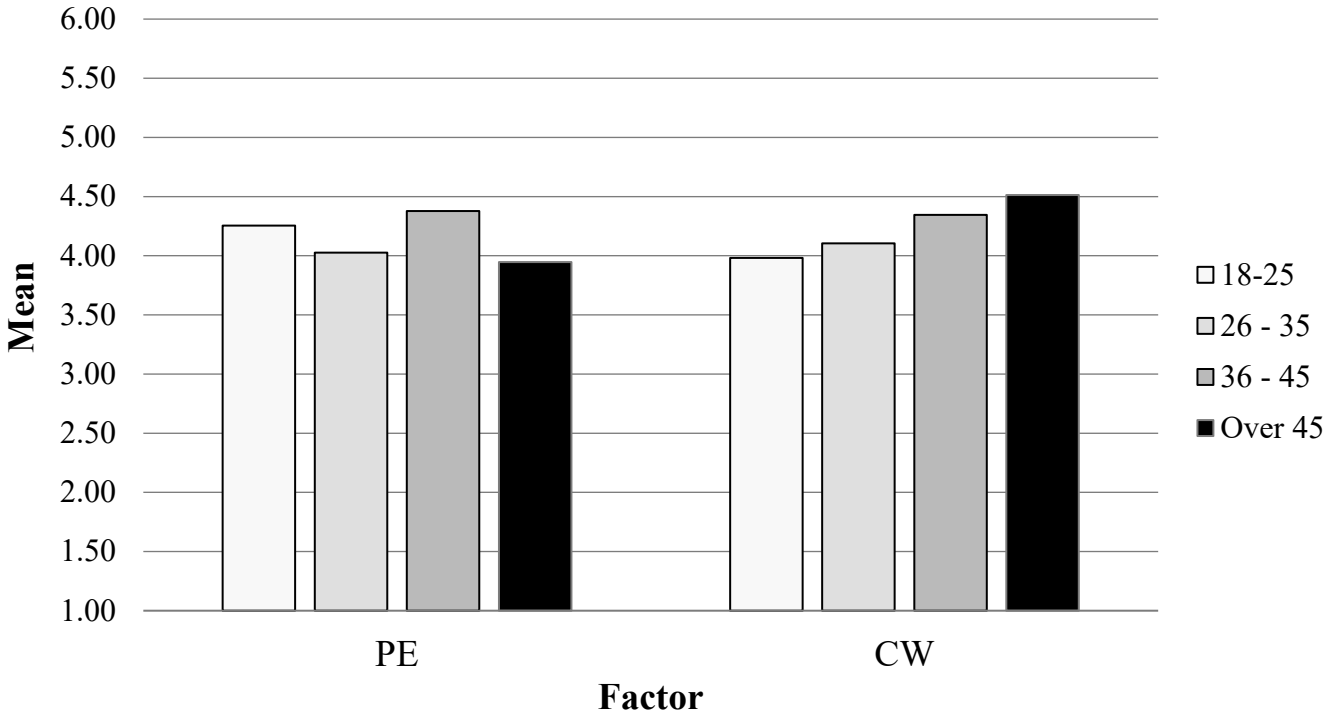
Note. *$p < .05$. Age groups: 18-25 $n = 41$, 26-35 $n = 120$, 36-45 $n = 84$, over 45 $n = 58$.

Participants aged between 36 and 45 years had the most favorable perceptions of Performance Evaluation in their organizations, while those over the age of 45 had the least favorable perceptions ($\overline{X}_{(1)}$= 4.38, SD = .74; $\overline{X}_{(2)}$= 3.95, SD = .90, *d* = .50). Such perceptions included those of the staff performance conclusions based on their performance evaluation or the alignment of feedback and professional development with organizational goals.

Participants over 45 years old had the most favorable perceptions of the rates of unplanned lateness and absence of critical staff while those aged between 18 and 25 years had the least favorable perceptions ($\overline{X}_{(1)}$= 4.51, SD = .98; $\overline{X}_{(2)}$= 3.98, SD = 1.30, *d* = .47). Figure 32 shows the small effect, according to Cohen (1988), of participant age.

Figure 32

Survey 2 Statistically Significant Effect of Participant Age



Note. $N = 303$. 18-25 $n = 41$, 26-35 $n = 120$, 36-45 $n = 84$, over 45 $n = 58$. PE = Performance Evaluation. CW = Critical staff Withdrawal

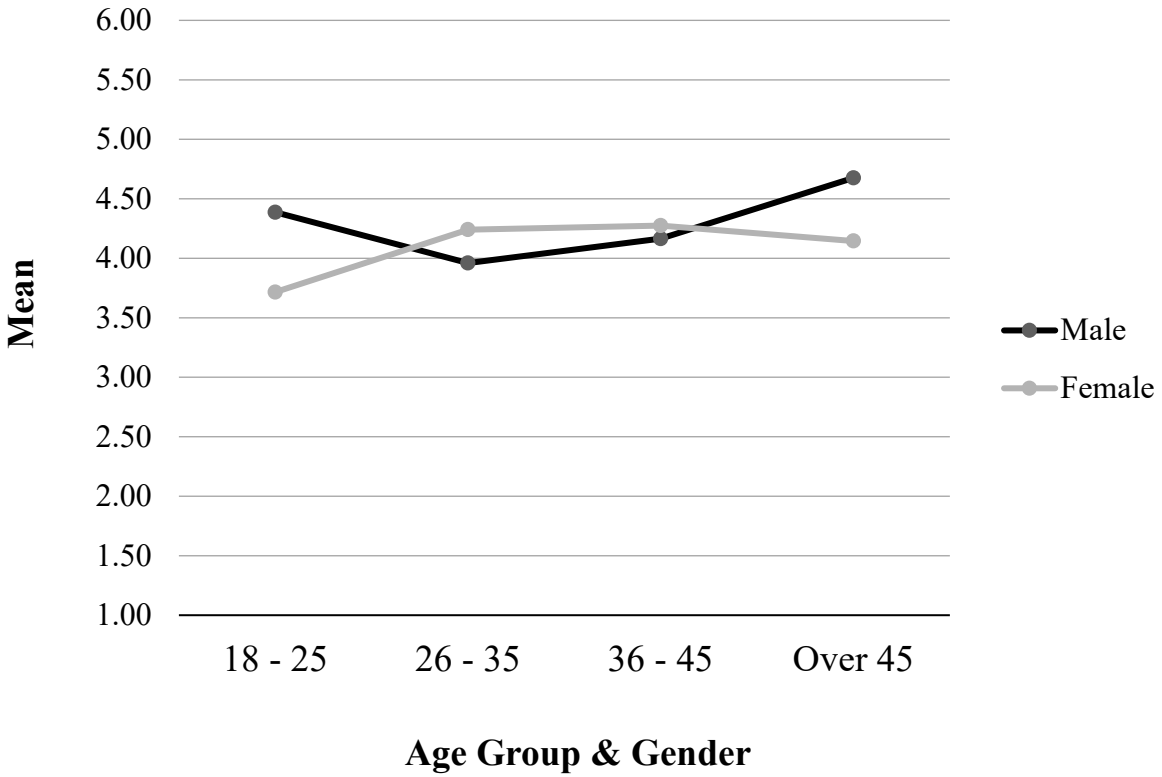
Age Group and Gender. Multivariate tests of the interaction effect of respondent's age group and gender on the eight lower order factors, Λ = .91, *Mult. F (*24, 835.89) = 1.22, p = .22, showed no statistical significance. However, the multivariate tests on the two higher order factors showed a statistically significant interaction effect, Λ = .95**,** *Mult. F* (6, 588) = 2.45, p = .03. This is not unusual because higher order factors represent constructs that are more parsimonious and perhaps more theoretically salient than the lower order factors and therefore tend to have more explanatory or predictive power (Musek, 2017).

The univariate ANOVAs indicated that these overall differences were primarily related to perception of performance management results (F = 4.75, df = 3, p < .01). Moreover, as shown in Table 14 and Table 15, respondent gender had no statistically significant effect on neither the lower order factors nor the higher order factors. This suggests that, at least for the higher order factors, a respondent's age group affects the influence of gender on perception.

Male participants over the age of 45 had the most favorable perceptions of performance management results, while female respondents who were between 18 and 25 years of age had the least favorable perceptions ($\overline{X}_{(1)}$= 4.68, SD = .68; $\overline{X}_{(2)}$= 3.72, SD = .87). As shown in Figure 33, the statistically significant interaction effect of participant age and gender on their perception of performance management results was disordinal. As Figure 33 shows, perception differences between men and women were starker in the youngest and oldest age groups than in the mid-range age groups. Moreover, perception changes over time (age) for both men and women fluctuated. In the mid-range age groups, men had less favorable perceptions of performance management results than women, but in the youngest and oldest age groups, they had more favorable perceptions than women. A discussion of the possible reasons for these patterns is offered in Chapter 7.

Figure 33

Survey 2 Statistically Significant Interaction Effect of Age Group and Gender on Perception of Results.



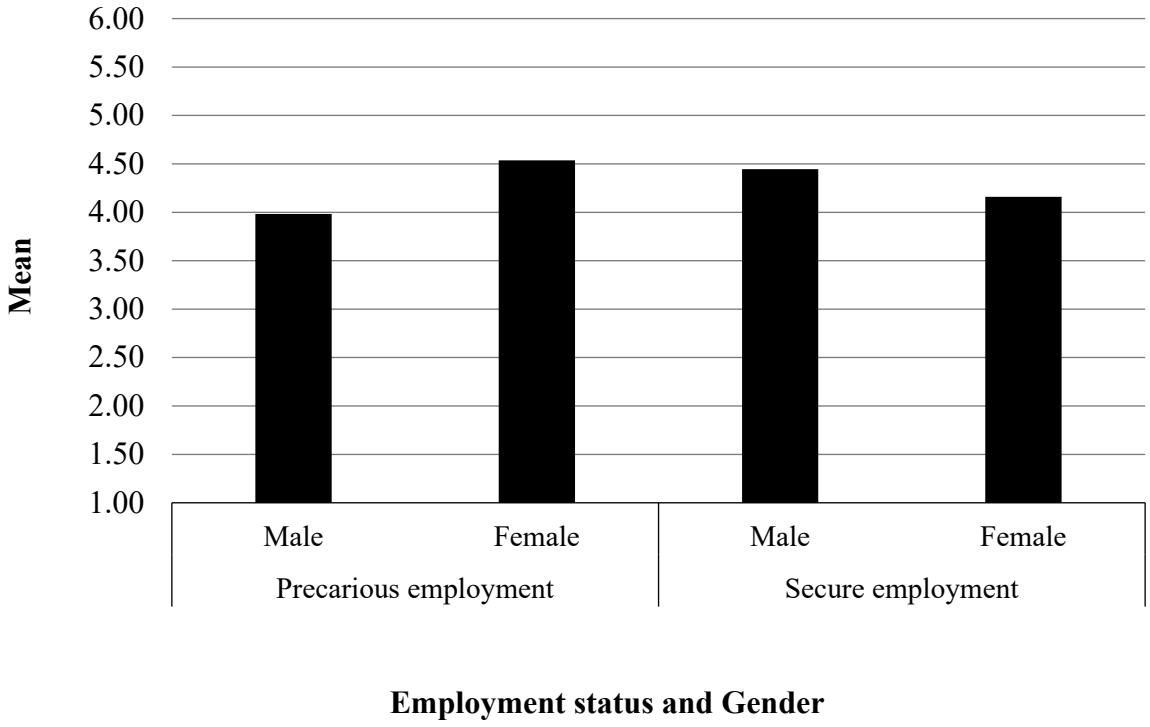
Note. N = 303. Females 18 - 25 *n* = 27, 26 - 35 *n* = 75, 36 - 45 *n* = 51, over 45 *n* = 36. Males 18 - 25 *n* = 14, 26 - 35 *n* = 45, 36 - 45 *n* = 33, over 45 *n* = 22.

Employment Status and Gender. Although neither gender nor employment status showed statistically significant effects on participant responses (see Table 14 and Table 15), there was a statistically significant interaction effect. The multivariate test results for the two higher order factors were not statistically significant**,** $\Lambda = .99$**,** *Mult. F* (2, 298) = .99, $p = .37$, the interaction effect was significant for the eight lower order factors**,** $\Lambda = .95$**,** *Mult. F (*8, 292) = 2.06, $p = .04$. The univariate ANOVAs indicated that these overall differences were largely related to perceptions of performance management documentation ($F = 5.77$, $df = 1$, $p = .02$).

Female participants who were in precarious employment at the time of the survey had the most favorable perceptions of documentation, while males, also in precarious employment had the least favorable perceptions ($\overline{X}_{(1)} = 4.54$, SD = 1.02; $\overline{X}_{(2)} = 3.98$, SD = 1.36). It must be noted that due to the small number of men in precarious employment, these results must be interpreted with caution. In Chapter 7, these results, in particular, are discussed in terms of future research directions rather than educated generalizations. Figure 34 shows the interaction effect of gender and employment status on perception of performance management documentation.

Figure 34

Interaction Effect of Gender and Employment Status on Perception of Documentation



Note. $N = 303$, precarious employment (males $n = 12$, females $n = 56$), secure employment (males $n = 100$, females $n = 133$).

Organization Size. The size of a respondent's organization was categorized into four groups according to the number of staff; 64 (21%) of respondents worked in small organizations (1-50 workers), 114 (38%) worked in medium-sized organizations (51-1000 workers), 50 (17%) worked in large organizations (1001- 4000 workers), and the remaining 75 (25%) worked in very large organizations (over 4000 workers).

Similar to the first survey study, multivariate tests of the eight lower order factors, Λ = .86, *Mult. F* (24, 847.49) = 1.85, p = .01, showed that that the size of a respondent's organization had a statistically significant impact on their responses. The univariate ANOVAs, shown in Table 12, indicated that these overall differences were primarily related to the responses regarding perception of Performance Evaluation, work Climate and Turnover rates. A similar pattern was noted in the multivariate tests of the two higher order factors (Λ = .95, *Mult. F* (6, 596) = 2.39, p = .03). The univariate ANOVAs (Table 12) showed that the overall differences were mainly related to Perception of Results (F = 3.93, df = 3, p = .01).

Table 12

Survey 2 Statistically Significant Effects of Organization Size

Factor	*F* ratio	*df*
Performance Evaluation	4.03**	3
Climate	3.11*	3
Turnover	5.34**	3
Perception of Results	3.93*	3

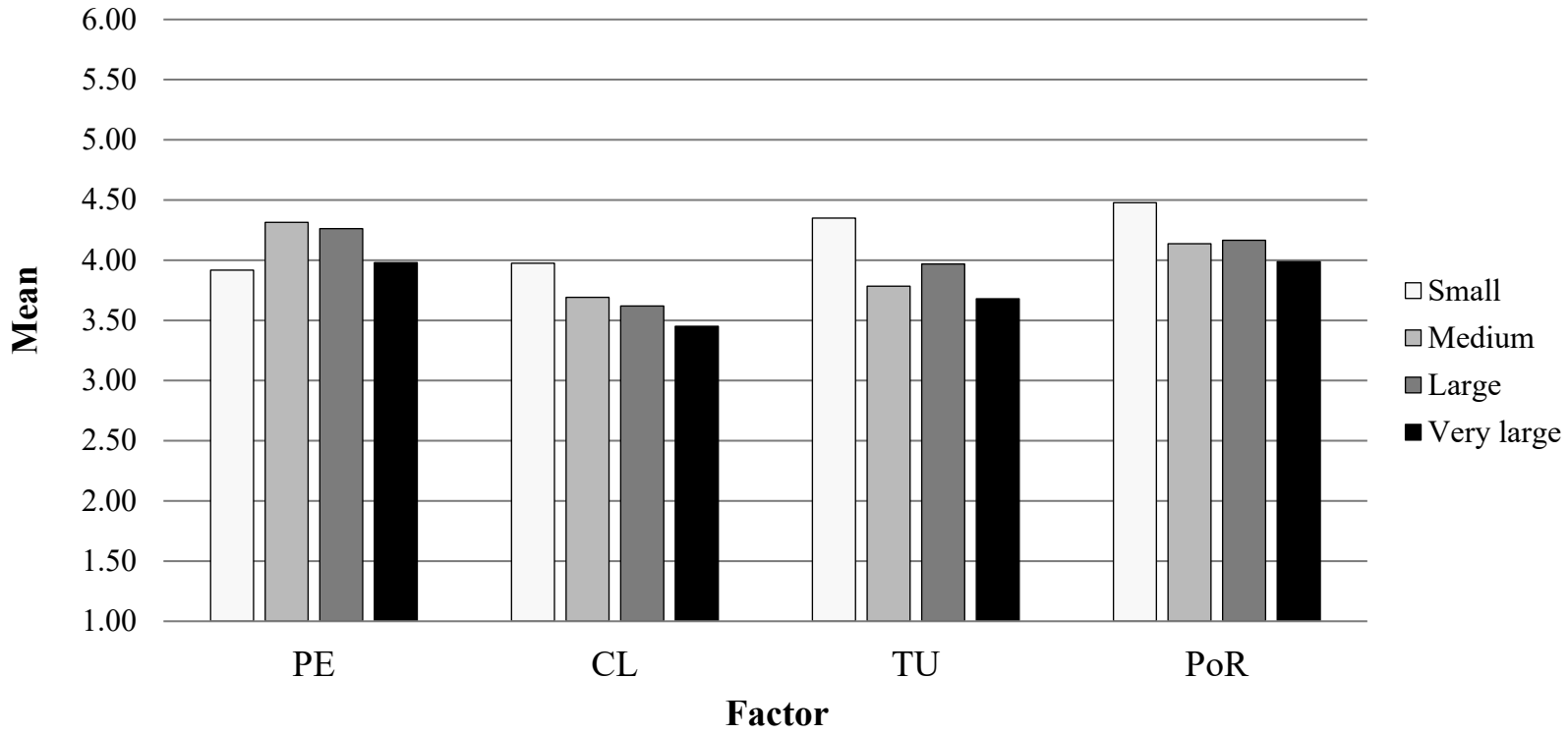
Note. *p < .05. **p < .01. Organization size: Small n = 64, Medium n = 114, Large n = 50, Very Large n = 75.

Participants who worked in small organizations at the time of the survey had the most favorable perceptions of their work climate while those in very large organizations had the least favorable perceptions ($\overline{X}_{(1)}$= 3.98, SD = 1.11; $\overline{X}_{(2)}$= 3.45, SD = .98, *d* = .52). Likewise, those in small organizations had the most favorable perceptions of turnover rates at their organization while those in very large organizations had the least favorable perceptions ($\overline{X}_{(1)}$= 4.35, SD = .97; $\overline{X}_{(2)}$= 3.68, SD = 1.05, *d* = .65).

A similar pattern was noted in perceptions of performance management results; those in small organizations had the most favorable perceptions, while those in very large organizations had the least favorable perceptions ($\overline{X}_{(1)}$= 4.48, SD = .76; $\overline{X}_{(2)}$= 3.99, SD = .88, *d* = .59). However, a different trend was noted in perceptions of Performance Evaluation in that those in large organizations had the most favorable perceptions while those in small organizations had the least favorable perceptions ($\overline{X}_{(1)}$= 4.32, SD = .82; $\overline{X}_{(2)}$= 3.92, SD = 1, *d* = .46). Figure 35 shows the effect of organization size on perception of Performance Evaluation, Turnover and Climate, noting the larger effect size on Turnover. These results showed similar trends to the first survey study; in particular the most positive perceptions of work climate and rates of turnover were consistently noted in respondents from small organizations.

Figure 35

Survey 2 Statistically Significant Effect of Organization Size



Note. $N = 303$. Small $n = 64$, Medium $n = 114$, Large $n = 50$, Very large $n = 75$. PE = Performance Evaluation, CL = Climate, TU = Turnover, PoR = Perception of Results

Workforce Experience. Participant workforce experience was categorized into three groups according to the range of number of years they had been in the workforce at the time of the survey. Ninety-five (31%) participants had been in the workforce of ten years or less, 118 (39%) had been in the workforce between 11 and 20 years, and the remaining 90 (30%) had been in the workforce for over 20 years. Multivariate tests of the eight lower order factors, $\Lambda = .92$, *Mult. F* (16, 586) = 1.67, $p = .05$, showed that respondent workforce experience had a statistically significant impact on their responses. The univariate ANOVAs, shown in Table 13, indicated that these overall differences were primarily related to the responses regarding perception of work climate and withdrawal behavior of critical staff. A similar pattern was noted in the multivariate tests of the two higher order factors ($\Lambda = .95$, *Mult. F* (4, 598) = 3.65, $p = .01$). The univariate ANOVAs showed that the overall differences were mainly related to Perception of Results ($F = 5.09$, $df = 2$, $p = .01$).

Table 13

Survey 2 Statistically Significant Effects of Workforce Experience

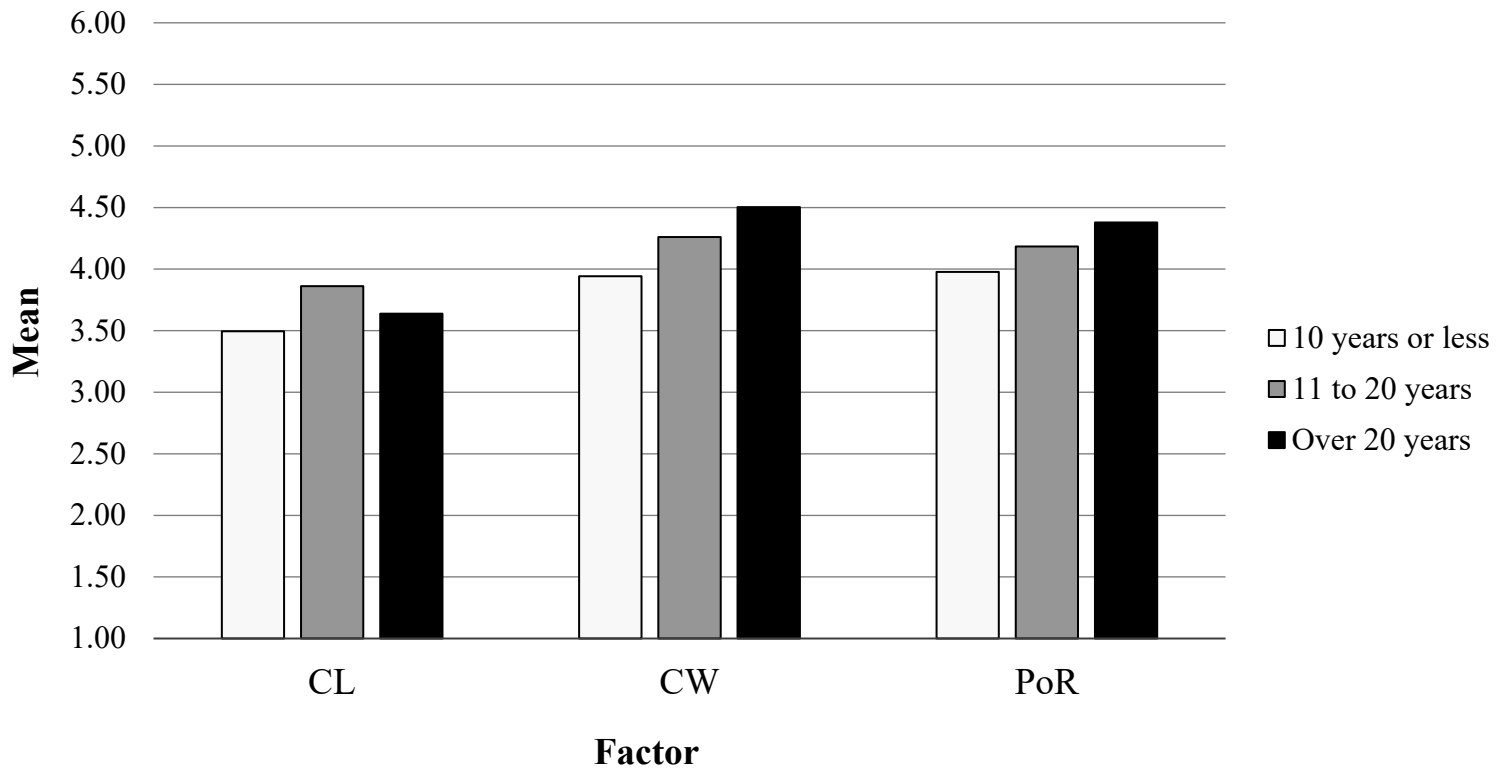
Factor	*F* ratio	*df*
Climate	3.49*	3
Critical staff Withdrawal	6.05**	3
Perception of Results	5.09*	2

Note. $^{*}p < .05$. $^{**}p < .01$. Years in workforce: 10 years or less $n = 95$, 11-20 years $n = 118$, over 20 years $n = 90$.

Participants who had been in the workforce for 11 to 20 years had the most favorable perceptions of their work climate, while those with ten years or less workforce experience had the least favorable perceptions ($\overline{X}_{(1)}$= 3.86, SD = .95; $\overline{X}_{(2)}$= 3.94, SD = 1.10, *d* = 0.36). However, perceptions of rates of unplanned lateness and absence of critical staff had a somewhat different pattern in that participants who had been in the workforce for over 20 years had the most favorable perceptions while those with ten years or less workforce experience had the least favorable perceptions ($\overline{X}_{(1)}$= 4.50, SD = 1.05, $\overline{X}_{(2)}$= 3.94, SD = 1.25, *d* = .51). Similarly, participants who had been in the workforce for over 20 twenty years had the most favorable perceptions of performance management results while those who had been in the workforce for ten years or less had the least favorable perceptions ($\overline{X}_{(1)}$= 4.38, SD = .80; $\overline{X}_{(2)}$= 3.98, SD = .97, *d* = .47). Figure 36 shows the effect of workforce experience on perception of work climate and withdrawal of critical staff, illustrating the larger effect size on Climate.

Figure 36

Survey 2 Statistically Significant Effect of Workforce Experience



Note. $N = 303$. 10 years or less $n = 95$, 11-20 years $n = 118$, over 20 years $n = 90$. CL = Climate, CW = Critical staff Withdrawal, PoR = Perception of Results.

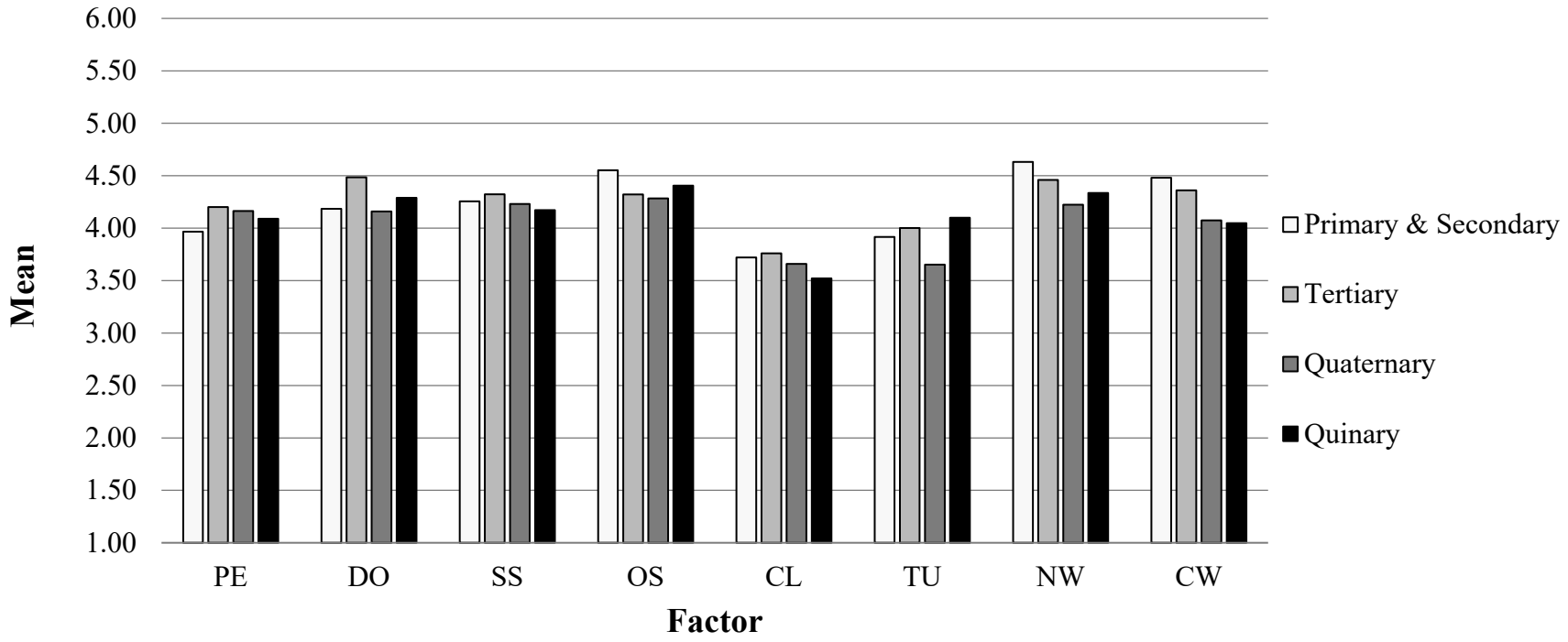
Industry Sector. The industry in which a participant was working at the time of the survey was coded into one of four economic sectors. Thirty-eight participants (13%) worked in Primary and Secondary industries; 119 participants (39%) worked in Tertiary industries, 88 participants (29%) worked in Quaternary industries, and the remaining 58 participants (19%) worked in Quinary industries.

Multivariate tests of the eight lower order factors, $\Lambda = .89$, *Mult. F* (24, 847.49) = 1.52, $p = .05$, showed that the industry in which a respondent worked had a statistically significant impact on their responses. The univariate ANOVAs indicated that these overall differences were more likely related to the responses regarding perception of turnover rates ($F = 2.53$, $df = 3$, $p = .06$) and withdrawal behavior of critical staff ($F = 2.29$, $df = 3$, $p = .08$). Though the p values are close to .05 it would be prudent to interpret these results with caution. The multivariate results of the higher order factors showed that industry sector had no statistically significant effects on responses ($\Lambda = .97$, *Mult. F* (6, 596) = 1.37, $p = .22$).

Participants who worked in quinary industries had the most favorable perceptions of turnover rates at their organization, while those in quaternary industries had the least favorable perceptions ($\overline{X}_{(1)} = 4.10$, $SD = 1.08$; $\overline{X}_{(2)} = 3.65$, $SD = 1.15$, $d = .42$). Participants who worked in primary and secondary industries had the most favorable perceptions of the rates of unplanned lateness or absence of critical staff, while those in quinary industries had the least favorable perceptions ($\overline{X}_{(1)} = 4.48$, $SD = 1.11$; $\overline{X}_{(2)} = 4.05$, $SD = 1.04$, $d = .39$). Figure 37 shows the effect of industry sector on perception of turnover rates and withdrawal behavior of critical staff.

Figure 37

Survey 2 Statistically Significant Effects of Industry Sector on the Lower Order Factor Model



Note. Primary & secondary n = 38, tertiary n = 119, quaternary n = 88, quinary n = 58. PE = Performance Evaluation, DO = Documentation, SS = Supervisory Support, OS = Organizational Support, CL = Climate, TU = Turnover, NW = New staff Withdrawal, CW = Critical staff Withdrawal.

Non-Statistically Significant Effects. Multivariate tests of the eight lower order factors and the two higher order factors showed that tenure, employment status, location, and gender had no statistically significant effect on participant responses as shown in Table 14 and Table 15.

Table 14

Survey 2 Non-Statistically Significant Effects on the Eight Lower Order Factors

Independent Variable	Λ	*F* (df1, df2)	*p*
Tenure	.96	1.37 (8, 294)	.21
Employment status	.97	1.29 (8, 294)	.25
Location	.89	1.48 (24, 847)	.07
Gender	.97	1 (8, 294)	.44

Note. Λ = Wilk's multivariate test criterion, *p* = exact *p* value, *F* (df1, df2) = *F* distribution with hypothesis degrees of freedom and error degrees of freedom.

Table 15

Survey 2 Non-Statistically Significant Effects on the Two Higher Order Factors

Independent Variable	Λ	*F* (df1, df2)	*p*
Tenure	.98	2.29 (2, 300)	.10
Employment status	.99	1.14 (2, 300)	.32
Location	.96	2.04 (2, 596)	.06
Gender	.99	1.93(2, 300)	.15

Note. Λ = Wilk's multivariate test criterion, *p* = exact *p* value, *F* (df1, df2) = *F* distribution with Hypothesis degrees of freedom and Error degrees of freedom.

4. Chapter Summary

This chapter described the results of the second online survey in which the eight factor structure of the EMI scale, developed in the first online survey was verified. The items were further reduced from 55 to more manageable and defensible set of 41items. Estimates of reliability using Cronbach's alpha calculations were high enough for all eight factors, indicating good internal consistency, to confidently proceed with further analysis.

Per the recommendations of the first online survey, a higher order factor model was investigated and resulted in two higher order factors. The first, Perception of Effort, subsumed the following lower order factors; Performance Evaluation, Documentation, Supervisory Support, Organizational Support, and Climate. The second higher order factor, Perception of Results, subsumed the following lower order factors; Turnover, Critical staff Withdrawal, and New staff Withdrawal. The placement of the Climate factor within the higher order factors was debated because its items, theoretically, related to performance management effort or results. For example, while the item "At my workplace, the communication climate is dominated by complaints" can be indicative of genuine dissatisfaction with organizational communication efforts, it can also refer to negative co-worker habits that impacts the general mood in the workplace. In this way, the Climate factor could, theoretically, be placed under the Perception of Results higher order factor as it relates to outcomes of performance management activities. However, the confirmatory factor analysis strongly supported this factor's placement with Perception of Efforts.

The MANOVA showed that a participant's work role, workforce experience, age, industry sector and the size of their organization were statistically significant independent variables that effected respondents' perceptions, and showed some trends consistent with the

MANOVA results from the first survey. It also showed interaction effects when age and gender with considered and when age and employment status were considered together, despite gender and employment status having no statistically significant effect when considered alone.

Some of the MANOVA results showed dissimilarities between the higher order factors and lower order factors, which are not uncommon for two reasons. Because lower order models are usually more discriminatory than higher order models so, in cases where independent variables show a significant effect on only one of several first order factors, the effect is likely neutralized by the other factors when they are grouped into their higher order structure. In cases where an independent variable showed a statistically significant effect on the higher order factor model but not the lower order factor model this was most likely due to the higher explanatory power that higher order models have over lower order models (Musek, 2017). The next chapter discusses these findings, their contributions and presents several recommendations for future research.

Chapter 7: Discussion

Organizations often involve complex and, at times, disorderly dynamics both between and within their departments. They are influenced by changes to their internal environment, like staff changes, as well as their external environment, like political shifts (Soldan & Nankervis, 2014). In particular, technological advancements, changes in consumer demands and global labor migration continue to force organizations to adapt in order to succeed. In addition, the economic change in developed countries from a dominance of manufacturing to knowledge services cemented the value of human resources to organizational performance. As such, it became necessary for HRM, in particular, performance management to become an integral and strategic organizational operation.

Research has shown that, despite the call for performance management to be incorporated in strategic organizational functions, the perceived value of performance management is not advancing at the rate it should because of issues in determining its effectiveness (Handa & Garima, 2014). To elaborate, performance management, unlike other organizational functions such as marketing or finance, does not have the sufficient analytics nor metrics for accurate evaluation of its effectiveness (Fink & Sturman, 2017). Hence, data from staff performance reviews are often used as the barometer of performance management effectiveness which is a flawed approach (Baker, 2013; Boris & Kopczynski Winkler, 2013; Boudreau & Ramstad, 2007; Carlson & Kavanagh, 2018). This problem saw the genesis of the evidence-based management movement, which requires the use of reliable academic and evaluation data to inform HRM decisions.

This chapter recaps the scope of the thesis and the research stages, followed by a discussion of the significant contributions made to scholarly work as well as evidence-based performance management. Following which, priority future directions are discussed and concluding comments are offered.

1. Research Problem and Thesis Scope

Performance management started in the late 1980's in response to negative reactions towards control-based HRM using merit rating systems and management-by-objectives (Armstrong, 2015; Armstrong & Taylor, 2014). It offered a much-welcomed attitude of continuous improvement and has become a prominent feature in organizational management discourse (Sachane et al., 2018). In essence, through performance management, organizations aim to advance overall performance by improving employee performance in a sustainable way.

In spite of its promises, the success of performance management remains highly debatable. Research has shown that many employees do not believe that their organization's performance management improves their performance, hence it often receives the lowest ratings in staff surveys (Sachane et al., 2018). It is unsurprising that performance management is considered as the proverbial 'Achilles' heel' of organizational operations because, although perceptions of its success are poor, organizations have a legal obligation to use it even if only to justify reward allocation or disciplinary actions (Armstrong, 2015; Falcone & Tan, 2013). This partial implementation of performance management is often seen in the form of the annual performance review.

By and large, performance management can fail in two ways; design and implementation. Design errors can occur when performance management policies are so prescriptive that they turn what is meant to be a fluid process into a bureaucratic nightmare. Implementation errors can flow from such design errors, however, the two are not intrinsically linked. As Armstrong (2014) observed, even well designed policies can be poorly implemented if they are misunderstood, or there is a negative perception of their purpose, value, or feasibility.

There is little guarantee that organizational policies will be interpreted as intended because people tend to assign their own meaning to a policy in terms of its purpose, practicality and the ways in which to implement it (Al Shirawi, 2015; Guest, 2002; Ohemeng et al., 2018). Thus, performance management is prone to a process of reciprocal adaptation, in which the objectives of the designers are re-interpreted to a degree by the implementers (Guest, 2002). This can be problematic if there are significant perceptual discrepancies, as they can lead to a detrimental gap between what is officially mandated and what is done. As Nishii and Wright (2008) concisely suggested, “the effect of HR practices is not likely to be automatic and always as expected; instead, their effect will reside in the meanings that employees attach to those practices” (p. 504). The authors highlighted an important reality of perception, which is that most reactions to an external stimulus, like a conversation with a supervisor, are far more contingent on subjective perception of the stimulus than objective reality. Therefore, the implementers’ perceptions, and subsequent translation, of performance management policies are equally important as the quality of the policies themselves (Purcell & Hutchinson, 2007). Thus, it is not enough to have comprehensive performance management policies; one must also consider the way in which the policies are delivered to staff, especially the implementers. In essence, the content of organizational communication not only has to be accurate, it must also be convincing.

As Gary and Laidlaw (2004) found, implementers must trust that they and their staff will derive value proportionate to the effort needed to implement their organization's performance management policies.

Increasingly, the concept of evidence-based management has been suggested as the best way forward to improve performance management and help resolve issues of partial implementation. As proponents of evidence-based management have stated, to improve the effectiveness of performance management one needs to use reliable evidence from four sources; expertise, external stakeholders, academic research and evaluation (Barends et al., 2014; Barends et al., 2012; Briner & Barends, 2016). Despite wide agreement on its benefits, evidence-based practice is not the norm in many organizations (Rousseau, 2012). Several authors have suggested this to be due to the different domains in which practitioners and academics work, as well as the limited use of evaluation data (Bartunek, 2007; Bartunek & Rynes, 2014; Cascio & Boudreau, 2014). Issues of time restrictions to conduct a search for academic research, the abstract nature of scholarly publications, and inaccessibility to those publications have contributed to the limited use of research findings in practice (Werner, 2012). Moreover, despite the call for evaluation in performance management, issues of measurement, evaluation skills, and the costs associated with quality evaluation are proving to be formidable obstacles to the use of evaluation in evidence-based management (Armstrong & Baron, 1998; Lawler, 2007). As a result, performance managers tend to rely on their experiential knowledge, collegial advice and the internet; unfortunately, this limits the reliability of the evidence that informs practitioners' decisions (Bezzina et al., 2017; Stetler et al., 2014).

Given that effective performance management is vital for organizational performance, the success of which is contingent on its implementation, understanding the implementers'

perceptions towards performance management would contribute greatly to evidence-based practice. This research aimed to narrow the gap in the literature by using a threefold lens in its inquiry into performance management perception; the *situation*, the *perceiver*, and the *target* (Robbins et al., 2016; Robbins et al., 2013). For the purposes of this thesis, the perceivers were the implementers / performance managers, the overall target was performance management, and the situation was the organizational context in which perception occurs.

2. Research Summary

The ultimate goal of this thesis was to contribute to evidence-base performance management by bridging the gap between practice, evaluation, and academic research. In doing so, a theoretically grounded and empirically validated scale, called EMI, was created to measure performance managers' perceptions of performance management for organizations to use, irrespective of their evaluation capacity, alongside other existing organizational metrics. Guided by an empiricist epistemology, theory of change and utilization-focused evaluation, three sequential stages were carried out, illustrated in Figure 38, to address the following research questions:

RQ 1: What perceptions of performance management should be measured in evaluation efforts for evidence-based performance management?

RQ1a: "what perceiver characteristics influence perception of performance management and to what extent?"

RQ1b: "what situation characteristics influence perception of performance management and to what extent?"

RQ 2: How can these perceptions be measured in a valid and reliable measurement scale?

With respect to the second research question, the validity in question was construct validity and reliability referred to internal consistency. To address the research questions, three sequential research stages were done, namely; item generation, scale development and, scale refinement and validation. The empirical study in Stage 1 was a scoping literature review, whereas in Stages 2 and 3, online surveys were done recruiting from an international pool of participants. The target population consisted of staff who manage the performance of other workers. This includes team leaders, junior managers, middle managers and senior managers. The achieved samples were representative of the target population specifically in that only those who manage staff performance were invited to participate. The recruitment process explicated this requirement so, the demographic items in each survey did not require further screening of participants.

Stage 1: Item Generation. This stage encompassed findings from the narrative literature review (Chapter 2) to conceptualize the perception targets, and the results from the scoping literature review (Chapter 4) to generate items for the perception targets. Performance management is innately a complex organizational intervention and any inquiry on best practices must consider the outcomes as well as the context in which the outcomes were observed for the findings to be useful and accurate (Berwick, 2008; Kovner, 2012). When evaluating complex interventions, evaluators often use a theory of change and a logic model so, to navigate the literature, a theory of change and a logic model were used for deductive thematic analysis. The model (see Figure 8, Chapter 2) showed nine perception targets relating to overall performance management perception, namely perception of; propriety, feasibility, utility, accuracy, evidence-based practice, organizational communication, organizational support, work engagement, and others' behavior. The first four perception targets were from the Joint Committee's evaluation

standards, while the remaining five were from various seminal pieces of literature relating to perception and employee behavior (Barends et al., 2014; Downs & Hazen, 1977; Eisenberger et al., 1986; Joint Committee on Standards for Educational Evaluation, 2008; Nagadevara et al., 2008; Schaufeli et al., 2002). This stage provided the theoretical basis for answering both research questions, after which the data from Stages 2 and 3 provided empirical support.

Following the conceptual identification of the perception targets, the first empirical study, shown in Figure 38, used a scoping literature review to identify relevant and psychometrically validated scale items that could be used for the EMI scale. In cases where an insufficient number of items were found, new items were generated guided by DeVellis (2017) and Gable and Wolf (1993). As a result, the initial EMI scale contained 130 items on a six-point Likert scale to measure nine perception targets.

For the purposes of data analysis during the second and third stage of this research, the perception targets were considered as the dependent variables, while characteristics of the perceiver and the situation were the independent variables i.e., antecedents. It was paramount to achieve a balance between parsimony (to avoid superfluous items) and conceptual plausibility (to meaningfully account for data patterns), As such, item reduction was based on statistical analysis and conceptual congruity.

Stage 2: Scale Development. The aim of this stage was to develop the EMI scale based on empirical data as well as explore the antecedent factors. As Figure 38 shows, an online survey, using 130 scale items, was conducted with 286 international participants (performance managers). Though the sample size was less than initially desired, the similarities between the sample characteristics and the target population, as well as results from the KMO test and Bartlett's test of Sphericity provided enough confidence to proceed with further analysis. To

develop the scale's structure and content, an exploratory factor analysis was done, followed by Cronbach's alpha calculations to ascertain the scale's estimates of reliability in terms of internal consistency. After several iterations of the factor analysis, 75 items were deleted. The analyses supported an eight factor solution that proposed the perception targets to be; Performance Evaluation, Documentation, Organizational Support, Supervisory Support, Climate, Turnover, Critical staff Withdrawal and New staff Withdrawal. Finally, to explore characteristics of the perceiver and the situation, a MANOVA was done using six antecedent factors, only three of which were statistically significant; location, organization size, and work role.

Stages 3: Scale Refinement and Validation. The purpose of this stage was to refine and validate the EMI scale based on new empirical data, and determine the statistically significant antecedent factors. Shown in Figure 38 shows, a second online survey, using 55 scale items, was conducted with 303 international participants (performance managers). Despite the less than ideal sample size, the consistency in the demographic characteristics between the sample and the target population provided additional assurance that the sample was representative of the target population. To verify the EMI scale's structure and content, a confirmatory factor analysis was used, deleting 14 items, and supported the eight factor model as well as a two higher order factor model. Though the model fit indices supported the fit of both models, the higher order model was considered to fit better from a conceptual plausibility perspective. Cronbach's alpha calculations revealed sufficient estimates of internal consistency to confidently proceed with further analysis. Because this survey had a significantly smaller number of items than the first survey, more demographic items were used to determine the statistical significance of perceiver and situation characteristics as antecedent variables. A MANOVA, using nine categorical

variables, revealed seven statistically significant antecedents; age, work role, workforce experience, organization size, industry sector, age and gender, employment status and gender.

The final EMI scale evolved from its conceptual form to an empirically validated scale in a few notable ways as shown in Figure 26 in Chapter 5. After an assessment of conceptual congruity as well as the statistical results, two higher order factors encompassing eight first order factors formed the final EMI scale. The first higher order factor, Perception of Efforts, encompassed perception of Performance Evaluation (eleven items), Documentation (four items), Supervisory Support (four items), Organizational Support (four items), and Climate (five items). The second higher order factor, Perception of Results, encompassed perception of Turnover (five items), Critical staff Withdrawal (four items) and, New staff Withdrawal (four items).

The scale contained a defensible set of 41 items, 14 of which were from the original Joint Committee standards (propriety – three items, feasibility – two items, utility – four items and accuracy – five items). The remaining 27 items were from four of the five remaining perception targets; organizational support (nine items), organizational communication (three items), work engagement (two items) and others' behavior regarding withdrawal (thirteen items). No items from evidence-based practice were included in the final EMI scale as discussed in Section 3 of this chapter.

Figure 38

Summary of Research Stages and Results

	Stage One – Qualitative	**Stage Two – Quantitative**	**Stage Three – Quantitative**
Name	Item Generation	Scale Development	Scale Refinement & Validation
Aims	• Identify perception targets* • Generate items	• Determine perception targets • Refine items • Explore significant antecedents	• Verify perception targets • Refine items • Determine significant antecedents
Data Collection	Scoping literature review	Online Survey 1	Online Survey 2
Data Analysis	Thematic analysis	Exploratory factor analysis Estimates of reliability Multivariate analysis of variance	Confirmatory factor analysis Estimates of reliability Multivariate analysis of variance
Software	Online databases MS Word MS Excel	Prolific Academic LinkedIn Qualtrics MS Excel SPSS 24	Prolific Academic LinkedIn Qualtrics MS Excel SPSS 24 AMOS 24
Results	120 publications Nine perception targets 130 items	286 performance managers Eight factors (perception targets) 55 items 3 significant antecedents	303 performance managers Two higher order factors (perception targets) Eight lower order factors (perception targets) 41 items 7 significant antecedents

Note. * The purpose of stage one included identifying perception targets and generating items for them. However, it was through the narrative literature review that the perception targets were identified, and through scoping study that the items were generated.

3. Substantive Contributions

This thesis presented three substantive contributions. The first was the empirical determination of performance management perception targets that may affect the connection between performance management policies and implementation. The second was its contribution to narrowing the gap in instrumental relevance of performance management research by developing a theoretically-based and empirically-tested scale for organizations to use in evaluation for evidence-based performance management. The third contribution was the enrichment of the extant knowledge on the antecedents of perception relating to characteristics of the perceiver and the situation in which perception occurs. The following provides a detailed discussion of these substantive contributions.

3.1 *Performance Management Perception Targets*

Central to the thesis' contributions is the conclusion that perception of performance management is not a unidimensional construct, in fact, it is composed of several perception targets. As mentioned in Section 2 of this chapter, the results from the first stage of this research presented nine perception targets which were conceptually categorized into four components of a logic model (see Figure 8, Chapter 2); inputs, activities, outputs and outcomes of performance management. In essence, the results from Stage 1 suggested that perception of performance management contains four higher order constructs (inputs, activities, outputs, and outcomes), which, in turn, encompassed nine lower order perception targets. The empirical data from the second and third stage of this research supported the notion that perception of performance management as a multidimensional construct but contains two higher order constructs (factors), which, in turn, encompass eight lower order perception targets (factors).

The first higher order factor, Perception of Efforts, related to perceptions of the organization's investment and practices in performance management. It subsumed five lower order factors; the first related to perception of Performance Evaluation in terms of the accuracy, propriety, feasibility, and utility of the organization's performance management efforts. The second lower order factor related the perception of Documentation in terms of the clarity, accessibility, and strategic alignment of employee performance information. The third lower order factor related to the perception of Supervisory Support in terms of fair treatment and feedback one gets from their supervisor. The fourth lower order factor related to the perception of Organizational Support in terms of like job security and fair treatment. The fifth lower order factor related to the perception of Climate in terms of the communication culture and work opportunities in one's work environment.

The second higher order factor, Perception of Results, related to perception of outcomes of the organization's performance management efforts, in particular the withdrawal behaviors of other employees within the perceiver's work environment. This factor subsumed three lower order factors; the first was perception of Turnover in terms of the general rates of others' turnover. The second was perception of New staff Withdrawal relating to the rates of unplanned absence and unplanned lateness of staff new to the organization and recent graduates. The third was perception of Critical staff Withdrawal relating to the unplanned lateness and unplanned absence of permanent and tenured staff.

A comparison of the logic model used in the literature review and the final structure of the EMI scale showed important similarities. First, the logic model from Stage 1 proposed the higher order level contain perception of performance management inputs, activities, outputs, and outcomes. The empirical data, on the other hand, suggested the higher order level contain

perception of performance management efforts and results, which are, arguably, an amalgamation of the aforementioned logic model categories. Second, the theorized relationship between the higher order levels in the first logic model was supported, albeit only in terms of correlation. Specifically, Perception of Efforts was positively correlated with Perception of Results (r = .38).

A comparison of the nine conceptual performance management perception targets from the literature review and the eight factors from the survey studies showed five noteworthy changes. The first was seen upon examining perception of other's behavior, which included items that described perceptions of others' performance and the three withdrawal behaviors; turnover, unplanned lateness, and unplanned absence. However, items regarding the performance of others were not in the final EMI scale due to their indefensible statistical weakness. This finding supported the literature on the significant effects of negative behavior compared to positive behavior of other staff on one's perceptions (Basford & Offermann, 2012; Hung et al., 2009; Robinson et al., 2014). In particular, Robinson et al. (2014) noted that people are more likely affected by observed negative behavior than positive behavior because negative experiences are, psychologically, more powerful than positive ones. This is an important finding because, as the literature suggested, poor behaviour of others not only distances the perceiver from the 'offender' but the organization itself as it is seen to be the source of justice (Chan & Lai, 2017; Hung et al., 2009; Hutchinson et al., 2009; Robinson et al., 2014). So, if undesirable behaviour is perceived to be high, it is likely that the perceiver will lose trust in the organization.

The second noteworthy change was seen in the transformation of perception of turnover, unplanned lateness and unplanned absence from one sub-construct (withdrawal) to three individual latent factors. Moreover, unlike the perception of Turnover factor, unplanned lateness and absence emerged as two factors relating to specific types of staff; new staff and critical staff. As the literature suggested, it is known that undesirable behavior of certain workers can be rationalized by the perceiver to be outside the control of the organization, which may reduce the likelihood that the perceiver will view the organization negatively (Basford & Offermann, 2012; Eder & Eisenberger, 2008; Hung et al., 2009; Robinson et al., 2014; Zhou & George, 2001). For example, unplanned lateness of colleagues with young families may be rationalized by the perceiver to be due to the often-unpredictable demands of young children rather than a failure of performance management. Likewise, withdrawal behavior of staff new to the workforce may be rationalized by the perceiver to be due to the workers' unrealistic expectations of the workplace.

The emergence of New staff Withdrawal and Critical staff Withdrawal as perception targets is a strong contribution because it allows for specific organizational intervention. In particular, the withdrawal of critical staff like tenured staff and those on a permanent contract, should be a red flag for organizations because critical staff usually require significant organizational investment and their withdrawal can be more costly (Tracey, 2010; Tracey & Hinkin, 2008; Walker, 2017; Whetstone, 2017).

The third distinction between the conceptual targets and the empirical factors arose when the original organizational support construct from Eisenberger et al., (1986) emerged as two factors; Supervisory Support and Organizational Support. This may be a somewhat divisive finding because it has been well documented that supervisors are often seen as organizational agents, therefore Organizational Support and Supervisory Support may in fact be components of

a single construct (Eisenberger et al., 1986; Kottke & Sharafinski, 1988). More recently, Alfes, Shantz, et al. (2013) found that the combination of experience with the perception of the organization one's supervisor is more associated with employee behavior than either perception target on its own. Their suggestion for researching perceptions of the general organizational climate and the supervisor may support the argument for one construct, however, it could equally support the presence of two constructs that must be targeted systematically. The findings of this thesis support the proposition that the extent to which employees see their supervisor as an agent of the organization plays an important role in perceptions of support to such a degree that support constitutes two constructs (Allen et al., 2003; Eisenberger et al., 2002; Purcell & Hutchinson, 2007; Rhoades & Eisenberger, 2002).

The fourth important change from the conceptual to the empirical stages of this research was noted upon the emergence of perception of Climate as a stand-alone factor, rather than as a sub-construct of organizational communication. The items for perception of Climate stemmed from perception of organizational communication as well as organizational support and related to evaluative perceptions of feedback, work opportunities and informal communication within one's work environment. This finding supports the notion that the effect of the intangible work atmosphere, despite measurement difficulties, is worthy of investigation because of the insights it can provide (Teclaw et al., 2014). For example, the item "When it comes to opportunities at work favoritism plays a big role" can explain staff complaining, poor team work, disengagement with role promotion opportunities and intention to quit. Moreover, the item "Creative solutions & new ideas are not welcomed at work" can explain the behavior of staff who, for instance, do not contribute to team discussion or take initiative to solve problems.

The fifth notable change was the dissipation of the original evidence-based practice construct. Despite the many iterations of the exploratory factor analysis in Stage 2 (see Chapter 5), there was little justification for retaining any of its eleven items because of their extremely poor factor loadings, across multiple factors and conceptual weakness compared to other similar items from the other original constructs. For example, the item "Our performance management system is not regularly evaluated to ensure its continued usefulness & positive impact on staff performance" loaded below .30 across multiple factors. It was also conceptually repetitive with the Joint Committee's accuracy standard item "A system is in place to allow the periodic review of the performance evaluation system to ensure its continued usefulness", which had excellent factor loading (.83) on one factor (Performance Evaluation). This finding questions the original assumption that evidence-based practice is a unidimensional construct, rather it is likely that all the EMI items were underpinned by evidence-based practice.

3.2 *The EMI Scale*

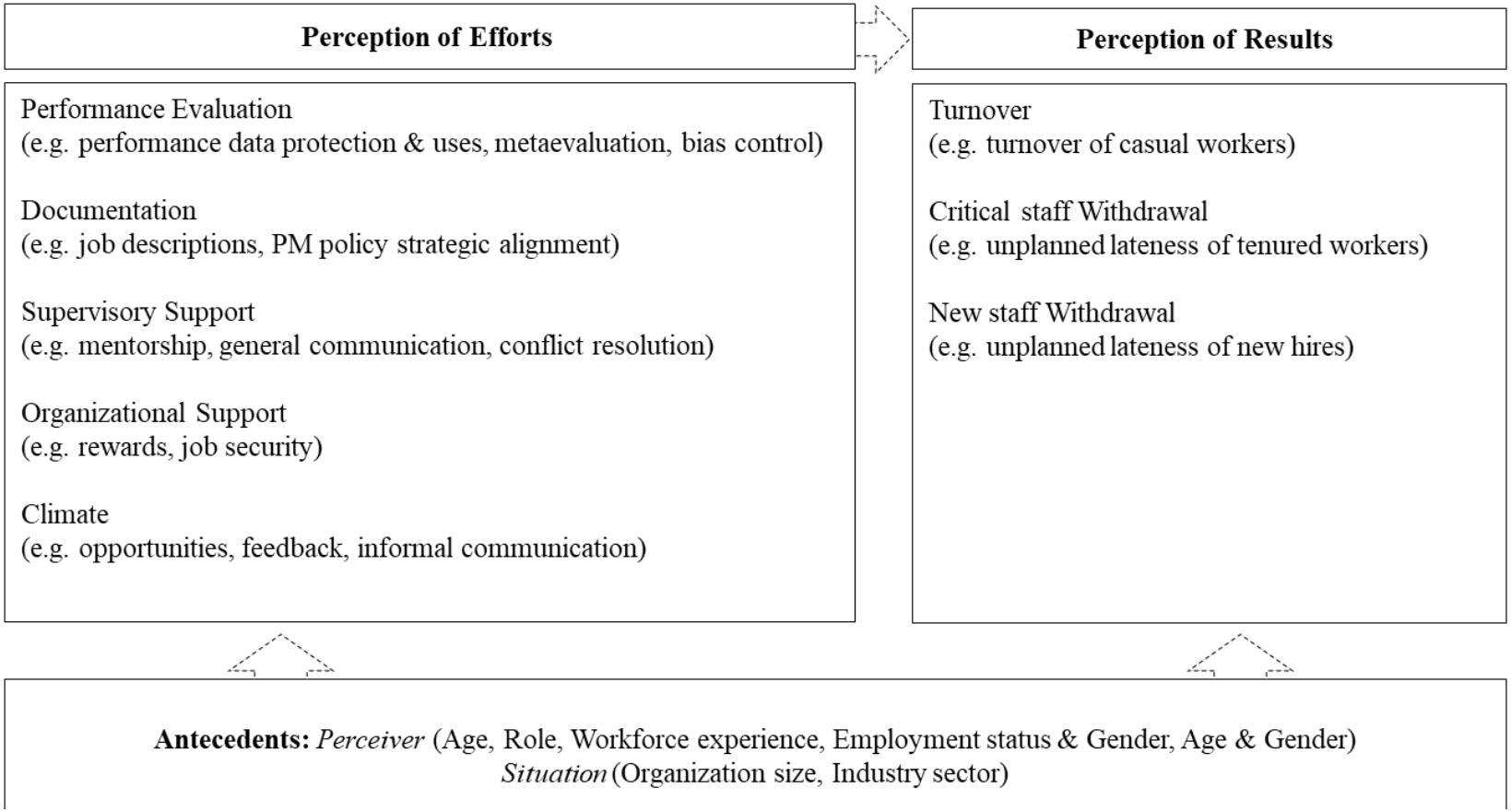
As Nicolai and Seidl (2010) and Wang et al. (2020) suggested, issues of instrumental relevance have not been, historically, central in management research. However, a demand for specific solutions for performance management problems is on the rise. The EMI scale was developed to resolve issues in measuring perceptions to enhance evaluation for evidence-based performance management. In particular, the scale's strengths stem from its theoretical basis and empirically validated multidimensionality.

Theoretical Basis. The research presented in this thesis was guided by a theory of change, commonly used by program evaluators, and utilization-focused evaluation. As discussed in Chapter 2 Section 4.3, a theory of change explains how performance management is expected to achieve its desired goals. In essence, using a theory of change allows an evaluator to discover

not only what was achieved by the intervention, but also why and how it was achieved (DuBow & Litzler, 2019; Rogers & Weiss, 2007; Weiss, 1997a, 1997b). Performance management, like many initiatives with ambitious goals for behavior change, is inherently complex and difficult to evaluate, making it all the more necessary to carefully construct and use a theory of change, which is usually accompanied by an intervention-specific logic model (DuBow & Litzler, 2019). By following the logic of a theory of change to navigate the literature, the initial item pool for the EMI scale had a strong theoretical basis, and the interpretation of the latent factors that emerged in the exploratory factor analysis was meaningful. Figure 39 shows the empirically informed logic model displaying the hierarchy of the perception targets (latent factors), as well as the newly hypothesized relationships between the perception targets, characteristics of the perceiver and the situation (antecedent factors). It must be noted that, typically, a logic model has more components than the one presented in Figure 39, like an intervention's need, aims, activities, outputs, and outcomes to enable evaluation stakeholders to carefully monitor their progress. However, this logic model was simple as its purpose was to illustrate the research findings coherently. Moreover, logic models require as much detail as possible in order to add precision to any theorized causal links, indicated by arrows from one component to the next. Though this research did not test for causality, the following logic model contains arrows, however, as shown in Figure 39, they have a dotted outline to further indicate that they are correlation links not causal links.

Figure 39

Simplified Logic Model of the Empirically Tested Factors



Note. Dotted arrows indicate correlation relationships / effects for examination in future research on causal relationships.

This thesis was also guided by utilization-focused evaluation because the limited use of evaluation data in evidence-based performance management was highlighted many times over in the literature (Bourgeois et al., 2013; Cousins et al., 2014). For instance, a recurring theme in the evaluation literature concerns the partial or misuse of evaluation, and issues of organizational capacity to do and use evaluation (Cousins et al., 2008). The goal of utilization-focused evaluation is centered on the intended uses by the intended users, in this case, the use of the EMI scale by organisations to evaluate performance managers' perceptions likely to influence their implementation behavior and enhance their evidence-based practice (Ridde et al., 2005). Patton (2000) contended that even the most technically sound evaluations are useless if they end with an evaluation report sitting on a shelf collecting dust. Therefore, every step of a utilization-focused evaluation is focused on how people in their real-life settings would use the evaluation's findings to lead to desired outcomes (Ridde et al., 2005). Hence, the development and refinement of the EMI scale was heavily influenced by considerations of utility. The scale needed to be understandable, easy to use, and easily incorporated with existing organizational metrics. As a result, the EMI scale was designed to include items not typically included in common organizational metrics (discussed in Chapter 2) to avoid redundancy

Multidimensionality and Item Structure. In addition to a theoretical basis and empirical testing, the EMI scale was designed to measure multiple targets of perception. Though other instruments like the UWES may be considered superior to the EMI scale as it can delve deeper into its construct (work engagement), the aim of this thesis and the EMI scale, in particular, was to allow organizations to identify areas of priority for further investigation, much like a triage process. The Joint Committee's list of key questions also enables a triage process; however the questions have not be quantitatively validated. Therefore, this thesis extended the

committee's work by psychometrically validating the EMI scale items, which contained many items from the Joint Committee's list of key questions.

Another level of dimensionality stemmed from Bowen and Ostroff's (2004) work on HRM strength, in which they argued for content and process as categories for investigation. Wang et al. (2020) expanded Bowen and Ostroff's (2004) seminal work by identifying a third consideration; HRM motivation (they 'why' as they called it). This referred to employee perceptions regarding the reasons and organizational motivation behind HRM policies and processes (Wang et al., 2020). This thesis built on Wang et al.'s (2020) contribution, by adding a fourth component; HRM outcome. The final EMI scale contained five items on content and motivation, four items on content and process, two items on process and motivation, 14 items on outcomes, and 16 items on process. Building further on Wang et al.'s (2020) notes on the orientation of scale items and how they can affect survey responses, the final EMI scale items were varied in their orientation to allow for as accurate a measurement as possible. Four items were descriptive observational, ten items were evaluative experiential items, and 27 items were evaluative observational.

3.3 *Antecedents*

As previously stated, perception can be affected by characteristics of the perceiver, the target and the situation. While researchers have recognized that perceptions influence behavior (Liao et al., 2009; Nishii & Wright, 2008), scholars know much less about the characteristics of the perceiver and the situation and the extent to which they can influence one's perception of a target (K. Jiang et al., 2017). As Hewett et al. (2018) recently showed, there is a rich body of research about the consequences of employee perception but not about its antecedents or potential drivers. Wang et al.'s (2020) literature review also found a disproportionate amount of research favoring the consequences of employee perception over the antecedents. Based on their findings, Wang et al. (2020) urged future research to identify antecedents to HRM perceptions, especially those that are under the discretion of the organization, like communication structures (the situation).

Furthermore, as previous research has proposed, there is significant variability in management perceptions among performance managers of all levels within the organizational hierarchy, even within the same organization (Bartram et al., 2007; Ostroff & Bowen, 2016). Therefore, identifying antecedents pertaining to the perceiver is also important because, though they may not be at the discretion of the organization, they provide insight that can help organizations to target their efforts, efficiently, to improve performance management perceptions. Therefore, in addition to the identification of perception targets related to performance management, this thesis contributed to narrowing the gap in the literature regarding instrumental relevance by exploring and determining the statistical significance of seven antecedents.

It has been long accepted that perceiver characteristics like age, gender, and work experience play an important role in perception and subsequent behavior (Bellou, 2010; Hofstede, 1998). However, the research on which factors have an effect and the statistical significance of their effect remains mixed. For example, the multivariate tests employed in this research showed that tenure had no statistically significant impact on perceptions; a finding supported by Teclaw et al. (2014) but contradicted by Lundmark and Biorkman (2011) and Ogunkuade (2018). Given the variety of industries, sample sizes and cultures between these studies and this research, it is not certain if the inconsistencies in the findings are due to theoretical or measurement issues.

It is imperative to note that, as stated in Chapter 3, one of the purposes of the first survey was to explore perception antecedents, whereas in the second survey the purpose was to determine the statistically significant antecedents. This meant that demographic items were not as prioritized in the first survey as they were in the second survey, which resulted in the former having fewer items and missing data. Therefore, the following provides a discussion of the educated assumptions explaining the MANOVA findings from second survey with some reference to the results from the first survey where applicable.

3.4 *Perceiver Characteristics*

The findings showed that a perceiver's age, role and workforce experience had statistically significant effects on their perceptions, whilst their gender, employment status, and tenure did not. However, when gender was considered with employment status or age, a statistically significant effect emerge.

Gender. As shown in Chapter 6, there was no statistically significant effect of participant gender on their perceptions. This finding supports the work of Burke et al. (2015), who found little gender variance on perceptions of leadership, some work outcomes, and learning opportunities. The authors noted that the few gender differences were most likely because the men in their study occupied roles within higher levels of organizational hierarchy and had been with their employer for longer than the sample of women. It must be noted that their study sample came from the textile and furniture sector in Turkey, and that the number of females was much smaller than the male sample (215 male and 46 female). Whereas this thesis, the samples were from multiple countries, industries and organization sizes. This could indicate that the effect of gender on perception is not as prone to contextual and sample variation as other antecedent factors.

By contrast, Juris Bennett et al. (2015) found significant gender differences in perceptions of information technology among 623 healthcare providers in Saudi Arabia. Their finding must be taken with consideration of the workforce setting in Saudi Arabia, and as the authors noted, females might have different perceptions because of their recent inclusion in the healthcare workforce. Therefore, the difference could be due to variable workforce experience rather than gender.

A study by Marinaccio et al. (2013) included 4,123 participants in Italy found that women, especially those in full-time roles, had more deleterious perceptions of factors related to work stress than men, or women in part-time roles. The authors supposed that this could be due to the gender divide present in the Italian culture. In particular, the pressure that women feel to balance their careers with their family responsibilities may have reduced their resilience and augmented their perception of job stressors (Carrier, 1995; Nair et al., 2017; Travis et al., 2016).

Indeed, the contradiction between the thesis' findings and Marinaccio et al.'s (2013) may stem from differences in cultural context and sample sizes. However, given that the gender divide in social norms is also apparent in countries like turkey, it is more likely that the contradictory findings are due to the level of specificity of the constructs and their scale items. To elaborate, Marinaccio et al.'s (2013) study used the Italian version of the Health and Safety Executive Management Standards Indicator Tool, which contained many generic items like "I have unrealistic time pressures". The perception measured by such items could refer to factors outside the workplace like family obligations rather than job-specific stressors. This is not to claim that their study items were unsuitable; on the contrary, because their topic was the perception of work stress risk, which is a broad concept compared to perception of performance management, one could argue that their broad items were quite suitable. Nevertheless, this means that their results on the effect of gender are not comparable to the results from this thesis. The EMI scale, despite having a few broad items, lends itself more to specific phenomena as most of its items refer to unambiguous experiences or observations of real-life performance management policies, practices, and outcomes.

Work Role. The influence of work role is well noted in the literature on performance management, HRM, and employee behaviour (Kang & Busser, 2018; Op de Beeck et al., 2016; Singh, 2000; van Dijke et al., 2018). However, the influence of organizational hierarchy on perceptions remains relatively uncharted, specifically, in terms of how authority level influences perception and subsequent performance (Kang & Busser, 2018). In the first survey (Chapter 5), work role was grouped into HR staff and team leaders, while in the second survey, it was categorized into four groups according to their hierarchical levels within the organization; administration, junior management, middle management and senior management. The

multivariate tests from both surveys showed that role had a statistically significant effect on responses despite the difference in sample size, coding, and number of items.

As discussed in Chapter 6, the multivariate tests of the second survey data showed some statistically significant effects of work role. The differences were particularly significant in perceptions of performance management results regarding the withdrawal behavior of other workers, specifically in terms of unplanned lateness and unplanned absence of critical staff and new staff. Participants in middle management had the most favorable perceptions of performance management results while those in senior management had the least favorable perceptions. These findings supported previous assertions by Corley (2004) and Kang and Busser (2018) that staff within each hierarchical organizational level have divergent perceptions. These differences are generally driven by the day-to-day tasks associated with each role, which affect the employee's understanding of events in the work place (Buss & Kuyvenhofen, 2011; Özleblebici & Çetin, 2015). They can also be caused by the varying degrees of autonomy, responsibilities, and access to organizational data. For instance, senior managers tend to have broader perspectives of organizational operations and access to data that middle managers do not. The latter tend to have access to information about their team or their department, hence their perspective is more localized than that of senior managers.

The EMI items within Critical staff Withdrawal and New staff Withdrawal required respondents to state their level of agreement with statements claiming that the respective behavior of others is 'high' or 'frequent', however, a definition for these terms was not provided. Therefore, the response variance could be due to different meanings that senior managers and middle managers associate with the term 'high' or 'frequent'. It is likely that senior managers' perceptions may be influenced by their overarching perspective of organizational functioning,

compared to the more localized perspective of middle managers. Therefore, middle managers might only know about their immediate team's withdrawal behavior or only have access to their own department's performance data (Birken et al., 2015; Buss & Kuyvenhofen, 2011; Olson et al., 2005; Whalley, 2000). While senior managers may have access to data from more if not all departments as well as the financial data on the costs of withdrawal behavior that other managers do not. This, in turn, and may also lead to different thresholds for defining 'high' or 'frequent' rates of withdrawal between senior managers and middle managers.

Age. Participant age was categorized into four age ranges; 18 to 25, 26 to 35, 36 to 45 and, over 45 years of age. Multivariate tests (Chapter 6) showed that participant age had a statistically significant impact on their perception of Performance Evaluation and the Critical staff Withdrawal, but not on Perception of Efforts or Results. Dissimilarities between higher order factors and lower order factors are not uncommon. Lower order models are usually more discriminatory than higher order models so, in cases where independent variables show a significant effect on only one of several first order factors, the effect is likely neutralized by the other factors when they are grouped into their higher order structure (Brown, 2015; Franck, 2002).

Participants aged between 36 and 45 years had the most favorable perceptions of Performance Evaluation in their organizations, while those over the age of 45 had the least favorable perceptions. This could be due to a difference in career stages and associated levels of engagement with performance management (Chaudhary & Rangnekar, 2017). To elaorate, those in the early to middle part of their careers tend to be invested in career development and, thus, actively enagage with performance management processes (Malik & Subramanian, 2015; Vlastarakos & Nikolopoulos, 2008). Other (older) employees may not engage as much because

they see less relevance to their work or career (Wang et al., 2020). In addition, organizations may invest more in workers who are in early to mid-career stages as they have greater workforce longevity compared to older workers (Kooij & Boon, 2018; Mainiero & Gibson, 2017; Twenge, 2010). Therefore, for example, the Performance Evaluation item "All the performance managers have received appropriate training in the evaluation process" may have been poorly rated by older works because they may not have received as much training as others. This finding is supported by Juris Bennett et al. (2015), who found that tenured (older) staff had lower perceptions of their organization's information technology because the organization prioritized training for new staff.

By contrast, participants over the age of 45 had the most favorable perceptions of the rates of unplanned lateness and unplanned absence of critical staff, while those aged between 18 and 25 years had the least favorable perceptions. This could be due to generational differences in item interpretation. For example, older workers who have more experience in the workforce may have experienced higher rates of others' withdrawal in other organizations, whereas the younger workers have fewer points of comparison. That said, this finding was still surprising because the literature on generational differences strongly suggested that Millennials tend to value work flexibility, while Baby Boomers tend to value adherence to traditional work hours (Chen & Choi, 2008; Christopher et al., 2018; Cucina et al., 2018; Zopiatis et al., 2012).

Age and Gender. Despite gender having no statistically significant effects on responses, the multivariate tests in Chapter 6 showed a statistically significant interaction effect when both age and gender were considered, specifically, on Perception of Results. For both men and women, the perceptions within the oldest participants on performance management results were more favorable than the perception within the youngest participants. This indicates the

possibility that perception gaps between men and women are greater among Millennials and Baby Boomers. Whereas, men and women who were predominantly from Generation X, the difference in perception was far smaller.

The significance of the interaction effect is supported by other studies like Morris et al. (2005). However, Morris et al. (2005) found that the pattern of gender differences varied with age such that the variance was more distinct with increasing age, which is not completely consistent with the results from Survey 2. Moreover, there was a notable inverted pattern between the extreme age groups and the mid-range age groups in that, women from the mid-range age group had more favorable perceptions than men, but in either extreme age groups, it was the men who had the more favorable perceptions. This could be because the research from Morris et al. (2005) focused specifically on technology adoption rather than HRM perceptions, or because their research was more than a decade older than this thesis, the demographic dynamics of the populations are no longer comparable. For example, according to the UK Office for National Statistics, in 2018, the average age of first-time parenthood rose for the tenth consecutive year in to 30.6 (mothers) and 33.6 years (fathers) (Office for National Statistics, 2019). As a result, the female participants in Survey 2, who had the more favorable perceptions of results, fall into the category of new mothers (26-35) and mothers with young children (36-45). Therefore, it is possible that their tolerance for the withdrawal of others is higher than that of women under 25 or over 45 years of age, especially if the behavior is observed among peers with the same family commitments as their own. It could also be that their interpretation of the words 'high' or 'frequent' in the relevant scale items is of a higher threshold than the other age groups. Male participants of the same age groups (25-35, 36-45) had less favorable perceptions of others' withdrawal which could be because they have the added pressure of providing for their young

families, especially if their partners had decreased their work hours to look after their young children. As such, they may be more sensitive to unplanned withdrawal of co-workers and assign different meanings to the words ‘high’ or ‘frequent’.

Workforce Experience. Categorized by the number of years in the workforce, workforce experience showed a statistically significant effect on participant responses, especially in Perception of Results, and two lower order factors; Climate and Critical staff Withdrawal. Participants were grouped into three categories; ten years of experience or less, 11 to 20 years, and over 20 years in the workforce. In particular, participants who had been in the workforce for 11 to 20 years had the most favorable perceptions of their work climate, while those with ten years or less workforce experience had the least favorable perceptions.

Moreover, those who had been in the workforce for over 20 years had the most favorable perceptions of Perception of Results, especially in Critical staff Withdrawal, while those with ten years or less workforce experience had the least favorable perceptions. These differences could be due to the same reasons for the different perceptions between the various age groups, in addition to the different perceptions developed through the variety of experiences and workplaces during their years in the workforce.

Tenure. As perception can be influenced by learning, peer socialization, and work experiences, all of which are affected by time, it is possible that new staff and tenured staff would have different perceptions (Graham & Nafukho, 2007). In particular, it was expected that tenured staff would have less favorable perceptions of performance management efforts than new staff because organizations typically invest more into the new cohort of workers (Alnıaçık et al., 2012; Juris Bennett et al., 2015; Teclaw et al., 2014). However, the thesis’ findings showed no statistically significant effects of tenure on perception. This supports Rhoades and

Eisenberger's (2002) suggestion that dissatisfied workers tend to leave their organizations, therefore tenured staff typically have more stable positive perceptions.

Contradictions were found with other studies like Graham & Nafukho (2007), who investigated employee perception of culture associated with organizational learning readiness. The authors found that tenure (longevity, as they called it) had a significant effect of perception of culture associated with organizational learning readiness; specifically that newer employees had more positive perceptions. The contradiction in findings is most likely due to the fact that Graham & Nafukho's (2007) study involved 150 employees from one small manufacturing organization in the United States, compared to the 303 participants in Survey 2 (Chapter 6) who were from various countries, organizational sizes and industries.

Employment Status. The literature indicated that the nature of one's employment status (precarious vs secure) can have an effect on employee perceptions (Kim et al., 2012). In this thesis, precarious employment included part-time, fixed term, and casual work, while secure employment included full-time and permanent work. Although the two can overlap, for example, a worker can be employed full-time but on a fixed-term contract, the main distinction between the two categories was in the level of income security. Therefore, even part-time workers on a permanent contract can experience income security issues due to the lower pay compared to that of permanent full-time employment. Further, depending on a country's work legislation, precarious employment tends to have far fewer benefits than secure employment like training opportunities, annual leave, and sick leave. The multivariate results from both surveys showed that employment status had no statistically significant effects on responses. This contradicts other research like Marinaccio (2013) who found that workers in fixed-term employment had favorable perceptions of their work demands and relationships. The contradiction is most likely

because their findings related to perceptions of the risk of work-related stress, which is a construct with a higher level of abstraction than performance management.

Nonetheless, in Survey 2, the multivariate test results showed a statistically significant interaction effect when gender was also considered. As noted in Chapter 6, due to the small number of men in precarious employment within the study sample, these results must be interpreted with caution. Female participants who worked in precarious employment at the time of the survey had the most favorable perceptions of performance management documentation like their job descriptions, while males, also in precarious employment, had the least favorable perceptions. This could be because women occupy part-time roles more than men, usually due to childcare needs, hence they may be more engaged with the relevant documentation (Workplace Gender Equality Agency, 2016, 2018). Moreover, as research from the United Kingdom found, men are more likely to negotiate their starting salary whereas women are more likely to ask for changes in working hours, so they may be more engaged with aspects of their job descriptions that go beyond their pay (McGovern et al., 2007). Though the wage-gender gap is prominent in the literature, items about perceptions of pay were excluded from this thesis, because its aim was to gage perceptions about performance management relating to more intrinsic than extrinsic incentives for performance.

3.5 *Situation Characteristics*

The findings revealed that organizational size had a statistically significant effect on perception, while industry sector had some statistical significance and location had none.

Organization Size. The size of a respondent's organization was categorized into four groups according to the number of staff. In both survey studies, the size of the organization had a statistically significant effect on responses. Survey 2 multivariate tests of the eight lower first order factors showed that the size of a respondent's organization had a statistically significant effect on their responses. This effect was most notable in perception of Performance Evaluation, Climate, Turnover and Perception of Results.

Participants who worked in small organizations at the time of the survey had the most favorable perceptions of their work climate while those in very large organizations had the least favorable perceptions. This could be because small organizations are able to engage their employees more easily and at a lower cost than their larger counterparts (Tracey & Hinkin, 2008; Witt et al., 2000). For example, personal and informal communication as well as familiar social events and decentralized decision-making are prevalent in small organizations (Clampitt & Girard, 1993; Karanges et al., 2015). Whereas larger organizations tend to have formal channels of communication, and centralized decision-making, which, as Schminke et al. (2000) suggested, can negatively influence employee perceptions.

Similarly, those in small organizations had the most favorable perceptions of turnover rates at their organization, while those in very large organizations had the least favorable perceptions. A similar pattern was noted in Perception of Results; those in small organizations had the most favorable perceptions, while those in very large organizations had the least favorable perceptions. This could be because early intervention to reduce turnover is easier in

small organizations. To elaborate, in small organizations, it may be easier to notice early signs of turnover due to the smaller number of staff. Another explanation is that, as Tracey and Hinkin (2008) suggested, though larger organizations can afford greater work opportunities and income stability, they also have more job demands, which can add to employee stress levels and increase their likelihood of leaving.

Industry Sector. The industry in which a participant was working at the time of the survey was coded into one of four economic sectors. Primary and secondary industries included mining or manufacturing, respectively; tertiary industries included retail or finance, quaternary industries included research, and, quinary industries included government services.

Multivariate tests of the eight lower order factors showed that the industry in which a respondent worked had a statistically significant impact on their responses, however, the multivariate results of the higher order factors showed that industry sector had no statistically significant effects on responses. Furthermore, the univariate ANOVAs of the eight lower order factors showed that these overall differences were more likely related to the responses regarding Turnover and Critical staff Withdrawal which, due to the high p value (p = .06 and .08, respectively) must be interpreted with caution.

Participants who worked in quinary industries had the most favorable perceptions of turnover rates at their organization, while those in quaternary industries had the least favorable perceptions. This finding supports the hypothesis that there is higher rates of turnover in quaternary industries due to the nature of the work contracts inherent in the industry. For example, quaternary sector roles include consultancy, research, entertainment and media; all of which have a culture of freelance, fixed term and casual employment which, in turn, have lower levels of income security than permanent contract roles (Vosko et al., 2009). The finding also

supports other research like that from the United States Bureau of Statistics, which found that the median job tenure in 2018 for roles in film and sound recording was only 2.9 years (U.S. Bureau of Labor Statistics, 2018). Meanwhile, utilities organizations, like electrical power or natural gas, had a 9.5 year median job tenure which was much higher than any other industry sector in 2018 (U.S. Bureau of Labor Statistics, 2018). Moreover, those in the quinary industry, like those in federal government, had a median tenure of 8.3 years. In fact, quaternary industry roles had much lower median tenure in 2018 than quinary industry roles (U.S. Bureau of Labor Statistics, 2018).

By contrast, participants who worked in primary and secondary industries had the most favorable perceptions of the rates of unplanned lateness or absence of critical staff, while those in quinary industries had the least favorable perceptions. This could be due to the higher levels of job autonomy and flexible work conditions in quinary industries compared to primary and secondary industries. Bredehöft et al. (2015) noted that flexibility in task completion is especially notable in jobs that require highly skilled knowledge workers. In essence, such flexibility results in the proverbial ‘double-edged sword’ of job autonomy, in that workers may work offsite but not necessarily report their whereabouts to anyone but their direct supervisor. This can be perceived, by others, as unplanned lateness or absence.

The difference in withdrawal behavior perceptions between quinary and primary and secondary industry sectors could also be due to the level of complexity of the associated work roles. To elaborate, organizations may dedicate fewer resources to fill positions that require cognitively simple tasks, like mining or manufacturing, than complex roles like those in government or scientific research. Though this may not apply to a performance manager’s role, the EMI scale items referred to the behavior of others but did not specify the work roles of the

other employees in question, so the responses could have referred to other performance managers or other employees in general.

Task complexity may explain the variance in perception between the industry sectors because staff required for complex roles offer organizations, in knowledge industries, a more competitive edge while demanding competitive salaries and more training on the specific organizational operations associated with their roles (Huselid et al., 2005). So, while overall turnover may be higher for simpler job positions, the financial implications may not be as great. As a consequence, organizations tend to invest more in attaining and retaining talented staff for strategically complex roles (Tracey & Hinkin, 2008). In fact, Tracey & Hinkin (2008) posited that organizations may need to spend double the amount of money and time to replace a competent worker in a complex role than one who has mastered all aspects of a simpler work role.

4. Future Directions

This section provides a synthesis of the work undertaken in this thesis to recommend how it can be extended through future research to enhance its findings.

Scale Structure. The literature posited that employees tend to have two dominant relationships at work; one with their direct supervisor, and the other with the organization as a whole (Gok et al., 2015; Maertz et al., 2007). The distinction between these relationships is, arguably, situational, and previous research has focused on the two relationships as components of one construct; organizational support. This is because employees may consider it a supervisor's responsibility to provide support because they are agents of the organization (Cole et al., 2006; Katsaros et al., 2014). It reasonably follows that employees would view a supervisor's support as an indicator organizational support, however, this is highly contingent on

the extent to which a supervisor is perceived to be an agent of the organization (Eisenberger et al., 1986; Kottke & Sharafinski, 1988). The thesis findings supported the theory that supervisory support is a construct, separate but, positively correlated with organizational support. Further investigation is needed into how employees see their supervisor's relationship with the organization to better understand their responses to the Organizational Support and Supervisory Support items in the EMI scale. Such research would help to substantiate the theory that when employees see their supervisor as distant from the organization then supervisory support and organizational support are two separate constructs.

Further, to better understand results of the EMI scale, future research, in conjunction with the EMI scale, should survey employees using items regarding perceptions of their own skills and abilities. Because, as Katsaros et al. (2014) further argued, one's perception of their own knowledge, abilities and skills plays a role in their performance. This perception can be influenced by the information given to the worker by the organization, professional development, but also by their own personal characteristics (Katsaros et al., 2014).

Antecedents. The statistical significance of the characteristics of the perceivers and the situation in which perception occurs requires further investigation. For instance, the effect of industry sector on perceptions was found to be questionable in this thesis as it was inconsistent between Survey 1 and Survey 2. Future research should include more detailed demographic items to examine the characteristics of each industry sector, which may provide further insight into how this might affect perceptions.

Both survey studies presented in this thesis recruited performance managers who could have been involved in the design as well as implementation of performance management. Since involvement can range from periodic consultation to extensive participation and influence, future

research should gage the level of involvement of the performance managers in the design or redesign of performance management policies to further understand the effect of job tasks on their perception of performance management.

Farndale and Sanders (2017) and Sanders et al. (2008) offered that employee perceptions of HRM process can vary specifically due to the cultural and national surroundings of the perceiver. Though this research did not delve deeply into the effect of the perceiver's location and upbringing, the survey findings showed generational differences in perception which could be due to age, cultural upbringing or both. For that reason, it is decidedly recommended for future research to include items that explore the effect of the perceiver's cultural environment on their perception of performance management.

Validity. There are several recommendations for future research to enhance the validity of the thesis' results. First, the scoping review study was limited due to having one reviewer so, some literature may have been missed and there was the possibility of selection bias. However, this can be the case with scoping reviews because the researcher(s) are typically unfamiliar with the extant literature and may not be able to apply the focused criteria common in systematic reviews. Indeed, scoping reviews can be useful precursors to systematic reviews (Munn et al., 2018) so, future research should build on the findings of this study by conducting a meta-analysis or systematic review.

Second, aside from correlation, the findings did not offer evidence for causal relationships between the antecedent factors and perception targets because each survey study collected data at one point in time (cross-sectional). Hence, future research must use longitudinal study designs to substantiate causal relationships. Longitudinal research is vital because self-report data are vulnerable to the respondent's state of mind and mood at the time of a survey. In

particular, a respondent may have forgotten pertinent information, and/or their mood at the time of the survey may have impacted their perceptions (positively or negatively). Because this research was completely anonymous and participants were untraceable, it cannot be determined if the participants of the first survey also undertook the second survey. Therefore, future longitudinal research should collect data from the same participants at different time intervals to account for variance caused by to transient mood.

Third, there is a possibility of common method bias that can lead to the inflation or deflation of correlations because both surveys used self-report items (Batista-Foguet et al., 2014; Jordan & Troth, 2020; Reio, 2010). This was, perhaps, unavoidable because the research topic was on individual employee perceptions for which self-reporting may be the most valid form of measurement (Alfes, Shantz, et al., 2013; Alfes, Truss, et al., 2013; Marinaccio et al., 2013; Ostroff & Bowen, 2016). However, common method variance stemming from social desirability bias was unlikely in this research due to the high degree of anonymity so, participants were less likely influenced by the urge to been seen in a favorable light. Therefore, future research should adhere to anonymity but also collect supporting data like organizational performance data and analyze performance management policy documents.

The characteristics of the scale items could have also led to common method variance. As Podsakoff et al. (2009), Reio (2010), and Worthington and Whittaker (2006) stated, abstract or general items can be difficult to answer and are more likely to lead to bias. This source of bias was not a great concern in this thesis because, although the EMI scale contained some general items, the vast majority were designed to be as specific as possible, and variable in their orientation (descriptive, evaluative, experiential and observational). Therefore, future research, if it is to add more items to the EMI scale, should consider striking a balance of item orientation.

The scope of this thesis did not extend to the assessment of actual implementation of performance management and organizational performance for two reasons. First, the identification of performance management perception targets that may influence implementation was lacking in the literature and, therefore, considered a research priority. Second, the assessment of real-life performance management systems, practices, and organizational outcomes requires longitudinal research ideally and / or case study research within several organizations. Both of which were not feasible for this thesis, but are highly recommended for future research. In particular, new research is needed to ascertain the EMI scale's capacity to predict performance management implementation by using it along with measures of implementation behavior (performance data) and organizational outcomes. Furthermore, during times of organizational uncertainty, like internal restructuring or market fluctuations, employees need to perceive the support they get from their organization positively more than ever because this can affect how they perceive the organization's situation, their job security, and their role in implementing suggested changes (Katsaros et al., 2014). Future research on the EMI scale should include items about the organization's internal and external environments to better understand employee perceptions and implementation behavior.

In sum, along with the EMI scale, the main direction for future research on the validity of the EMI scale centers on longitudinal studies in a variety of organizations, countries, and industries, with a focus on determining causal links and potential mediation and moderation effects of perception, the perceiver and the situation on implementation of performance management. To do so, future research should, along with the EMI scale, collect data on the objective reality of performance management within an organization. This could include reviewing HRM policy documents, staff performance data, and organizational performance data.

Intended Use. The EMI scale is intended to help organizations in triaging performance managers' perceptions for early intervention to improve performance management implementation. The nature of the intervention was outside the scope of this thesis so, future developments of the EMI scale could benefit from developing a repository for suitable interventions based on the scale's results. For example, for medical researchers and practitioners, the Cochrane database gives information about diagnoses as well as field-tested alternative treatments from which to choose (van Aken & Romme, 2012). Likewise, the EMI scale could provide a 'diagnosis' of the perceptions of performance managers, but to effectively advance evidence-based performance management it should be paired with a resource like the Cochrane database.

Intended Users. The utility of the EMI scale needs to be tested in real-life setting to ascertain its capacity to work with existing organizational metrics, its feasibility and the utility of its data for the intended users. This should be tested in organizations with variable degrees of evaluation capacity to further support the thesis' claim that the EMI scale's utility for organizations is irrespective of their evaluation capacity to do or use evaluation.

Involving primary intended users in developing or reviewing the theory of change is also highly recommended for future research. As the literature review suggested, in practice, initiatives like performance management, rarely have an explicated theory so, an evaluator must first identify or construct the theory. As Donaldson and Lipsey (2006) speculated, when useful information already exists in the literature, ignoring it can be a very inefficient way to evaluate. However, while drawing on existing theories in the social sciences is beneficial, evaluation experts like Weiss advocated for stakeholder involvement in the construction of the theory as they are the ones who understand the context (Wildschut, 2014). As such, further investigations

are needed, specifically with input from performance managers to enhance the theoretical basis of the EMI scale.

Performance Evaluation. The evaluation standards from the Joint Committee, have been an integral part of evaluation discourse and practice for decades (Howard & Harman, 2007; Joint Committee on Standards for Educational Evaluation, 2008; Loup & Ellett, 1997). There are, however, a few limitations of the standards that have caught the eyes of critics. Perhaps the strongest limitation of the standards and, in particular, the list of key questions, is their psychometric validity. The Joint Committee, with the help of a validation panel, was as transparent as possible in the development and validation of the standards, however, the validation process referred to face and content validity only, and despite the mention of field tests, the nature of those tests is unknown (Linn et al., 1989). Nonetheless, the items adapted from the list of key questions from the Joint Committee were invaluable to the EMI scale despite the lack of previous quantitative testing. After the factor analyses, the factor that accounted for most variance was the Performance Evaluation factor, which had items only from the Joint Committee. This indicated that the Joint Committee's (2008) list of key questions have psychometric qualities that, at least, merit further investigation. As such, future research on the construct validity and estimates of reliability of the list of key questions would, more likely than not, result in a validated scale for performance evaluation based on evaluation expertise and scientific research, which would be a substantial contribution to the evaluation discipline.

5. Concluding Comments

Changes in organizational performance have called for commensurate changes in HRM and performance management in order to recruit and retain talented staff. It is important to remember that unlike other resources, human resources are fundamentally complex and their performance can be affected by a multitude of intangible forces, some of which, like family dynamics, are beyond organizational control. Nonetheless, performance management can impact employee behavior to such a degree that a push for evidence-based practice became prevalent in the literature. Specifically, the literature identified that evidence-based practice is essential to the effectiveness of performance management however, it is challenged by the limited use of evidence from academic research or quality evaluation.

The issue of effectiveness has been linked to poor implementation of performance management, which by many accounts is influenced by the implementers' perception of performance management. Performance managers, especially those in the lower levels of an organization's chain of command, play an essential role in performance management implementation because their daily interactions with other staff can create a positive or negative experience with performance management, which can influence its effectiveness in managing staff behavior and impacting organizational performance (De Winne et al., 2013; Frenkel & Bednall, 2016; Frenkel, Restubog, et al., 2012; Purcell & Hutchinson, 2007). Therefore, in addition to existing organizational metrics, this thesis strongly suggested that any evaluation or research for evidence-based performance management should include the measurement of the performance managers' perceptions.

This thesis presented new knowledge about the three components of perception in the performance management context; the perceiver, the situation and the target. It showed that

performance management as a target of perception contains multiple levels; eight lower order factors and two higher order factors. Furthermore, seven characteristics of the perceiver and the situation were found to influence perception to a statistically significant degree. The thesis also presented a new measurement scale, called EMI, which was based on theory, empirical data, as well as lessons from Patton's (2000) utilization-focused evaluation. The new EMI scale, was designed to work alongside existing performance management metrics and frameworks like the balanced scorecard, thereby narrowing the extensive gap between research and practice (Dulebohn & Johnson, 2013). Overall, the findings provided a substantial foundation for future longitudinal research that especially aims to ascertain the causal relationships between perception and performance management effectiveness.

Irrespective of the nature of future inquiry into performance management, it is vital to remember that performance management is, innately, a complex organizational intervention and any research on evidence-based practice must consider the outcomes as well as the context in which the outcomes were observed (Ferris et al., 2008; Kovner, 2012). It is envisaged that the EMI scale will help organizations to enhance their evidence-based practice in performance management by providing data, not commonly collected, in a feasible, understandable, and accurate way early enough for meaningful intervention to improve the implementation of performance management.

Reference List

Abma, T. A. (2005). Responsive evaluation: Its meaning and special contribution to health promotion. *Evaluation and Program Planning, 28*, 279-289. https://doi.org/10.1016/j.evalprogplan.2005.04.003

Abu Bakar, H., & Mustaffa, C. S. (2013). Organizational communication in Malaysia organizations: Incorporating cultural values in communication scale. *Corporate Communications, 18*, 87-109. https://doi.org/10.1108/13563281311294146

Academy of Human Resource Development. *About AHRD*. https://www.ahrd.org/page/about_ahrd_2

Adams, J. S. (1965). Inequity in social exchange. *Advances In Experimental Social Psychology, 2*, 267-299.

Ahmady, S., Changiz, T., Brommels, M., Gaffney, F. A., Thor, J., & Masiello, I. (2009). Contextual adaptation of the personnel evaluation atandards for assessing faculty evaluation systems in developing countries: The case of Iran. *BMC Medical Education, 9*, 18. https://doi.org/10.1186/1472-6920-9-18

Akan, B., Er Ülker, F., & Ünsar, A. S. (2016). The effect of organizational communication towards resistance to change: A case study in banking sector. *Economic Review: Journal of Economics & Business / Ekonomska Revija: Casopis za Ekonomiju i Biznis, 14*, 53-67.

Akdol, B., & Arikboga, F. S. (2015). The effects of leader behavior on job satisfaction: A research on technology Fast50 Turkey companies. *Procedia - Social and Behavioral Sciences, 195*, 278-282.

Al Shirawi, T. M. (2015). *Strategy implementation: Exploring roles, perceptions, and expectations of middle managers' practices* [PhD, Brunel University].

Alfes, K., Shantz, A. D., Truss, C., & Soane, E. C. (2013). The link between perceived human resource management practices, engagement and employee behaviour: A moderated mediation model. *International Journal of Human Resource Management, 24*(2), 330-351. https://doi.org/10.1080/09585192.2012.679950

Alfes, K., Truss, C., Soane, E. C., Rees, C., & Gatenby, M. (2013). The relationship between line manager behavior, perceived HRM practices, and individual performance: Examining the mediating role of engagement. *Human Resource Management, 52*(6), 839. https://doi.org/10.1002/hrm.21512

Alkin, M. C. (Ed.). (2013). *Evaluation roots: A wider perspective of theorists' views and influences* (2nd ed.). Sage Publications.

Allen, D. G., Shore, L. M., & Griffeth, R. W. (2003, 1/1/2003). The role of perceived organizational support and supportive human resource practices in the turnover process. *Journal of Management, 29*, 99-118. https://doi.org/10.1016/S0149-2063(02)00222-2

Alliger, G. M., & Janak, E. A. (1989). Kirkpatrick's levels of training criteria: Thirty years later. *Personnel Psychology, 42*(2), 331-342.

Alnıaçık, Ü., Alnıaçık, E., Akçin, K., & Erat, S. (2012). Relationships between career motivation, affective commitment and job satisfaction. *Procedia - Social and Behavioral Sciences, 58*, 355-362. https://doi.org/https://doi.org/10.1016/j.sbspro.2012.09.1011

Alvarez, J. L., & Mazza, C. (1998). Haute couture or pret-a-porter: Creating and diffusing management practices through the popular press. *IESE Research Papers, D/368*. http://ideas.repec.org/p/ebg/iesewp/d-0368.html

American Evaluation Association. (n.d.). *Guiding principles for evaluators*. https://www.eval.org/p/cm/ld/fid=51

Andersson Arntén, A.-C., Algafoor, N. A., Al Nima, A., Schütz, E., Archer, T., & Garcia, D. (2016). Police personnel affective profiles: Differences in perceptions of the work climate and motivation. *Journal of Police and Criminal Psychology, 31*(1), 2-14. https://doi.org/10.1007/s11896-015-9166-5

Ang, S. H., Bartram, T., McNeil, N., Leggat, S. G., & Stanton, P. (2013). The effects of high-performance work systems on hospital employees' work attitudes and intention to leave: A multi-level and occupational group analysis. *International Journal of Human Resource Management, 24*(16), 3086-3114. https://doi.org/10.1080/09585192.2013.775029

Ang, S. H., Cavanagh, J., Southcombe, A., Bartram, T., Marjoribanks, T., & McNeil, N. (2017). Human resource management, social connectedness and health and well-being of older and retired men: The role of men's sheds. *International Journal of Human Resource Management, 28*(14), 1986-2016. https://doi.org/10.1080/09585192.2015.1088886

Appelbaum, E., Bailey, T., Berg, P., & Kalleberg, A. L. (2000). *Manufacturing advantage: Why high performance work systems pay off*. ILR Press.

Ariely, D., Kamenica, E., & Prelec, D. (2008). Man's search for meaning: The case of legos. *Journal of Economic Behavior & Organization, 67*, 6.

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *Int J Soc Res Methodol, 8*, 19-32. https://doi.org/10.1080/1364557032000119616

Armstrong, M. (1996). *A handbook of personnel management practice* (6th ed.). Kogan Page.

Armstrong, M. (2015). *Armstrong's handbook of performance management: An evidence-based guide to delivering high performance.* (5th ed.). Kogan Page. https://doi.org/10.1097/PRS.0b013e31822adba3

Armstrong, M., & Baron, A. (1998). *Performance management: The new realities*. London: Institute of Personnel and Development.

Armstrong, M., & Baron, A. (2005). *Managing performance: Performance management in action* (1st ed.). Chartered Institute of Personnel and Development.

Armstrong, M., & Taylor, S. (2014). *Armstrong's handbook of human resource management practice.* (13th ed.). Kogan Page.

Aromataris, E., & Munn, Z. (Eds.). (2019). *Joanna Briggs institute reviewer's manual*. The Joanna Briggs Institute. https://reviewersmanual.joannabriggs.org/

Atwater, L. E., & Brett, J. F. (2006). 360-Degree feedback to leaders: Does it relate to changes in employee attitudes? *Group & Organization Management, 31*(5), 578-600. https://doi.org/10.1177/1059601106286887

Auerbach, M. P. (2019). Economic sectors. In *Salem Press Encyclopedia* (pp. 7). Great Neck Publishing.

Australian Bureau of Statistics. (2013). *1220.0 - ANZSCO - Australian and New Zealand Standard Classification of Occupations, 2013, Version 1.2* https://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/83CF605BED750F35CA257B95001310EF?opendocument

Australian Human Resources Institute. (2008). *'Dangling the carrot' - Investigating performance management practices* (HRPulse, Issue 2008).

Australian Human Resources Institute. (2019). *Performance management*. https://www.ahri.com.au/resources/ahriassist/performance-management/

Australian Human Resources Institute. (n.d.-a). *AHRI: ASSIST*. https://www.ahri.com.au/resources/ahriassist/

Australian Human Resources Institute. (n.d.-b). *HRM*. Mahlab. https://www.hrmonline.com.au/

Australian Human Rights Commission. (n.d.). *Performance management*. https://www.humanrights.gov.au/quick-guide/12072

Australian Public Service Commission. (2018). *2018 APS employee census data*. Australian Public Service Commission. https://data.gov.au/data/dataset/aps_employee_census_2018

Azzam, T., & Levine, B. (2015). Politics in evaluation: Politically responsive evaluation in high stakes environments. *Evaluation and Program Planning, 53*, 44-56. https://doi.org/http://dx.doi.org/10.1016/j.evalprogplan.2015.07.002

Bailey, T., Berg, P., & Sandy, C. (2001). The effect of high performance work practices on employee earnings in the steel, apparel and medical electronics and imaging industries. *Industrial and Labor Relations Review, 54*(2A), 525-543.

Baker, T. (2013). *The end of the performance review: A new approach to appraising employee performance*. Basingstoke: Palgrave Macmillan.

Bakker, A. B. (2011). An evidence-based model of work engagement. *Current Directions in Psychological Science, 20*, 265-269. https://doi.org/10.1177/0963721411414534

Bakker, A. B., & Demerouti, E. (2007). The job demands-resources model: State of the art. *Journal of Managerial Psychology, 22*, 309-328. https://doi.org/10.1108/02683940710733115

Bakker, A. B., Demerouti, E., & Sanz-Vergel, A. I. (2014). Burnout and work engagement: The JD–R approach. *Annual Review of Organizational Psychology and Organizational Behavior, 1*(1), 389-411. https://doi.org/10.1146/annurev-orgpsych-031413-091235

Bakker, A. B., & Leiter, M. P. (2010). *Work engagement: A handbook of essential theory and research*. Taylor and Francis.

Balducci, C., Fraccaroli, F., & Schaufeli, W. B. (2010). Psychometric properties of the Italian version of the Utrecht Work Engagement Scale (UWES-9): A cross-cultural analysis. *European Journal of Psychological Assessment, 26*(2), 143-149. https://doi.org/10.1027/1015-5759/a000020

Baluch, A. M. (2017). Employee perceptions of HRM and well-being in nonprofit organizations: Unpacking the unintended. *International Journal of Human Resource Management, 28*(14), 1912-1937. https://doi.org/10.1080/09585192.2015.1136672

Baluch, A. M., Salge, T. O., & Piening, E. P. (2013). Untangling the relationship between HRM and hospital performance: The mediating role of attitudinal and behavioural HR outcomes. *International Journal of Human Resource Management, 24*(16), 3038-3061. https://doi.org/10.1080/09585192.2013.775027

Bamford, D., & Hall, C. (2007). A case study into labour turnover within an NHS trust. *Health Services Management Research, 20*, 9-21. https://doi.org/10.1258/095148407779615007

Bandaranayake, D. (2003). Assessing performance management of human resource for health in south-east asian countries aspects of quality and outcome. In W. Van Lerberghe, G. Kegels, & V. De Brouwere (Eds.), *Towards a Global Health Workforce Strategy* (pp. 239-256).

Bangchokdee, S., & Mia, L. (2016). The role of senior managers' use of performance measures in the relationship between decentralization and organizational performance. *Journal of Accounting & Organizational Change, 12*, 129-151.

Barbars, A. (2016). Interaction between organizational culture and work engagement in the information and communication technology sector in Latvia. *Journal of Business Management, 2016*(12), 84-100.

Barends, E., Rousseau, D., & Briner, R. (2014). *Evidence-Based management: The basic principles* 2014). Center for Evidence-Based Management. https://www.cebma.org/wp-content/uploads/Evidence-Based-Practice-The-Basic-Principles.pdf

Barends, E., ten Have, S., & Huisman, F. (2012). Learning from other evidence-based practices: The case of medicine. In D. M. Rousseau (Ed.), *The Oxford Handbook of Evidence-Based Management*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199763986.013.0002

Barends, E., Villanueva, J., Rousseau, D., Briner, R., Jepsen, D., Houghton, E., & ten Have, S. (2017). Managerial attitudes and perceived barriers regarding evidence-based practice: An international survey. *Plos One, 12*(10). https://doi.org/10.1371/journal.pone.0184594

Baritz, L. (1960). *The servants of power: A history of the use of social science in American industry*. Wesleyan University Press.

Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management Studies, 17*(1), 99-120.

Barratt, M. J., Ferris, J. A., & Lenton, S. (2015). Hidden populations, online purposive sampling, and external validity: Taking off the blindfold. *Field Methods, 27*(1), 3-21. https://doi.org/https://doi.org/10.1177/1525822X14526838

Bartlett, D. (2011). The neglect of the political: An alternative evidence-based practice for I–O psychology. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 4*, 27–31.

Bartram, T., Stanton, P., & Leggat, S. (2007). Lost in translation: Exploring the link between HRM and performance in healthcare. *Human Resource Management Journal, 17*(1), 21-41.

Bartunek, J. M. (2007). Academic-practitioner collaboration need not require joint or relevant research: Toward a relational scholarship of integration. *Academy of Management Journal, 50*(6), 1323.

Bartunek, J. M., & Rynes, S. L. (2014). Academics and practitioners are alike and unlike: The paradoxes of academic–practitioner relationships. *Journal of Management, 40*(5), 1181-1201. https://doi.org/10.1177/0149206314529160

Baruch, Y. (2004). The desert generation: Lessons and implications for the new era of people management. *Personnel Review, 33*(2), 241-256.

Basford, T. E., & Offermann, L. R. (2012). Beyond leadership: The impact of coworker relationships on employee motivation and intent to stay. *Journal of Management and Organization, 18*, 807-817.

Bateman, T. S., & Organ, D. W. (1983). Job Satisfaction and the Good Soldier: The Relationship Between Affect and Employee "Citizenship". *Academy of Management Journal, 26*, 587-595. https://doi.org/10.2307/255908

Batista-Foguet, J. M., Revilla, M., Saris, W. E., Boyatzis, R., & Serlavos, R. (2014). Reassessing the effect of survey characteristics on common method bias in emotional and social intelligence competencies assessment. *Structural Equation Modeling A Multidisciplinary Journal, 21*(4), 596-607. https://doi.org/10.1080/10705511.2014.934767

Baughman, W. A., Dorsey, D. W., & Zarefsky, D. (2011). Putting evidence in its place: A means not an end. *Organizational Psychology: Perspectives on Science and Practice, 4*, 62-64.

Baumeister, R. F. (2016). Toward a general theory of motivation: Problems, challenges, opportunities, and the big picture. *Motivation and Emotion*, 1. https://doi.org/10.1007/s11031-015-9521-y

Baysal, O., Holmes, R., & Godfrey, M. W. (2013). Developer dashboards: The need for qualitative analytics. *IEEE Software, 30*, 46-52. https://doi.org/10.1109/MS.2013.66

Beatty, R. W., Huselid, M. A., & Schneier, C. E. (2003). New HR metrics: Scoring on the business scorecard. *Organizational Dynamics, 32*(2), 107-121.

Becker, B., & Gerhart, B. (1996). The impact of human resource management on organizational performance: Progress and prospects. *Academy of Management Journal, 39*, 779–801.

Bedford, D. S., Malmi, T., & CPA Australia. (2009). Best practice in performance management. Melbourne, Australia.

Bednall, T. C., Sanders, K., & Runhaar, P. (2013). Stimulating informal learning activities through perceptions of performance appraisal quality and human resource management system strength: A two-wave study. *Academy of Management Learning & Education, 13*(1), 45-61. https://doi.org/10.5465/amle.2012.0162

Behson, S. J. (2005). The relative contribution of formal and informal organizational work–family support. *Journal of Vocational Behavior, 66*, 487-500. https://doi.org/10.1016/j.jvb.2004.02.004

Bellou, V. (2010). Organizational culture as a predictor of job satisfaction: The role of gender and age. *Career Development International, 15*(1), 4-19.

Bennett, J. (2003). *Evaluation methods in research*. New York; London: Continuum, 2003.

Bennis, W. G., & O'Toole, J. (2005). How business schools lost their way. *Harvard Business Review, 83*(5), 96–104.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238-246.

Berry, C. M., Lelchook, A. M., & Clark, M. A. (2012). A meta-analysis of the interrelationships between employee lateness, absenteeism, and turnover: Implications for models of withdrawal behavior. *Journal of Organizational Behavior, 33*(5), 678-699. https://doi.org/10.1002/job.778

Berwick, D. (2008). The science of improvement. *Journal of the American Medical Association, 299*, 1182–1184.

Bezzina, F., Cassar, V., Tracz-Krupa, K., Przytuła, S., & Tipurić, D. (2017). Evidence-Based human resource management practices in three EU developing member states: Can managers tell truth from fallacy? *European Management Journal*. https://doi.org/http://dx.doi.org/10.1016/j.emj.2017.02.010

Biddix, J. P. (2009). Instrument, validity, reliability. *Research Rundowns, 2014*.

Birken, S. A., Lee, S.-Y. D., Weiner, B. J., Chin, M. H., Chiu, M., & Schaefer, C. T. (2015). From strategy to action: How top managers' support increases middle managers' commitment to innovation implementation in health care organizations. *Health Care Management Review, 40*, 159.

Biron, M., Farndale, E., & Paauwe, J. (2011). Performance management effectiveness: Lessons from world-leading firms. *International Journal of Human Resource Management, 22*(6), 1294-1311. https://doi.org/10.1080/09585192.2011.559100

Biswas, S., Varma, A., & Ramaswami, A. (2013). Linking distributive and procedural justice to employee engagement through social exchange: A field study in India. *International Journal of Human Resource Management, 24*(8), 1570-1587. https://doi.org/http://dx.doi.org/10.1080/09585192.2012.725072

Bjørnstad, S., Patil, G. G., & Raanaas, R. K. (2016). Nature contact and organizational support during office working hours: Benefits relating to stress reduction, subjective health complaints, and sick leave. *Work, 53*(2016), 9-20. https://doi.org/10.3233/WOR-152211

Blackman, D., Buick, F., O'Donnell, M., O'Flynn, J., & West, D. (2014). *Strengthening the performance framework: Towards a high performing Australian public service*. https://www.apsc.gov.au/sites/default/files/FINAL-strengthening-performance-framework-diagnostic-implementation-18-07-14.pdf

Blackman, D., Buick, F., O'Donnell, M., O'Flynn, J., & West, D. (2012). *Developing high performance: Performance management in the Australian public service* (Crawford School Working Paper No. 12-09). Australian National University. http://ssrn.com/abstract=2130232

Blau, P. M. (1964). *Exchange and power in social life*. Wiley.

Bloom, B. S. (Ed.). (1956). *Taxonomy of educational objectives: The classification of educational goals* (1st ed.). Longman.

Bloom, D. E., Canning, D., & Fink, G. (2011). Implications of population aging for economic growth. *Oxford Review of Economic Policy, 26*(4), 583-612. https://doi.org/10.3386/w16705

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health, 6*. https://doi.org/10.3389/fpubh.2018.00149

Boon, C., Hartog, D. N. D., Boselie, P., & Paauwe, J. (2011). The relationship between perceptions of HR practices and employee outcomes: Examining the role of person-organisation and person-job fit. *International Journal of Human Resource Management, 22*(1), 138-162. https://doi.org/10.1080/09585192.2011.538978

Boris, E. T., & Kopczynski Winkler, M. (2013). The emergence of performance measurement as a complement to evaluation among U.S. foundations. *New Directions for Evaluation, 2013*, 69-80. https://doi.org/10.1002/ev.20047

Bos-Nehles, A. C., & Veenendaal, A. A. R. (2019). Perceptions of HR practices and innovative work behavior: The moderating effect of an innovative climate. *International Journal of Human Resource Management, 30*(18), 2661-2683. https://doi.org/10.1080/09585192.2017.1380680

Boswell, W. R., Olson-Buchanan, J. B., & LePine, M. A. (2004). Relations between stress and work outcomes: The role of felt challenge, job control, and psychological strain. *Journal of Vocational Behaviour, 64*, 165 - 181.

Bothma, C. F. C., & Roodt, G. (2013). The validation of the turnover intention scale. *South African Journal of Human Resource Management, 11*(1) e1-e12 https://doi.org/10.4102/sajhrm.v11i1.507

Boudreau, J. W., & Ramstad, P. M. (2002). *Strategic HRM measurement in the 21st century: From justifying HR to strategic talent leadership* (Working Paper #02-15). Center for Advanced Human Resource Studies. http://digitalcommons.ilr.cornell.edu/cahrswp/56

Boudreau, J. W., & Ramstad, P. M. (2007). *Beyond HR: The new science of human capital*. Harvard Business School Press.

Bourgeois, I., & Cousins, J. B. (2013). Understanding dimensions of organizational evaluation capacity. *American Journal of Evaluation, 34*(3), 299-319. https://doi.org/10.1177/1098214013477235

Bourgeois, I., Toews, E., Whynot, J., & Lamarche, M. K. (2013). Measuring organizational evaluation capacity in the Canadian federal government. *Canadian Journal of Program Evaluation, 28*, 1.

Bourgeois, I., Whynot, J., & Thériault, É. (2015). Application of an organizational evaluation capacity self-assessment instrument to different organizations: Similarities and lessons learned. *Evaluation and Program Planning, 50*, 47-55. https://doi.org/10.1016/j.evalprogplan.2015.01.004

Bowen, D. E., & Ostroff, C. (2004). Understanding HRM-Firm performance linkages: The role of the "strength" of the HRM system. *Academy of Management Review, 29*, 203-221. https://doi.org/10.5465/AMR.2004.12736076

Boxall, P. F. (2007). The goals of HRM. In P. Boxall, J. Purcell, & P. Wright (Eds.), *Oxford Handbook of Human Resource Management* (pp. 48-67). Oxford University Press.

Boxall, P. F., & Purcell, J. (2011). *Strategy and human resource management*. (3rd ed.) Hampshire [England]; New York: Palgrave Macmillan.

Boxall, P. F., Purcell, J., & Wright, P. (Eds.). (2007). *Human resource management: Scope, analysis and significance*. Oxford University Press.

Branham, L. (2012). *The 7 hidden reasons employees leave: How to recognize the subtle signs and act before it's too late* (2nd ed.). American Management Association.

Bredehöft, F., Dettmers, J., Hoppe, A., & Janneck. (2015). Individual work design as a job demand: The double-edged sword of autonomy. . *Psychology of Everday Activity, 8*(2), 1-14.

Breuer, K. A., Sliwka, P. D., & Kuntz, P. L. (2011). *Human capital, personnel economics, incentives, subjective performance evaluations, target agreements, intra-firm trainings* [159]. Universität zu Köln [University of Cologne], Germany.

Briner, R., & Barends, E. (2016). The role of scientific findings in evidence-based HR. *People & Strategy, 39*, 16-20.

Briner, R., Denyer, D., & Rousseau, D. (2009). Evidence-based management: Concept cleanup time? *Academy of Management Perspectives, November*, 19-32.

Brown, J., Gosling, T., Sethi, B., Sheppard, B., Stubbings, C., Sviokla, J., . . . Fisher, L. (2017). *Workforce of the future:The competing forces shaping 2030*. PwC. https://www.pwc.com/gx/en/services/people-organisation/publications/workforce-of-the-future.html

Brown, T. A. (2015). Confirmatory factor analysis for applied research. In T. D. Little (Ed.), *Methodology in the social sciences* (2nd ed.). The Guilford Press.

Brug, J., Tak, N. I., & Te Velde, S. J. (2011). Evaluation of nationwide health promotion campaigns in the Netherlands: An exploration of practices, wishes and opportunities. *Health Promotion International, 26*(2), 244-254. https://doi.org/10.1093/heapro/daq058

Bujang, M. A., Ab Ghani, P., Soelar, S. A., & Zulkifli, N. A. (2012). *Sample size guideline for exploratory factor analysis when using small sample: Taking into considerations of different measurement scales* 2012 International Conference on Statistics in Science, Business and Engineering (ICSSBE),

Burke, R. J., Koyuncu, M., Wolpin, J., Acar, F. T., & Koyuncu, K. (2015). Gender differences in work experiences and work and learning outcomes among employees in the manufacturing sector in Turkey: An exploratory study. *Journal of Industrial Relations & Human Resources, 17*(3), 168-183. https://doi.org/10.4026/1303-2860.2015.0285.x

Buss, W., & Kuyvenhofen, R. (2011). Perceptions of european middle managers of their role in strategic change. *Global Journal of Business Research, 5*(5), 11.

Butitova, D. (2019). Perceptions of organizational justice among state government employees: The effect of years of public service. *Public Personnel Management, 48*(3), 345-368. https://doi.org/10.1177/0091026018816348

Byrne, Z. S., Peters, J. M., & Weston, J. W. (2016). The struggle with employee engagement: Measures and construct clarification using five samples. *Journal of Applied Psychology, June*(9), 1-28. https://doi.org/http://dx.doi.org/10.1037/apl0000124

Caesens, G., Marique, G., Hanin, D., & Stinglhamber, F. (2016). The relationship between perceived organizational support and proactive behaviour directed towards the organization. *European Journal of Work & Organizational Psychology, 25*, 398-411.

Carden, F., & Alkin, M. C. (2012). Evaluation roots: An international perspective. *Journal of Multi-Disciplinary Evaluation, 8*(17), 102-118. https://pdfs.semanticscholar.org/d69f/cfbd989ae33a2a473fae3710590ad5acb19d.pdf

Carlson, K. D., & Kavanagh, M. J. (2018). HR Metrics and workforce analytics. In M. J. Kavanagh & R. D. Johnson (Eds.), *Human Resource Information Systems* (4th ed.). Sage.

Carman, J. G. (2007). Evaluation practice among community-based organizations: Research into the reality. *American Journal of Evaluation, 28*, 60-75.

Carpenter, M., Bauer, T., & Erdogan, B. (2012). Motivating employees. In *Management Principles* (pp. 635-673). Creative Commons.

Carrier, S. (1995). Family status and career situation for professional women. *Work, Employment & Society, 9*(2), 343-358. https://doi.org/10.1177/0950017095009002008

Carriere, J., & Bourque, C. (2009). The effects of organizational communication on job satisfaction and organizational commitment in a land ambulance service and the mediating role of communication satisfaction. *Career Development International, 14*, 29-49. https://doi.org/10.1108/13620430910933565

Cascio, W. F. (1991). *Costing human resources: The financial impact of behavior in organizations* (3rd ed.). Boston, Mass.: PWS-Kent Pub. Co.

Cascio, W. F. (2008). To prosper, organizational psychology should... bridge application and scholarship. *Journal of Organizational Behavior, 29*(4), 455-468. www.jstor.org/stable/30162639

Cascio, W. F., & Boudreau, J. (2011). *Investing in people: financial impact of human resource initiatives* (2nd ed.). Pearson Education Inc.

Cascio, W. F., & Boudreau, J. (2014). *Evidence-based management at the bottom of the pyramid: Why human resources standards and research must connect more closely*. Oxford handbooks online. http://www.oxfordhandbooks.com/view/lo.1093/oxfordhb/9780199935406

Castanheira, F., & Story, J. (2016). Making good things last longer: The role of savoring on the relationship between HRM and positive employee outcomes. *Human Resource Management, 55*(6), 985-1000. https://doi.org/10.1002/hrm.21704

Cawley, B. D., Keeping, L. M., & Levy, P. D. (1998). Participation in the performance appraisal process and employee reactions: A meta-analytic review of field investigations *Journal of Applied Psychology, 83*(4), 615-633.

Caykoylu, S., & Egri, C. P. (2011). Key organizational commitment antecedents for nurses, paramedical professionals and non-clinical staff. *Journal of Health Organization and Management, 25*, 7-33.

Center for Evidence-Based Management. (2019). *About CEBMa*. https://cebma.org/

Ceschi, A., Demerouti, E., Sartori, R., & Weller, J. (2017). Decision-Making processes in the workplace: How exhaustion, lack of resources and job demands impair them and affect performance. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.00313

Chan, S. H. J., & Lai, H. Y. I. (2017). Understanding the link between communication satisfaction, perceived justice and organizational citizenship behavior. *Journal of Business Research, 70*, 214-223.

Chaudhary, R., & Rangnekar, S. (2017). Socio-Demographic factors, contextual factors, and work engagement: Evidence from India. *Emerging Economy Studies, 3*(1), 1-18.

Chelimsky, E. (1995). The political environment of evaluation and what it means for the development of the field. *Evaluation Practice, 16*(3), 215-225. https://doi.org/https://doi.org/10.1177/109821409501600301

Chen, P.-J., & Choi, Y. (2008). Generational differences in work values: A study of hospitality management. *International Journal of Contemporary Hospitality Management, 20*(6), 595-615.

Chen, Y. Y., Zhang, Y., & Fey, C. (2011). When collaborative HR practices may not work well: The moderating role of social capital in the Chinese life insurance industry. *International Journal of Human Resource Management, 22*(2), 433-456. https://doi.org/10.1080/09585192.2011.540164

Cheng, M. M., & Coyte, R. (2014). The effects of incentive subjectivity and strategy communication on knowledge-sharing and extra-role behaviours. *Management Accounting Research, 25*, 119-130. https://doi.org/https://doi.org/10.1016/j.mar.2013.07.003

Chouinard, J. A. (2013). The case for participatory evaluation in an era of accountability. *American Journal of Evaluation, 34*, 237-253.

Christie, C. A., Alkin, M. C., Coghlan, A. T., King, J. A., Henry, G. T., Ridde, A. V., . . . Johnson, H. (2005). Participatory evaluation. In *Encyclopedia of Evaluation* (pp. 282-286). Sage Publications, Inc. https://doi.org/doi: http://dx.doi.org/10.4135/9781412950558

Christopher, S. A., Fethney, J., Chiarella, M., & Waters, D. (2018). Factors influencing turnover in GenX nurses: Results of an Australian survey. *Collegian, 25*(2), 217-225. https://doi.org/https://doi.org/10.1016/j.colegn.2017.06.003

Čikeš, V., Ribarić, H. M., & Črnjar, K. (2018). The determinants and outcomes of absence behavior: A systematic literature review. *Social Sciences, 7*, 120. https://doi.org/10.3390/socsci7080120

Civilidag, A. (2014). Hotel employees' mobbing, burnout, job satisfaction and perceived organizational support: A research on hospitality in Turkey. *European Scientific Journal, 10*(35), 1-22.

Clampitt, P. G., & Downs, C. W. (2004). *Downs-Hazen Communication Satisfaction Questionnaire*. Guilford Press.

Clampitt, P. G., & Girard, D. (1993). Communication satisfaction: A useful construct? *New Jersey Journal of Communication, 1*(2), 84-102.

Claus, L., & Briscoe, D. (2009). Employee performance management across borders: A review of relevant academic literature. *International Journal of Management Reviews, 11*(2), 175-196. https://doi.org/https://doi.org/10.1111/j.1468-2370.2008.00237.x

Cohen, D. (2007, 10/01). The very separate worlds of academic and practitioner publications in human resource management: Reasons for the divide and concrete solutions for bridging the gap. *The Academy of Management Journal, 50*, 1013-1019. https://doi.org/10.5465/AMJ.2007.27151946

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Cole, M. S., Bruch, H., & Vogel, B. (2006). Emotion as mediators of the relations between perceived supervisor support and psychological hardiness on employee cynicism. *Journal of Organizational Behavior, 27*, 21.

Collins, C. J., & Clark, K. D. (2003). Strategic human resource practices, top management team social networks, and firm performance: The role of human resource practices in creating organizational competitive advantage. *The Academy of Management Journal, 46*, 740.

Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C. O. L. H., & Ng, K. Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology, 86*, 425.

Comrey, A. L. (1988, Oct). Factor-Analytic methods of scale development in personality and clinical psychology. *J Consult Clin Psychol, 56*(5), 754-761. https://doi.org/10.1037//0022-006x.56.5.754

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis*. Hillsdale, NJ: Erlbaum.

Conway, E., Fu, N., Monks, K., Alfes, K., & Bailey, C. (2016). Demands or resources? The relationship between HR practices, employee engagement, and emotional exhaustion

within a hybrid model of employment relations. *Human Resource Management, 55*(5), 901-917. https://doi.org/10.1002/hrm.21691

Conway, E., & Monks, K. (2009). Unravelling the complexities of high commitment: An employee-level analysis. *Human Resource Management Journal, 19*(2), 140-158. https://doi.org/10.1111/j.1748-8583.2009.00090.x

Conway, J. M., & Huffcutt, A. I. (2003). A review and evaluation of exploratory factor analysis practices in organizational research. *Organizational Research Methods, 6*(2), 147-168. https://doi.org/10.1177/1094428103251541

Corley, K. G. (2004). Defined by our strategy or our culture? Hierarchical differences in perceptions of organizational identity and change. *Human Relations -New York, 57*, 1145.

Cousins, J. B., & Chouinard, J. A. (Eds.). (2012). *Participatory evaluation up close: An integration of research-based knowledge*. Information Age Pub.

Cousins, J. B., & Earl, L. M. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis, 14*(4), 397-418. https://doi.org/10.3102/01623737014004397

Cousins, J. B., Elliott, C., Amo, C., Bourgeois, I., Chouinard, J., Goh, S. C., & Lahey, R. (2008). Organizational capacity to do and use evaluation: Results of a pan-canadian survey of evaluators. *Canadian Journal of Program Evaluation, 23*, 1-35.

Cousins, J. B., Goh, S. C., Elliott, C., Aubry, T., & Gilbert, N. (2014). Government and voluntary sector differences in organizational capacity to do and use evaluation. *Evaluation and Program Planning, 44*, 1-13. https://doi.org/10.1016/j.evalprogplan.2013.12.001

Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation, 80*, 5-23. https://doi.org/10.1002/ev.1114

Cousins, J. B., Whitmore, E., & Shulha, L. (2013). Arguments for a common set of principles for collaborative inquiry in evaluation. *American Journal of Evaluation, 34*, 7-22.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Thousand Oaks: SAGE Publications.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Los Angeles: SAGE Publications.

Crino, M. D., & White, M. C. (1981). Satisfaction in communication: An examination of the Downs-Hazen measure. *Psychological Reports, 49*, 831-838.

Cucina, J. M., Byle, K. A., Martin, N. R., Peyton, S. T., & Gast, I. F. (2018). Generational differences in workplace attitudes and job satisfaction: Lack of sizable differences across

cohorts. *Journal of Managerial Psychology, 33*, 246-264. https://doi.org/10.1108/JMP-03-2017-0115

Curd, E. (2011). *The New Zealand ageing workforce: An analysis of the export manufacturing industry* [Master of Business, Auckland University of Technology].

Curtis, A. B., Harvey, R. D., & Ravden, D. (2005). Sources of political distortions in performance appraisals: Appraisal purpose and rater accountability. *Group and Organization Management, 30*, 42-60. https://doi.org/10.1177/1059601104267666

Dalal, R. S., Baysinger, M., Brummel, B. J., & LeBreton, J. M. (2012). The relative importance of employee engagement, other job attitudes, and trait affect as predictors of job performance. *Journal of Applied Social Psychology, 42*, E295-E325.

Datta, A., & Singh, R. (2018). Determining the dimensions of organizational climate perceived by the hotel employees. *Journal of Hospitality and Tourism Management, 36*, 40-48. https://doi.org/10.1016/j.jhtm.2018.07.001

David, F. R. (2013). *Strategic management concepts and cases: A competitive advantage approach*. Pearson. https://books.google.com.au/books?id=W6tOYAAACAAJ

Dayaram, K., & Fung, L. (2014). Organizational learning in the Philippines: how do team and individual learning contribute? *Asia Pacific Journal of Human Resources, 52*(4), 420-442. https://doi.org/10.1111/1744-7941.12039

de Brito, R. P., & de Oliveira, L. B. (2016). The relationship between human resource management and organizational performance. *Brazilian Business Review (English Edition), 13*, 90-110. https://doi.org/10.15728/bbr.2016.13.3.5

de Bruin, G. P., & Henn, C. M. (2013). Dimensionality of the 9-item Utrecht Work Engagement Scale (UWES-9). *Psychological Reports, 112*(3), 788-799. https://doi.org/10.2466/01.03.PR0.112.3.788-799

de Grandbois, Y. (2016). The story of service science. In Y. de Grandbois (Ed.), *Service Science and the Information Professional* (pp. 27-46). Chandos Publishing. https://doi.org/https://doi.org/10.1016/B978-1-84334-649-4.00002-0

De Nobile, J. (2017). Organisational communication and its relationships with job satisfaction and organisational commitment of primary school staff in Western Australia. *Educational Psychology, 37*, 380.

De Nobile, J., & McCormick, J. (2008). Organizational communication and job satisfaction in Australian catholic primary schools. *Educational Management Administration and Leadership, 36*(1), 101-122. https://doi.org/10.1177/1741143207084063

De Sio, S., Cedrone, F., Trovato Battagliola, E., Buomprisco, G., Perri, R., & Greco, E. (2018). The perception of psychosocial risks and work-related stress in relation to job insecurity and gender differences: A cross-sectional study. *Biomed Research International, 2018*, 1-6. https://doi.org/10.1155/2018/7649085

De Winne, S., Delmotte, J., Gilbert, C., & Sels, L. (2013). Comparing and explaining HR department effectiveness assessments: Evidence from line managers and trade union representatives. *International Journal of Human Resource Management, 24*(8), 1708-1735. https://doi.org/10.1080/09585192.2012.725069

DeCoster, J. (1998). Overview of factor analysis. http://www.stat-help.com/notes.html

Dekker, I., & Barling, J. (1995). Workforce size and work-related role stress. *Work & Stress, 9*, 45-54.

Dello Russo, S., Miraglia, M., Borgogni, L., & Johns, G. (2013). How time and perceptions of social context shape employee absenteeism trajectories. *Journal of Vocational Behavior, 83*, 209.

Delmotte, J., De Winne, S., & Sels, L. (2012). Toward an assessment of perceived HRM system strength: Scale development and validation. *International Journal of Human Resource Management, 23*(7), 1481-1506. https://doi.org/10.1080/09585192.2011.579921

Demerouti, E., Nachreiner, F., Baker, A. B., & Schaufeli, W. B. (2001). The job demands-resources model of burnout. *Journal of Applied Psychology, 86*, 499-512.

Demir, K. (2015). The effect of organizational justice and perceived organizational support on organizational citizenship behaviors: The mediating role of organizational identification. *Eurasian Journal of Educational Research, 15*(60), 131-148.

Den Hartog, D. N., Boon, C., Verburg, R. M., & Croon, M. A. (2013). HRM, communication, satisfaction, and perceived performance: A cross-level test. *Journal of Management, 39*(6), 1637-1665. https://doi.org/10.1177/0149206312440118

Den Hartog, D. N., Boselie, P., & Paauwe, J. (2004). Performance management: A model and research agenda. *Applied Psychology: An International Review, 53*, 556-569. https://doi.org/10.1111/j.1464-0597.2004.00188.x

DeNisi, A. S. (2000). Performance appraisal and performance management: A multilevel analysis. In K. J. Klein & S. Kozlowski (Eds.), *Multilevel Theory, Research and Methods in Organizations* (pp. 121-156). Jossey-Bass.

DeNisi, A. S., Cafferty, T. P., & Meglino, B. M. (1984). A cognitive view of the performance appraisal process: A model and research propositions. *Organizational Behavior and Human Performance, 33*, 360-396.

Denton, G. A., & White, B. (2000). Implementing a balanced-scorecard approach to managing hotel operations: The case of white lodging services. *Cornell Hotel and Restaurant Administration Quarterly, 41*(1), 13.

Department of Employment. (2017). *Australian Jobs 2017*. https://cica.org.au/wp-content/uploads/Australian-Jobs-2017.pdf

Department of Employment Skills Small and Family Business. (2018). *Human resource managers*. Australian Government,. https://joboutlook.gov.au/Occupation?code=1323

Department of Jobs and Small Business. (2019). *Australian jobs 2019* (1832-7230).

Derven, M. (2012). Building a strategic approach to learning evaluation. *TD Magazine*, (11), 54.

DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). SAGE.

Dewettinck, K., & van Dijk, H. (2013). Linking Belgian employee performance management system characteristics with performance management system effectiveness: Exploring the mediating role of fairness. *The International Journal of Human Resource Management, 24*(4), 806–825. https://doi.org/10.1080/09585192.2012.700169

Dimaya, R., McEwen, M., Curry, L., & Bradley, E. (2012). Managing health worker migration: A qualitative study of the Philippine response to nurse brain drain. *Human Resources for Health, 10*, 47.

Donaldson, L. (2012). Evidence-Based management (EBMgt) using organizational facts. In D. M. Rousseau (Ed.), *The Oxford Handbook of Evidence-Based Management*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199763986.013.0014

Donaldson, S. I., & Grant-Vallone, E. J. (2002). Understanding self-report bias in organizational behavior research. *Journal of Business and Psychology, 17*(2), 245-260. https://doi.org/10.1023/A:1019637632584

Donaldson, S. I., & Lipsey, M. W. (2006). Roles for theory in contemporary evaluation practice: Developing practical knowledge. In I. F. Shaw & J. C. Greene (Eds.), *The SAGE Handbook of Evaluation* (pp. 57-76). London, England: SAGE Publications Ltd. https://doi.org/doi: http://dx.doi.org/10.4135/9781848608078

Dopson, S., Bennett, C., Fitzgerald, L., Ferlie, E., Fischer, M., Ledger, J., . . . McGivern, G. (2012). *Health care managers' access and use of management research: Final report*. National Institute of Health Research.

Dorenbosch, L., Reuver, R. d., & Sanders, K. (2006). Getting the HR message across: The linkage between Line – HR consensus and "commitment strength" among eospital employees. *Management Revue, 17*(3), 274-291. https://doi.org/10.5771/0935-9915-2006-3-274

Downs, C. W., & Hazen, M. D. (1977). A factor analytic study of communication satisfaction. *Journal of Business Communication, 14*, 63.

DuBow, W. M., & Litzler, E. (2019). The development and use of a theory of change to align programs and evaluation in a complex, national initiative. *American Journal of Evaluation, 40*, 231-248.

Dulebohn, J. H., & Johnson, R. D. (2013). Human resource metrics and decision support: A classification framework. *Human Resource Management Review, 23*, 12.

e-reward. (2014). *Survey of performance management practice*. e-reward. https://www.e-reward.co.uk/research/surveys/performance-management-part-2-survey

Eby, L. T., Adams, D. M., Russell, J. E., & Gaby, S. H. (2000). Perceptions of organizational readiness for change: Factors related to employees' reactions to the implementation of team-based selling. *Human Relations - New York, 53*, 419.

Eder, P., & Eisenberger, R. (2008). Perceived organizational support: Reducing the negative influence of coworker withdrawal behavior. *Journal of Management, 34*(1), 55-68. https://doi.org/10.1177/0149206307309259

Edgar, F., & Geare, A. (2005). HRM practice and employee attitudes: Different measures - different results. *Personnel Review* (5), 534 - 549. https://doi.org/10.1108/00483480510612503

Edgar, F., & Geare, A. (2014). An employee-centred analysis: Professionals' experiences and reactions to HRM. *International Journal of Human Resource Management, 25*(5), 673-695. https://doi.org/10.1080/09585192.2013.803137

Edwards, B., Stickney, B., Milat, A., Campbell, D., & Thackway, S. (2016). Building research and evaluation capacity in population health: The NSW Health approach. *Health Promotion Journal of Australia, 27*(3), 264-267.

Ehrnrooth, M., & Bjrkman, I. (2012). An integrative HRM process theorization: Beyond signalling effects and mutual gains. *Journal of Management Studies, 49*(6), 1109-1135. https://doi.org/10.1111/j.1467-6486.2012.01055.x

Eisenberger, R., Huntington, R., Hutchison, S., & Sowa, D. (1986). Perceived organizational support. *Journal of Applied Psychology* (3), 500-507.

Eisenberger, R., Stinglhamber, F., Vandenberghe, C., Sucharski, I. L., & Rhoades, L. (2002). Perceived supervisor support: contributions to perceived organizational support and employee retention. *Journal of Applied Psychology* (3), 565.

Eldor, L., & Harpaz, I. (2016). A process model of employee engagement: The learning climate and its relationship with extra-role performance behaviors. *Journal of Organizational Behavior, 37*, 213-235.

Eley, R., Buikstra, E., Plank, A., Hegney, D., & Parker, V. (2007). Tenure, mobility and retention of nurses in Queensland, Australia: 2001 and 2004. *Journal of Nursing Management, 15*, 285-293. https://doi.org/10.1111/j.1365-2834.2007.00652.x

Ellwardt, L., Labianca, G. J., & Wittek, R. (2012). Who are the objects of positive and negative gossip at work? A social network perspective on workplace gossip. *Soc. Netw., 34*, 193–205. https://doi.org/10.1016/j.socnet.2011.11.003

Erdogan, B., Bauer, T. N., & Taylor, S. (2015). Management commitment to the ecological environment and employees: Implications for employee attitudes and citizenship behaviors. *Human Relations, 68*, 1669-1691.

Erdogan, B., Kraimer, M. L., & Liden, R. C. (2004). Work value congruence and intrinsic career success: The compensatory roles of leader-member exchange and perceived organizational support. *Personnel Psychology, 57*, 305.

Fair Work Ombudsman. (n.d.). *Managing performance & warnings*. Australian Government. https://www.fairwork.gov.au/employee-entitlements/managing-performance-and-warnings

Falcone, P., & Tan, W. (2013). *The performance appraisal tool kit: Redesigning your performance review template to drive individual and organizational change*. AMACOM.

Farndale, E., & Kelliher, C. (2013). Implementing performance appraisal: Exploring the employee experience. *Human Resource Management, 52*(6), 879-897. https://doi.org/10.1002/hrm.21575

Farndale, E., & Sanders, K. (2017). Conceptualizing HRM system strength through a cross-cultural lens. *International Journal of Human Resource Management, 28*(1), 132-148. https://doi.org/10.1080/09585192.2016.1239124

Farrell, D., & Petersen, J. C. (1984). Commitment, absenteeism, and turnover of new employees: A longitudinal study. *Human Relations, 37*(8), 681-692.

Felps, W., Mitchell, T. R., & Byington, E. (2006). How, when, and why bad apples spoil the barrel: negative group members and dysfunctional groups. *Research in Organizational Behavior, 27*, 175-222.

Ferguson, M., & Barry, B. (2011). I know what you did: The effects of interpersonal deviance on bystanders. *Journal of Occupational Health Psychology, 16*, 80-94.

Ferris, G. R., Munyon, T. P., Basik, K., & Buckley, M. R. (2008). The performance evaluation context: Social, emotional, cognitive, political, and relationship components. *Human Resource Management Review, 18*, 146-163.

Fetterman, D., Rodríguez-Campos, L., Wandersman, A., & O'Sullivan, R. G. (2014). Collaborative, participatory, and empowerment evaluation: Building a strong conceptual foundation for stakeholder involvement approaches to evaluation (A response to Cousins, Whitmore, and Shulha, 2013). *American Journal of Evaluation, 35*, 144-148. https://doi.org/10.1177/1098214013509875

Field, A. P., Miles, J., & Field, Z. (2012). *Discovering statistics using R.* London; Thousand Oaks, CA: Sage.

Fink, A., & Sturman, M. C. (2017). HR metrics and talent analystics. *The Scholarly Commons*. Retrieved 2018, from https://scholarship.sha.cornell.edu/cgi/viewcontent.cgi?article=2153&context=articles

Flint, D. H. (1999). The role of organizational justice in multi-source performance appraisal: Theory-based applications and directions for research. *Human Resource Management Review, 9*, 1-20.

Foster, J., & Guttman, R. (2018). Perceptions of job security in Australia. *Bulletin*, *March 2018*, 21. https://www.rba.gov.au/publications/bulletin/2018/mar/pdf/perceptions-of-job-security-in-australia.pdf

Fournier, D. (2005). Evaluation. In S. Mathison (Ed.), *Encyclopedia of Evaluation*. SAGE Publications, Inc. https://doi.org/http://dx.doi.org/10.4135/9781412950558.n179

Francis, L. J., & Smith, B. J. (2015). Toward best practice in evaluation: A study of Australian health promotion agencies. *Health Promotion Practice, 16*(5), 715-723. https://doi.org/10.1177/1524839915572574

Franck, R. (2002). *The explanatory power of models: Bridging the gap between empirical and theoretical research in the social sciences*. Kluwer Academic Publishers.

Frechtling, J. A. (2015). Logic models. In J. D. Wright (Ed.), *International Encyclopedia of the Social & Behavioral Sciences* (2nd ed., pp. 299-305). Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-08-097086-8.10549-5

Frenkel, S. J., & Bednall, T. (2016). How training and promotion opportunities, career expectations, and two dimensions of organizational justice explain discretionary work effort. *Human Performance, 29*(1), 16-32. https://doi.org/10.1080/08959285.2015.1120306

Frenkel, S. J., Li, M., & Restubog, S. L. D. (2012). Management, organizational justice and emotional exhaustion among Chinese migrant workers: Evidence from two

manufacturing firms. *British Journal of Industrial Relations, 50*(1), 121-147. https://doi.org/10.1111/j.1467-8543.2011.00858.x

Frenkel, S. J., Restubog, S. L. D., & Bednall, T. (2012). How employee perceptions of HR policy and practice influence discretionary work effort and co-worker assistance: Evidence from two organizations. *International Journal of Human Resource Management, 23*(20), 4193-4210. https://doi.org/10.1080/09585192.2012.667433

Furnham, A. (2000). Secrets of success from the Heathrow School of Management. *Business Strategy Review, 11*(3), 61–67.

Gable, R. K., & Wolf, M. B. (1993). *Instrument development in the affective domain: Measuring attitudes and values in corporate and school settings.* (2nd ed.). Springer.

Gavino, M. C., Wayne, S. J., & Erdogan, B. (2012). Discretionary and transactional human resource practices and employee outcomes: The role of perceived organizational support. *Human Resource Management, 51*, 665.

George, J. M., & Jones, G. R. (2012). *Understanding and managing organizational behavior* (6th ed.). Prentice Hall.

Gerhart, B., Wright, P., Mahan, G. C., & Snell, S. A. (2000). Measurement error in research on human resources and firm performance: How much error is there and how does it influence effect size estimates? *Personnel Psychology, 53*(4), 803–834.

Ghosh, P., Rai, A., & Sinha, A. (2014). Organizational justice and employee engagement: Exploring the linkage in public sector banks in India. *Personnel Review*, 628-652. https://doi.org/10.1108/PR-08-2013-0148

Gibb, S. (2000). Evaluating HRM effectiveness: The stereotype connection. *Employee Relations, 22*, 58-71.

Gilfillan, G. (2018). *Trends in use of non-standard forms of employment* (Research Paper Series, Issue 2018–19). Parliament of Australia. https://apo.org.au/sites/default/files/resource-files/2018-12/apo-nid209706.pdf

Gillet, N., Gagné, M., Sauvagère, S., & Fouquereau, E. (2013). The role of supervisor autonomy support, organizational support, and autonomous and controlled motivation in predicting employees' satisfaction and turnover intentions. *European Journal of Work & Organizational Psychology, 22*, 450-460.

Glasman, N. S., & Martens, P. A. (1993). Personnel evaluation standards: The use in principal assessment systems. 47.

Glavas, A., & Godwin, L. N. (2013). Is the perception of 'Goodness' good enough? Exploring the relationship between perceived corporate social responsibility and employee

organizational identification. *Journal of Business Ethics, 114*(1), 15-27. https://doi.org/10.1007/s10551-012-1323-5

Gok, S., Karatuna, I., & Karaca, P. O. (2015). The role of perceived supervisor support and organizational identification in job satisfaction. *Procedia - Social and Behavioral Sciences, 177*, 38-42.

Gomez-Mejia, L. R., & Balkin, D. B. (1992). *Compensation: Organizational strategy and firm performance*. South-Western Publishing.

González-Romá, V., Schaufeli, W. B., Bakker, A. B., & Lloret, S. (2006). Burnout and work engagement: Independent factors or opposite poles? *Journal of Vocational Behavior, 68*, 165-174.

Gould-Williams, J., & Davies, F. (2005). Using social exchange theory to predict the effects of HRM practice on employee outcomes: An analysis of public sector workers. *Public Management Review, 7*, 1-24.

Graham, C. M., & Nafukho, F. M. (2007). Employees' perception toward the dimension of culture in enhancing organizational learning. *The Learning Organization, 14*(3), 281-292. https://doi.org/10.1108/09696470710739435

Gray, J., & Laidlaw, H. (2004). Improving the measurement of communication satisfaction. *Management Communication Quarterly, 17*, 425-448.

Green, L. (2010). *The internet: An introduction to new media*. Berg.

Greene, J. C. (2007). *Mixed methods in social inquiry* (1st ed.). San Francisco, Calif : Jossey-Bass, c2007.

Griffin, R. (2012). Twenty-First century evaluation. *Training Journal*, 50-53.

Guest, D. (1991). Personnel management: The end of orthodoxy? *British Journal of Industrial Relations, 29*, 149-175.

Guest, D. (1997). Human resource management and performance: A review and research agenda. *International Journal of Human Resource Management, 8*, 263-276. https://doi.org/10.1080/095851997341630

Guest, D. (2002). Human resource management, corporate performance and employee wellbeing: Building the worker into HRM. *Journal of Industrial Relations, 44*(3), 335-358.

Guglielmi, D., Avanzi, L., Chiesa, R., Mariani, M. G., Bruni, I., & Depolo, M. (2016). Positive aging in demanding workplaces: The gain cycle between job satisfaction and work engagement. *Frontiers in Psychology, 7*, 1-10. https://doi.org/10.3389/fpsyg.2016.0122

Gullickson, A. M., & Al-Nawab, H. F. (2017). Making formative assessment feasible: A collaborative approach supporting grant funded projects. *Evaluation Journal of Australasia, 17*(4), 4-12.

Gyekye, S. A., & Salminen, S. (2009). Perceived organizational support: An African perspective. *Journal of Applied Social Psychology, 39*, 2651-2668.

Haines, V., & St-Onge, S. (2012). Performance management effectiveness: practices or context? *International Journal of Human Resource Management, 23*, 1158.

Handa, D., & Garima. (2014). Human resource (HR) analytics: Emerging trends in HRM. *CLEAR International Journal of Research in Commerce & Management, 5*(6), 59-62.

Hansen, M., Alkin, M. C., & Wallace, T. L. (2013). Depicting the logic of three evaluation theories. *Evaluation and Program Planning, 38*, 34-43. https://doi.org/https://doi.org/10.1016/j.evalprogplan.2012.03.012

Harris, K. J., Lambert, A., & Harris, R. B. (2013). HRM effectiveness as a moderator of the relationships between abusive supervision and technology work overload and job outcomes for technology end users. *Journal of Applied Social Psychology, 43*, 1686-1695.

Harrison, R. (2009). *Learning and development.* (5th ed.). Chartered Institute of Personnel and Development.

Hassan, B., Maqsood, A., & Riaz, M. N. (2011). Relationship between organizational communication climate and interpersonal conflict management styles. *Pakistan Journal of Psychology, 42*, 23-40.

Hassan, S. (2013). The importance of role clarification in workgroups: Effects on perceived role clarity, work satisfaction, and turnover rates. *Public Administration Review, 73*(5), 716-725. www.jstor.org/stable/42003000

Hays, L. J., O'Brien-Pallas, L., Duffield, C., Shamian, J., Buchan, J., Hughes, F., . . . North, N. (2012). Nurse turnover: A literature review - An update. *International Journal of Nursing Studies, 7*, 887.

He, H., Pham, H. Q., Baruch, Y., & Zhu, W. (2014). Perceived organizational support and organizational identification: Joint moderating effects of employee exchange ideology and employee investment. *International Journal of Human Resource Management, 25*, 2772-2795.

Heath, C., & Heath, D. (2007). *Made to stick: Why some ideas survive and others die*. Random House.

Hegney, D., Plank, A., & Parker, V. (2006). Extrinsic and intrinsic work values: Their impact on job satisfaction in nursing. *Journal of Nursing Management, 14*, 271-281. https://doi.org/10.1111/j.1365-2934.2006.00618.x

Helms, M., & Stern, R. (2001). Exploring the factors that influence employees' perceptions of their organisation's culture. *Journal of Management in Medicine, 15*, 415.

Henderson, A., Schoonbeek, S., Ossenberg, C., Caddick, A., Wing, D., Capell, L., & Gould, K. (2014). Achieving success in intervention studies: an analysis of variable staff engagement across three midwifery settings. *Journal of Clinical Nursing, 23*(11/12), 1653-1661. https://doi.org/10.1111/jocn.12296

Heritage, B., Gilbert, J. M., & Roberts, L. D. (2016). Job embeddedness demonstrates incremental validity when predicting turnover intentions for Australian university employees. *Frontiers in Psychology, 7*(62), 582. https://doi.org/10.3389/fpsyg.2016.00582

Herrbach, O., Mignonac, K., Vandenberghe, C., & Negrini, A. (2009). Perceived HRM practices, organizational commitment, and voluntary early retirement among late-career managers. *Human Resource Management, 48*(6), 895-915. https://doi.org/10.1002/hrm.20321

Heslin, P. A., & VandeWalle, D. (2011). Performance appraisal procedural justice: The role of a manager's implicit person theory. *Journal of Management, 37*, 1694-1718.

Hewett, R., Shantz, A., Mundy, J., & Alfes, K. (2018). Attribution theories in human resource management research: A review and research agenda. *International Journal of Human Resource Management, 29*(1), 87-126. https://doi.org/10.1080/09585192.2017.1380062

Hodgkinson, G. P. (2012). The politics of evidence-based decision making. In D. M. Rousseau (Ed.), *The oxford handbook of evidence-based management*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199763986.013.0023

Hofstede, G. (1998). Identifying organizational subcultures. *Journal of Management Studies, 35*, 1-12.

Holtom, B. C., Mitchell, T. R., Lee, T. W., & Eberly, M. B. (2008). Turnover and retention research: A glance at the past, a closer of the present, and a venture into the future. *Academy of Management Annals, 2*, 231-274.

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods, 6*(1), 53-59.

Howard, B. B., & Harman, S. S. (2007). An application of the "Personnel Evaluation Standards" to determine the effectiveness of a system of teacher evaluation. *ERS Spectrum, 25*, 45-56.

Hung, T.-K., Chi, N.-W., & Lu, W.-L. (2009). Exploring the relationships between perceived coworker loafing and counterproductive work behaviors: The mediating role of a revenge motive. *Journal of Business and Psychology, 24*, 257–270. https://doi.org/10.1007/s10869-009-9104-6

Hunt, A., & Samman, E. (2019). *Gender and the gig economy: Critical steps for evidence-based policy* [Working Paper](Work, Learning and Wellbeing in the Digital Economy, Issue January 2019). Analysis and Policy Observatory. https://apo.org.au/sites/default/files/resource-files/2019-01/apo-nid217091.pdf

Hunter, D. E. K., & Nielsen, S. B. (2013). Performance management and evaluation: Exploring complementarities. *New Directions for Evaluation, 2013*, 7-17. https://doi.org/10.1002/ev.20042

Huselid, M. A. (1995). The impact of human resource management practices on turnover, productivity, and corporate financial performance. *Academy of Management Journal, 38*, 635–672.

Huselid, M. A., & Becker, B. E. (2000). Comment on "Measurement error in research on human resources and firm performance: How much error is there and how does it influence effect size estimates?" by Gerhart, Wright, McMahan, and Snell. *Personnel Psychology, 53*(4), 835-854. https://doi.org/10.1111/j.1744-6570.2000.tb02419.x

Huselid, M. A., Becker, B. E., & Beatty, R. W. (2005). *The workforce scorecard: Managing human capital to execute strategy*. Harvard Business School Press.

Huselid, M. A., Jackson, S. E., & Schuler, R. S. (1997). Technical and strategic human resource management effectiveness as determinants of firm performance. *Academy of Management Journal, 40*(1), 171-188. https://doi.org/10.2307/257025

Hutchinson, M., Vickers, M. H., Wilkes, L., & Jackson, D. (2009). "The worse you behave, the more you seem, to be rewarded": Bullying in nursing as organizational corruption. *Employee Responsibilities and Rights Journal, 21*, 213-229. https://doi.org/10.1007/s10672-009-9100-z

Ichniowski, C., Shaw, K., & Prennushi, G. (1997). The effects of human resource management practices on productivity: A study of steel finishing lines. *The American Economic Review, 87*(3), 291-313. www.jstor.org/stable/2951347

Iqbal, S., & Hashmi, M. S. (2015). Impact of perceived organizational support on employee retention with mediating role of psychological empowerment. *Pakistan Journal of Commerce and Social Sciences, 9*, 18-34.

Jackson, D. L., Gillaspy, A. J. J., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological Methods, 14*(1), 6-23. https://doi.org/10.1037/a0014694

Jackson, S. E., & Schüler, R. S. (1985). A meta-analysis and conceptual critique of research on role ambiguity and role conflict in work settings. *Organizational Behavior and Human Decision Processes, 36*, 16-78.

Jalalkamali, M., Ali, A. J., Sunghyup Sean, H., & Nikbin, D. (2016). Relationships between work values, communication satisfaction, and employee job performance: The case of international joint ventures in Iran. *Management Decision*(4), 796.

James, L., & Mathew, L. (2012). Employee retention strategies: IT industry. *SCMS Journal of Indian Management, 9*(3), 79-87.

Järlström, M., Saru, E., & Vanhala, S. (2018). Sustainable human resource management with salience of stakeholders: A top management perspective. *Journal of Business Ethics, 152*(3), 703-724. https://doi.org/10.1007/s10551-016-3310-8

Jayawardana, A. K. L., O'Donnell, M., & Jayakody, J. A. S. K. (2013). Job involvement and performance among middle managers in Sri Lanka. *International Journal of Human Resource Management, 24*, 4008-4025.

Jensen, J. M., Patel, P. C., & Messersmith, J. G. (2013). High-Performance work systems and job control: Consequences for anxiety, role overload, and turnover intentions. *Journal of Management, 39*(6), 1699-1724.

Jiang, K., Hu, J., Liu, S., & Lepak, D. P. (2017). Understanding employees' perceptions of human resource practices: Effects of demographic dissimilarity to managers and coworkers. *Human Resource Management, 56*, 69.

Jiang, K., Lepak, D. P., Jia, J. U., & Baer, J. C. (2012). How does human resource management influence organizational outcomes? A meta-analytic investigation of mediating mechanisms. *Academy of Management Journal, 55*, 1264-1294. https://doi.org/10.5465/amj.2011.0088

Jiang, L., Li, F., Li, Y., & Li, R. (2017). Leader-Member exchange and safety citizenship behavior: The mediating role of coworker trust. *Work, 56*, 387-395.

Joint Committee on Standards for Educational Evaluation. (2008). *The personnel evaluation standards: How to assess systems for evaluating educators* (2nd ed.). Sage Publications, c1988.

Joo, B.-K. B., Hahn, H.-J., & Peterson, S. L. (2015). *Turnover intention: The effects of core self-evaluations, proactive personality, perceived organizational support, developmental feedback, and job complexity* Great Britain, Taylor & Francis.

Jordan, P. J., & Troth, A. C. (2020). Common method bias in applied settings: The dilemma of researching in organizations. *Australian Journal of Management*(1), 3.

Juris Bennett, C., Walston, S. L., & Al-Harbi, A. (2015). Understanding the effects of age, tenure, skill, and gender on employee perceptions of healthcare information technology within a Middle Eastern hospital. *International Journal of Healthcare Management, 8*(4), 272-280. https://doi.org/10.1179/2047971915Y.0000000010

Kadoya, Y., Khan, M. S. R., & Sano, Y. (2018). Effects of China's one-child policy on group cooperation: Survey evidence. *Journal of the Asia Pacific Economy, 23*(3), 327-339.

Kahn, W. A. (1990). Psychological conditions of personal engagement and disengagement at work. *The Academy of Management Journal, 33*(4), 692 - 724.

Kalleberg, A. (2003). Flexible firms and labour market segmentation: Effect of workplace restructuring on jobs and workers. *Work and Occupations, 30*, 154–175.

Kang, H. J. A., & Busser, J. A. (2018). Impact of service climate and psychological capital on employee engagement: The role of organizational hierarchy. *International Journal of Hospitality Management, 75*, 1-9. https://doi.org/https://doi.org/10.1016/j.ijhm.2018.03.003

Kaplan, R. S., & Norton, D. P. (1996). *The balanced scorecard: Translating strategy into action.* Boston, MA: Harvard Business School Press.

Karanges, E., Johnston, K., Beatson, A., & Lings, I. (2015). The influence of internal communication on employee engagement: A pilot study. *Public Relations Review, 41*, 129-131. https://doi.org/10.1016/j.pubrev.2014.12.003

Karanika-Murray, M., Pontes, H. M., Griffiths, M. D., & Biron, C. (2015). Sickness presenteeism determines job satisfaction via affective-motivational states. *Social Science & Medicine, 139*, 100-106.

Kasalak, G., & Bilgin Aksu, M. (2014). The relationship between perceived organizational support and organizational cynicism of research assistants. *Educational Sciences: Theory & Practice, 14*, 125-133. https://doi.org/10.12738/estp.2014.1.1765

Kaše, R., Paauwe, J., & Zupan, N. (2009). HR practices, interpersonal relations, and intrafirm knowledge transfer in knowledge-intensive firms: A social network perspective. *Human Resource Management, 48*(4), 615-639.

Katou, A. A. (2015). The mediating effects of psychological contracts on the relationship between human resource management systems and organisational performance. *International Journal of Manpower, 36*, 1012-1033. https://doi.org/10.1108/IJM-10-2013-0238

Katou, A. A., Budhwar, P. S., & Patel, C. (2014). Content vs. process in the HRM-Performance relationship: An empirical examination. *Human Resource Management, 53*(4), 527-544. https://doi.org/10.1002/hrm.21606

Katsaros, K. K., Tsirikas, A. N., & Bani, S. N. (2014). Exploring employees' perceptions, job-related attitudes and characteristics during a planned organizational change. *International Journal of Business Science and Applied Management, 9*(1), 36-50.

Keeping, L. M., & Levy, P. E. (2000). Performance appraisal reactions: Measurement, modeling, and method bias. *Journal of Applied Psychology, 85*(5), 708-723.

Kehoe, R. R., & Collins, C. J. (2017). Human resource management and unit performance in knowledge-intensive work. *Journal of Applied Psychology, 102*, 1222-1236.

Kehoe, R. R., & Wright, P. (2013). The impact of high-performance human resource practices on employees' attitudes and behaviors. *Journal of Management, 39*(2), 366-391. https://doi.org/10.1177/0149206310365901

Kekahio, W., Lawton, B., Cicchinelli, L., & Brandon, P. R. (2014). *Logic models: A tool for effective program planning, collaboration, and monitoring. REL 2014-025*. National Center for Education Evaluation and Regional Assistance & Regional Educational Laboratory Pacific.

Kellerman, A. (1985). The evolution of service economies: A geographical perspective. *The Professional Geographer, 37*(2), 133-143. https://doi.org/10.1111/j.0033-0124.1985.00133.x

Kenessey, Z. (1987). The primary, secondary, tertiary and quaternary sectors of the economy. *Review of Income & Wealth, 33*(4), 359-385. https://doi.org/10.1111/j.1475-4991.1987.tb00680.x

Kerr, S. (1975). On the follow of rewarding A while hoping for B. *Academy of Management Journal, 28*, 769-783.

Khalil, H., Peters, M., Godfrey, C. M., McInerney, P., Soares, C. B., & Parker, D. (2016). An evidence-based approach to scoping reviews. *Worldviews on Evidence-Based Nursing*, 118. https://doi.org/10.1111/wvn.12144

Kilroy, S., Flood, P. C., Bosak, J., & Chênevert, D. (2016). Perceptions of high-involvement work practices and burnout: The mediating role of job demands. *Human Resource Management Journal, 26*(4), 408-424. https://doi.org/10.1111/1748-8583.12112

Kilroy, S., Flood, P. C., Bosak, J., & Chênevert, D. (2017). Perceptions of high-involvement work practices, person-organization fit, and burnout: A time-lagged study of health care employees. *Human Resource Management, 56*(5), 821-835. https://doi.org/10.1002/hrm.21803

Kim, I.-H., Muntaner, C., Vahid Shahidi, F., Vives, A., Vanroelen, C., & Benach, J. (2012). Welfare states, flexible employment, and health: A critical review. *Health Policy, 104*(2), 99-127. https://doi.org/https://doi.org/10.1016/j.healthpol.2011.11.002

Kim, J. S., & Han, S. H. (2017). Examining the relationship between civil servant perceptions of organizational culture and job attitudes: In the context of the new public management reform in South Korea. *Public Organization Review, 17*(1), 157–175. https://doi.org/https://doi.org/10.1007/s11115-016-0372-0

Kim, T., Henderson, A. C., & Eom, T. H. (2015). At the front line: Examining the effects of perceived job significance, employee commitment, and job involvement on public service motivation. *International Review of Administrative Sciences, 81*(4), 713-733. https://doi.org/10.1177/0020852314558028

King, J. A., Cousins, J. B., & Whitmore, E. (2007, 01/01/). Making sense of participatory evaluation: Framing participatory evaluation. *New Directions for Evaluation*(114), 83-105. http://dx.doi.org/10.1002/ev.226

King, J. A., Stevahn, L., Ghere, G., & Minnema, J. (2001). Forum: Toward a taxonomy of essential evaluator competencies. *American Journal of Evaluation, 22*, 229-247. https://doi.org/10.1016/S1098-2140(01)00129-1

Kinnie, N., Hutchinson, S., Purcell, J., Rayton, B., & Swart, J. (2005). Satisfaction with HR practices and commitment to the organisation: Why one size does not fit all. *Human Resource Management Journal, 15*, 9-29.

Kirkpatrick, D. L. (1959). Techniques for evaluating training programs. *Journal of American Society for Training and Development, 23*, 2.

Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating training programs: The four levels*. (3rd ed.). Berrett-Koehler Publishers.

Klatt, E. C., & Klatt, C. A. (2011). How much is too much reading for medical students? Assigned reading and reading rates at one medical school. *Academic Medicine: Journal Of The Association Of American Medical Colleges, 86*, 1079-1083. https://doi.org/10.1097/ACM.0b013e31822579fc

Kline, R. B. (2011). *Methodology in the social sciences: Principles and practice of structural equation modeling* (3rd ed.). Guilford Press.

Kooij, D. T. A. M., & Boon, C. (2018). Perceptions of HR practices, person-organisation fit, and affective commitment: The moderating role of career stage. *Human Resource Management Journal, 28*, 61. https://doi.org/10.1111/1748-8583.12164

Koontz, H., O'Donnell, C., & Weihrich, H. (1986). Essentials of management. *McGraw-Hill series in management*.

Köse, A. (2016). The relationship between work engagement behavior and perceived organizational support and organizational climate. *Journal of Education and Practice, 7*, 42-52.

Koster, F. (2011). Able, willing, and knowing: The effects of HR practices on commitment and effort in 26 European countries. *International Journal of Human Resource Management, 22*(14), 2835-2851. https://doi.org/10.1080/09585192.2011.599946

Kottke, J. L., & Sharafinski, C. E. (1988). Measuring perceived supervisory and organizational support. *Educational & Psychological Measurement, 48*, 1075-1079.

Kovner, A. R. (2012). Adventures in the evidence-based management trade. In D. M. Rousseau (Ed.), *The Oxford Handbook of Evidence-Based Management*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199763986.013.0010

Kruse, K. (2012). The performance appraisal: A workplace evil that must be destroyed like a blood sucking vampire. *Forbes 2015*.

Krywulak, T., Roberts, M., & Conference Board of Canada. (2009). *Winning the "generation wars": Making the most of generational differences and similarities in the workplace*. Ottawa: Conference Board of Canada. https://central.bac-lac.gc.ca/.item?id=winning_generation_wars&op=pdf&app=Library

Kuvaas, B., & Dysvik, A. (2009). Perceived investment in employee development, intrinsic motivation and work performance. *Human Resource Management Journal, 19*(3), 217-236. https://doi.org/10.1111/j.1748-8583.2009.00103.x

Kyriazos, T. A. (2018). Applied psychometrics: Sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology, 9*, 2207-2230. https://doi.org/10.4236/psych.2018.98126

Lamm, E., Tosti-Kharas, J., & King, C. E. (2015). Empowering employee sustainability: Perceived organizational support toward the environment. *Journal of Business Ethics, 128*, 207.

Lapoint, P. A., & Liprie-Spence, A. (2017). Employee Engagement: Generational Differences in the Workforce. *Journal of Organizational Psychology, 17*(5), 118 - 128.

Lavelle, J. J., McMahan, G. C., & Harris, C. M. (2009). Fairness in human resource management, social exchange relationships, and citizenship behavior: Testing linkages of the target similarity model among nurses in the United States. *International Journal of Human Resource Management, 20*, 2419-2434.

Law, D. R. (2007). Appraising performance appraisals: A critical look at an external control management technique. *International Journal of Reality Therapy, 26*, 18-25.

Lawler, E. E. I. (1994). Performance management: The next generation. *Compensation and Benefits Review, 26*, 16.

Lawler, E. E. I. (2007). Why HR practices are not evidence-based. *Academy of Management Journal, 50*, 1033-1036. https://doi.org/10.5465/AMJ.2007.27155013

Lawler, E. E. I., McDermott, M., & Benson, G. S. (2012). What makes performance appraisals effective? *Compensation and Benefits Review.44*, 191.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer Publushing Company.

Lebas, M. J. (1995). Performance measurement and performance management. *International Journal of Production Economics, 41*, 23-35. https://doi.org/http://dx.doi.org/10.1016/0925-5273(95)00081-X

Lee, C. D. (2005). Rethinking the goals of your performance-management system. *Employment Relations Today. 32*, 53.

Lee, J., & Peccei, R. (2007). Perceived organizational support and affective commitment: the mediating role of organization-based self-esteem in the context of job insecurity. *Journal of Organizational Behavior, 28*, 661-685.

Leeuw, F. L., & Donaldson, S. I. (2015). Theory in evaluation: Reducing confusion and encouraging debate. *Evaluation, 21*(4), 467-480. https://doi.org/https://doi.org/10.1177/1356389015607712

Legge, K. (1978). *Power, innovation and problem solving in personnel management*. McGraw-Hill.

Lencioni, P. (2002). *The five dysfunctions of a team: A leadership fable*. Jossey-Bass.

Lengnick-Hall, M. L., Lengnick-Hall, C. A., Andrade, L. S., & Drake, B. (2009). Strategic human resource management: The evolution of the field. *Human Resource Management Review, 19*, 64-85. https://doi.org/10.1016/J.HRMR.2009.01.002

Leung, O., & Bartunek, J. M. (2012). Enabling evidence-based management: Bridging the gap between academics and practitioners. In D. M. Rousseau (Ed.), *The Oxford Handbook of Evidence-Based Management*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199763986.013.0009

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science, 5*, 69-77. https://doi.org/10.1186/1748-5908-5-69

Levin-Rozalis, M. (2003). Evaluation and research: Differences and similarities. *The Canadian Journal of Program Evaluation, 18*, 1-31.

Lewis, S., Linganiso, X., & Karodia, A. M. (2015). Investigating employee perceptions of performance management at the department of radiology in a public hospital in Gauteng. *Arabian Journal of Business and Management Review (OMAN Chapter), 5*(1), 9-39. https://doi.org/10.12816/0019103

Li, F., Deng, H., Leung, K., & Zhao, Y. (2017). Is perceived creativity-reward contingency good for creativity? The role of challenge and threat appraisals. *Human Resource Management, 56*(4), 693–709. https://doi.org/10.1002/hrm.21795

Li, X., & Frenkel, S. J. (2017). Where hukou status matters: Analyzing the linkage between supervisor perceptions of HR practices and employee work engagement. *International Journal of Human Resource Management, 28*(17), 2375-2402. https://doi.org/10.1080/09585192.2015.1137613

Li, X., Frenkel, S. J., & Sanders, K. (2011). Strategic HRM as process: How HR system and organizational climate strength influence Chinese employee attitudes. *International Journal of Human Resource Management, 22*, 1825-1842. https://doi.org/10.1080/09585192.2011.573965

Liao, H., Toya, K., Lepak, D. P., & Hong, Y. (2009). Do they see eye to eye? Management and employee perspectives of high-performance work systems and influence processes on service quality. *Journal of Applied Psychology, 94*, 371-391.

Liden, R. C., Wayne, S. J., Jaworski, R. A., & Bennett, N. (2004). Social loafing: A field investigation. *Journal of Management, 30*(2), 285-304. https://doi.org/https://doi.org/10.1016/j.jm.2003.02.002

LinkedIn Corporation. (n.d.). *LinkedIn*. https://au.linkedin.com/

Linn, R., Buchmann, M., Gould, B., Kellaghan, T., Lawrence, D., Robinson, P., & Zirkel, P. (1989). The development, validation, and applicability of the personnel evaluation standards. *Journal of Personnel Evaluation in Education, 2*, 199.

Lischka, J. (2015). *Pressure and organisational change in the digital age*. Reuters Institute for the Study of Journalism.

Liu, Z., Rexachs, D., Epelde, F., & Luque, E. (2017). An agent-based model for quantitatively analyzing and predicting the complex behavior of emergency departments. *Journal of Computational Science, 21*, 11-23. https://doi.org/https://doi.org/10.1016/j.jocs.2017.05.015

Lobo, R., McManus, A., Brown, G., Hildebrand, J., & Maycock, B. (2010). Evaluating peer-based youth programs: Barriers and enablers. *Evaluation Journal of Australasia*(1), 36.

Lobo, R., Petrich, M., & Burns, S. K. (2014). Supporting health promotion practitioners to undertake evaluation for program development. *BMC Public Health, 14*, 1315-1315. https://doi.org/10.1186/1471-2458-14-1315

Loup, K. S., & Ellett, C. D. (1997). Application of the "Personnel Evaluation Standards" to local district teacher evaluation programs: Analyses of 14 cases. Annual Meeting of the American Educational Research Association, Chicago, IL.

Lu, H., Barriball, K. L., Zhang, X., & While, A. E. (2012). Job satisfaction among hospital nurses revisited: A systematic review. *International Journal of Nursing Studies, 49*, 1017-1038. https://doi.org/http://dx.doi.org/10.1016/j.ijnurstu.2011.11.009

Lundmark, E., & Biorkman, H. (2011). New jobs – new ideas: The relationships between tenure and perceived creative climate. *Human Resource Development International, 14*, 605-621. https://doi.org/10.1080/13678868.2011.618712

Lunenburg, F. C. (2011). Expectancy theory of motivation: Motivating by altering expectations. *International Journal of Management, Business, and Administration, 15*.

Ma, S., Silva, M. G., Callan, V. J., & Trigo, V. (2016). Control and commitment HR practices, job satisfaction and turnover intentions: A comparison between local and multinational firms in China. *International Journal of Human Resource Management, 27*(9), 974-990. https://doi.org/10.1080/09585192.2015.1052535

Mache, S., Vitzthum, K., Wanke, E., A., D., Klapp, B. F., & Danzer, G. (2014). Exploring the impact of resilience, self-efficacy, optimism and organizational resources on work engagement. *Work, 47*, 491-500.

Macky, K., & Boxall, P. (2007). The relationship between 'high-performance work practices' and employee attitudes: An investigation of additive and interaction effects. *The international journal of human resource management, 18*(4), 537–567.

Maden, C., & Kabasakal, H. (2014). The simultaneous effects of fit with organizations, jobs and supervisors on major employee outcomes in Turkish banks: Does organizational support matter? *International Journal of Human Resource Management, 25*, 341-366.

Maertz, C. P., Griffeth, R. W., Campbell, N. S., & Allen, D. G. (2007). The effects of perceived organizational support and perceived supervisor support on employee turnover. *Journal of Organizational Behavior, 28*(8), 1059-1075.

Mainiero, L. A., & Gibson, D. E. (2017). The kaleidoscope career model revisited: How midcareer men and women diverge on authenticity, balance, and challenge. *Journal of Career Development, 45*(4), 361-377. https://doi.org/10.1177/0894845317698223

Malach-Pines, A. (2005). The burnout measure, short version. *International Journal of Stress Management, 12*, 78-88.

Malik, S., & Subramanian, V. V. (2015). Job expectations and perception variations across career stages: An empirical study in Indian tractor industry. *Paradigm: the Journal of Institute of Management Technology, 19*, 212-231.

Malinen, S., Wright, S., & Cammock, P. (2013). What drives organisational engagement?: A case study on trust, justice perceptions and withdrawal attitudes. *Evidence-based HRM, 1*(1), 96-108. https://doi.org/10.1108/20493981311318638

Mao, Y., & Hale, C. L. (2015). Relating intercultural communication sensitivity to conflict management styles, technology use, and organizational communication satisfaction in multinational organizations in China. *Journal of Intercultural Communication Research 44*, 132.

Marin-Garcia, J. A., & Martinez Tomas, J. (2016). Deconstructing AMO framework: A systematic review. *Intangible Capital*, *12*(4). https://www.raco.cat/index.php/Intangible/article/view/314570

Marinaccio, A., Ferrante, P., Corfiati, M., Di Tecco, C., Rondinone, B. M., Bonafede, M., . . . Iavicoli, S. (2013). The relevance of socio-demographic and occupational variables for the assessment of work-related stress risk. *BMC Public Health, 13*(1), 1-18. https://doi.org/10.1186/1471-2458-13-1157

Markides, C. (2007). In search of ambidextrous professors. *Academy of Management Journal, 50*(4), 762–768.

Marler, J. H., & Boudreau, J. W. (2017). An evidence-based review of HR analytics. *International Journal of Human Resource Management, 28*, 3-26.

Marshall, J. (2016). *A comparison of evaluation practices based on e-learning and mobile learning delivery rates* E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2016, Washington, DC, United States. https://www.learntechlib.org/p/173924

Marshall, J., & Rossett, A. (2014). Perceptions of barriers to the evaluation of workplace learning programs. *Perf Improvement Qrtly, 27*(3), 7-26.

Mathison, S. (Ed.). (2005). *Encyclopedia of evaluation*. Sage Publications, Inc. https://doi.org/doi: http://dx.doi.org/10.4135/9781412950558.

Mathison, S. (2008). What is the difference between evaluation and research - and why do we care? In N. L. Smith & P. R. Brandon (Eds.), *Fundamental issues in evaluation*. New York: Guilford Press, c2008.

McCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*(1), 15.

McCarthy, A., Cleveland, J., Hunter, S., Darcy, C., & Grady, G. (2013). Employee work–life balance outcomes in Ireland: A multilevel investigation of supervisory support and perceived organizational support. *International Journal of Human Resource Management, 24*, 1257-1276.

McConnell, C. R. (1994). The reality of perception in employee relations: "If it looks like a duck ...". *The Health Care Supervisor, 12*, 75-86.

McGonagle, A. K., Fisher, G. G., Barnes-Farrell, J. L., & Grosch, J. W. (2015). Individual and work factors related to perceived work ability and labor force outcomes. *The Journal of Applied Psychology, 100*(2), 376-398. https://doi.org/10.1037/a0037974

McGovern, P., Hill, S., Mills, C., & White, M. (2007). *Market, class, and employment*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199213375.001.0001

McKegg, K., Wehipeihana, N., & Pipi, K. (2016). *The evaluation capacity building project: Developing an evaluation capacity assessment tool and process for New Zealand NGOs* (Using Evidence for Impact).

McLaughlin, J. A., & Jordan, G. B. (1999). Logic models: A tool for telling your program's performance story. *Evaluation and Program Planning, 22*, 65-72.

McLean, G. N. (2005). Examining approaches to HR evaluation. *Strategic HR Review, 4*, 24-27.

Meintjes, C., & Steyn, B. (2006). A critical evaluation of the Downs-Hazen instrument (CSQ) by measuring employee communication satisfaction at a private higher education institution in South Africa. *Communication: South African Journal for Communication Theory & Research, 32*(1), 152-188. https://doi.org/10.1080/02500160608537968

Mercer, S. L., DeVinney, B. J., Fine, L. J., Green, L. W., & Dougherty, D. (2007). Study designs for effectiveness and translation research: Identifying trade-offs. *American Journal of Preventive Medicine, 33*(2), 139-154. https://doi.org/http://dx.doi.org/10.1016/j.amepre.2007.04.005

Mertens, D. M. (2013). What does a transformative lens bring to credible evidence in mixed methods evaluations? *New Directions for Evaluation, 2013*, 27-35. https://doi.org/10.1002/ev.20055

Mertens, D. M., & Wilson, A. T. (2012). *Program evaluation theory and practice: A comprehensive guide* (1st ed.). Guilford Press.

Metheny, G. A. (2013). The value of leading from the middle. *Leadership Advance Online* (23). Retrieved 2019, from https://www.regent.edu/acad/global/publications/lao/issue_23/metheny.pdf

Meyer, J. P., Stanley, D. J., Herscovitch, L., & Topolnytsky, L. (2002). Affective, continuance, and normative commitment to the organization: A meta-analysis of antecedents, correlates, and consequences. *Journal of Vocational Behavior, 61*, 20-52. https://doi.org/10.1006/jvbe.2001.1842

Milkovich, G. T., & Wigdor, A. K. (1991). Pay for performance: Evaluating performance appraisal and merit pay. National Academy Press.

Mitic, S., Nikolic, M., Jankov, J., Vukonjanski, J., & Terek, E. (2017). The impact of information technologies on communication satisfaction and organizational learning in companies in Serbia. *Computers in Human Behavior, 76*, 87. https://doi.org/10.1016/j.chb.2017.07.012

Mone, E. M., & London, M. (2010). *Employee engagement through effective performance management a practical guide for managers*. Routledge.

Monks, K., Conway, E., Fu, N., Bailey, K., Kelly, G., & Hannon, E. (2016). Enhancing knowledge exchange and combination through HR practices: Reflexivity as a translation process. *Human Resource Management Journal, 26*(3), 304-320. https://doi.org/10.1111/1748-8583.12108

Morris, M. G., Venkatesh, V., & L, A. P. (2005). Gender and age differences in employee decisions about new technology: An extension to the theory of planned behavior. *IEEE Transactions on Engineering Management, 52*(1).

Morrow, P. C. (2011). Managing organizational commitment: Insights from longitudinal research. *Journal of Vocational Behavior, 79*, 18-35.

Mowday, R. T., Steers, R. M., & Porter, L. W. (1979). The measurement of organizational commitment. *Journal of Vocational Behavior, 14*, 224-247. https://doi.org/http://dx.doi.org/10.1016/0001-8791(79)90072-1

Mukkavilli, M., Kulkarni, S., Doshi, D., Reddy, S., Reddy, P., & Reddy, S. (2017). Assessment of work engagement among dentists in Hyderabad. *Work, 58*, 333-340.

Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018, 2018/11/19). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology, 18*(1), 143. https://doi.org/10.1186/s12874-018-0611-x

Munro, G. D. (2010). The scientific impotence excuse: Discounting belief-threatening scientific abstracts. *Journal of Applied Social Psychology, 40*, 579–600.

Musah, A. A., Zulkipli, G., & Ahmad, N. S. I. (2017). Relationship between organizational communication and job satisfaction in temporary work environment: An empirical study of plant turnaround workers. *Global Business and Management Research: An International Journal, 9*, 73.

Musek, J. (2017). *The general factor of personality*. Academic Press: An imprint of Elsevier.

Naccarella, L., Pirkis, J., Kohn, F., Morley, B., Burgess, P., & Blashki, G. (2007). Building evaluation capacity: Definitional and practical implications from an Australian case study. *Evaluation and Program Planning, 30*(3), 231-236. https://doi.org/10.1016/j.evalprogplan.2007.05.001

Nagadevara, V., Srinivasan, V., & Valk, R. (2008). Establishing a link between employee turnover and withdrawal behaviours: Application of data mining techniques. *Research & Practice in Human Resource Management, 16*(2), 81-99.

Nair, A. G., Jain, P., Agarwal, A., & Jain, V. (2017). Work satisfaction, burnout and gender-based inequalities among ophthalmologists in India: A survey. *Work, 56*, 221-228.

Nankervis, A. R. (2006). Performance management: Theory in practice? *Asia Pacific Journal of Human Resources, 44*, 83.

Netemeyer, R. G., Bearden, W. O., & Sharma, S. (2003). *Scaling procedures: Issues and applications*. London: Sage Publications.

Netemeyer, R. G., Boles, J. S., McKee, D. O., & McMurrian, R. (1997). An investigation Into the antecedents of organizational citizenship behaviors in a personal selling context. *Journal of Marketing, 61*, 85.

Neuman, A., Shahor, N., Shina, I., Sarid, A., & Saar, Z. Evaluation utilization research: Developing a theory and putting it to use. *Evaluation and Program Planning, 36*(1), 64-70. https://doi.org/https://doi.org/10.1016/j.evalprogplan.2012.06.001

Neves, P., & Eisenberger, R. (2012). Management communication and employee performance: The contribution of perceived organizational support. *Human Performance, 25*, 452.

New Zealand Council of Trade Unions. (2013). *Under pressure: A detailed report into insecure work in New Zealand.* https://www.union.org.nz/wpcontent/uploads/2016/12/CTU-Under-Pressure-Detailed-Report-2.pdf

Newman, A., Nielsen, I., Smyth, R., & Hooke, A. (2015). Examining the relationship between workplace support and life satisfaction: The mediating role of job satisfaction. *Social Indicators Research, 120*, 769-781.

Ngo, H. Y., Foley, S., Ji, M. S., & Loi, R. (2014). Work satisfaction of Chinese employees: A social exchange and gender-based view. *Social Indicators Research, 116*, 457.

Nicolai, A., & Seidl, D. (2010). That's relevant! Different forms of practical relevance in management science. *Organization Studies, 31*(9/10), 1257-1285. https://doi.org/10.1177/0170840610374401

Nishii, L., & Wright, P. (2008). Variability within organizations: Implications for strategic human resources management. In D. B. Smith (Ed.), *The people make the place: Dynamic linkages between individuals and organizations* (pp. 225-248). Lawrence Erlbaum.

Niven, P. (2014). *Balanced scorecard evolution: A dynamic approach to strategy execution.* (1st ed.) John Wiley & Sons.

Noe, R., Hollenbeck, J., Gerhart, B., & Wright, P. (2015). *Human resource management: Gaining a competitive advantage.* (9th ed.). Mcgraw-Hill Education.

Nordin, S. M., Sivapalan, S., Bhattacharyya, E., Ahmad, H. H. W. F. W., & Abdullah, A. (2014). Organizational communication climate and conflict management: Communications management in an oil and gas company. *Procedia - Social and Behavioral Sciences, 109*, 1046-1058.

North, N., Leung, W., Ashton, T., Rasmussen, E., Hughes, F., & Finlayson, M. (2013). Nurse turnover in New Zealand: Costs and relationships with staffing practises and patient outcomes. *Journal of Nursing Management, 21*(3), 419-428. https://doi.org/10.1111/j.1365-2834.2012.01371.x

NSW Industrial Relations. (n.d.). *Managing performance*. NSW Government. https://www.industrialrelations.nsw.gov.au/employers/managing-performance

Oberoi, M., & Rajgarhia, P. (2013). What your performance management system needs most. *Gallup Business Journal*, 2-2.

Odlyzko, A. (2002). The rapid evolution of scholarly communication. *Learned Publishing, 15(1), 7–19., 15*(1), 7-19.

Office for National Statistics. (2019). *Birth characteristics in England and Wales: 2018* (Statistical Bulletin) Executive Office of the UK Statistics Authority. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/bulletins/birthcharacteristicsinenglandandwales/2018#the-average-age-of-mothers-and-fathers-increased-for-the-10th-consecutive-year-in-2018

Ogunkuade, I. M. (2018). Influence of organisational affiliation, gender and length of service on the perception of organisational justice in Nigeria. *IFE PsychologIA, 26*, 9-18.

Ohemeng, F. L. K., Amoako‐Asiedu, E., & Obuobisa‐Darko, T. (2018). The impact of employee perception on the successful institutionalisation and implementation of

performance management systems in developing countries: The perspective from Ghana's public service. *Public Administration & Development, 38*, 75-86.

Olson, E. M., Slater, S. F., & Hult, G. T. M. (2005). The importance of structure and process to strategy implementation. *Business Horizons, 48*, 47-54. https://doi.org/https://doi.org/10.1016/j.bushor.2004.10.002

Op de Beeck, S., Wynen, J., & Hondeghem, A. (2016). HRM implementation by line managers: Explaining the discrepancy in HR-Line perceptions of HR devolution. *International Journal of Human Resource Management, 27*(17), 1901-1919. https://doi.org/10.1080/09585192.2015.1088562

Osborne, J. W., & Costello, A. B. (2004). Sample size and subject to item ratio in principal components analysis. *Practical Assessment, Research & Evaluation, 9*(11).

Osborne, J. W., & Costello, A. B. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, 10*(7), 1-9.

Osterman, P. (2009). Recognizing the value of middle management. *Ivey Business Journal, 73*(6). https://iveybusinessjournal.com/publication/recognizing-the-value-of-middle-management/

Ostroff, C., & Bowen, D. E. (2016). Reflections on the 2014 decade award: Is there strength in the construct of HR system strength? *Academy of Management Review, 41*(2), 196-214. https://doi.org/10.5465/amr.2015.0323

Özleblebici, Z., & Çetin, Ş. (2015). The role of managerial perception within strategic management: An exploratory overview of the literature. *Procedia - Social and Behavioral Sciences, 207*, 296-305.

Paauwe, J., & Richardson, R. (2001). Editorial introduction: HRM and performance: Confronting theory and reality. *International Journal of Human Resource Management, 12*, 1085-1091.

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance, 17*, 22-27.

Parijat, P., & Bagga, S. (2014). Victor Vroom's expectancy theory of motivation - an evaluation. *International Research Journal of Business and Management, 7*, 1-8.

Park, J. H., Newman, A., Zhang, L., Wu, C., & Hooke, A. (2016). *Mentoring functions and turnover intention: The mediating role of perceived organizational support.* Great Britain, Taylor & Francis.

Parveen, F., Jaafar, N. I., & Ainin, S. (2015). Social media usage and organizational performance: Reflections of Malaysian social media managers. *Telematics and Informatics, 32*(1), 67-78. https://doi.org/https://doi.org/10.1016/j.tele.2014.03.001

Patton, M. Q. (2000). Utilization-Focused evaluation. In D. L. Stufflebeam, G. F. Madaus, & T. Kellaghan (Eds.), *Evaluation Models*. Kluwer Academic Publishers.

Pattyn, V. (2014). Why organizations (do not) evaluate? Explaining evaluation activity through the lens of configurational comparative methods. *Evaluation, 20*(3), 348-367. https://doi.org/10.1177/1356389014540564

Pearce, J. L., & Porter, L. W. (1986). Employee responses to formal performance appraisal feedback. *Journal of Applied Psychology, 71*(2), 7.

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology, 70*, 153-163.

Peterson, R. A. (2000). Questions and answers. In R. A. Peterson (Ed.), *Constructing effective questionnaires* (pp. 1-12). Thousand Oaks, CA: SAGE Publications, Inc. https://doi.org/10.4135/9781483349022

Petrović, I. B., Vukelić, M., & Čizmić, S. (2017). Work engagement in Serbia: Psychometric properties of the serbian version of the Utrecht Work Engagement Scale (UWES). *Frontiers in Psychology, 8*(1799), 1-11. https://doi.org/10.3389/fpsyg.2017.01799

Pettinger, R. (2010). *Organisational behaviour: Performance management in perspective*. Routledge.

Pham, M. T., Rajić, A., Greig, J. D., Sargeant, J. M., Papadopoulos, A., & McEwen, S. A. (2014). A scoping review of scoping reviews: Advancing the approach and enhancing the consistency. *Research synthesis methods, 5*(4), 371-385. https://doi.org/10.1002/jrsm.1123

Piening, E. P., Baluch, A. M., & Ridder, H. G. (2014). Mind the intended-implemented gap: Understanding employees' perceptions of HRM. *Human Resource Management, 53*(4), 545-567. https://doi.org/10.1002/hrm.21605

Piening, E. P., Baluch, A. M., & Salge, T. O. (2013). The relationship between employees' perceptions of human resource systems and organisational performance: Examining mediating mechanisms and temporal dynamics. *Journal of Applied Psychology, 98*(6), 926-947. https://doi.org/10.1037/a0033925

Pihl-Thingvad, S. (2016). The inner workings of performance management in Danish job centers: Rational decisions or cowboy solutions? *Public Performance & Management Review, 40*(1), 48-70. https://doi.org/10.1080/15309576.2016.1177556

Podsakoff, N. P., LePine, J. A., & LePine, M. A. (2007). Differential challenge stressor-hindrance stressor relationships with job attitudes, turnover intentions, turnover, and withdrawal behavior: A meta-analysis. *Journal of Applied Psychology, 92*(2), 438-454.

Podsakoff, N. P., Whiting, S. W., Podsakoff, P. M., & Blume, B. D. (2009). Individual and organizational-level consequences of organizational citizenship behaviors: A meta-analysis. *Journal of Applied Psychology, 94*(1), 122-141. https://doi.org/https://doi.org/10.1037/a0013079

Polit, D. F., & Tatano Beck, C. (2017). *Essentials of nursing research: Methods, appraisal, and utilization* (9th ed.). Lipincott Williams and Wilkins.

Postmes, T., Tanis, M., & de Wit, B. (2001). Communication and commitment in organizations: A social identity approach. *Group Processes and Intergroup Relations, 4*(2), 227–246.

Profit, J., Sharek, P. J., Amspoker, A. B., Kowalkowski, M. A., Nisbet, C. C., Thomas, E. J., . . . Sexton, J. B. (2014). Burnout in the NICU setting and its relation to safety culture. *BMJ Quality & Safety, 23*(10), 806-813. https://doi.org/10.1136/bmjqs-2014-002831

Pulakos, E. D., Mueller-Hanson, R. A., & O'Leary, R. S. (2008). Performance management in the US. In A. Varma, P. S. Budhwar, & A. S. DeNisi (Eds.), *Performance Management Systems: A global perspective*. Routledge.

Purcell, J., & Hutchinson, S. (2007). Front-line managers as agents in the HRM-performance causal chain: Theory, analysis and evidence. *Human Resource Management Journal, 17*, 3-20.

Rahman, W., & Nas, Z. (2013). Employee development and turnover intention: Theory validation. *European Journal of Training and Development, 37*, 564.

Raina, R., & Roebuck, D. B. (2016). Exploring cultural influence on managerial communication in relationship to job satisfaction, organizational commitment, and the employees' propensity to leave in the insurance sector of India. *International journal of business communication, 53*, 97.

Rasheed, A., Khan, S.-U.-R., Rasheed, M. F., & Munir, Y. (2015). The impact of feedback orientation and the effect of satisfaction with feedback on in-role job performance. *Human Resource Development Quarterly, 26*, 31-51. https://doi.org/http://onlinelibrary.wiley.com/journal/10.1002/%28ISSN%291532-1096/issues

Ratcliffe, M. M. (2012). A sample theory-based logic model to improve program development, implementation, and sustainability of Farm to School programs. *Childhood Obesity, 8*, 315-322.

Rauf, A. (2016). *Measuring the effectiveness of HR Metrics on return on investment-An empirical study on pakistani organizations* IBIMA Conference on Vision 2020, Seville Spain.

Rayton, B. A., & Yalabik, Z. Y. (2014). Work engagement, psychological contract breach and job satisfaction. *International Journal of Human Resource Management, 25*, 2382-2400.

Ree, E., & Wiig, S. (2019). Employees' perceptions of patient safety culture in Norwegian nursing homes and home care services. *BMC Health Services Research, 19*(1), 607-607. https://doi.org/10.1186/s12913-019-4456-8

Reineke, R. A. (1991). Stakeholder involvement in evaluation: Suggestions for practice. *Evaluation Practice, 12*, 39-44.

Reio, T. G. (2010). The threat of common method variance bias to theory building. *Human Resource Development Review, 9*(4), 405-411. https://doi.org/https://doi.org/10.1177/1534484310380331

Reisman, J., & Gienapp, A. (2004). *Theory of change: A practical tool for action, results and learning*. Seattle, Washington.

Renger, R., Atkinson, L., Renger, J., Renger, J., & Hart, G. (2019). The connection between logic models and systems thinking concepts. *Evaluation Journal of Australasia. 19*, 79.

Reynolds, P. D., & Helfers, R. C. (2018). Differences in perceptions of organizational fairness based on job characteristics among police officers. *American Journal of Criminal Justice, 43*(2), 371-388. https://doi.org/10.1007/s12103-017-9404-8

Rhoades, L., & Eisenberger, R. (2002). Perceived organizational support: A review of the literature. *Journal of Applied Psychology*(4), 698.

Rich, B. L., Lepine, J. A., & Crawford, E. R. (2010). Job engagement: Antecedents and effects on job performance. *Academy of Management Journal, 53*(3), 617-635.

Richard, P. J., Devinney, T. M., Yip, G. S., & Johnson, G. (2009, 2009/06/01). Measuring organizational performance: Towards methodological best practice. *Journal of Management, 35*(3), 718-804. https://doi.org/10.1177/0149206308330560

Ridde, A. V., Ridde, F. V., Patton, M. Q., & Johnson, H. (2005). Utilization-Focused evaluation. In S. Mathison (Ed.), *Encyclopedia of Evaluation* (pp. 430-435). Thousand Oaks, CA: Sage Publications, Inc. https://doi.org/doi: http://dx.doi.org/10.4135/9781412950558

Riketta, M. (2002). Attitudinal organizational commitment and job performance: A meta-analysis. *Journal of Organisational Behavior, 23*(3), 257-266.

Riley, P., Zuber, A., Vindigni, S., Gupta, N., Verani, A., Sunderland, N., . . . Campbell, J. (2012). Information systems on human resources for health: A global review. *Human Resources for Health, 10*, 7.

Robbins, S., Robbins, S. P., Judge, T., Millett, B., & Boyle, M. (2016). *Organisational behaviour* (8th ed.). Pearson Education Australia. https://books.google.com.au/books?id=VK3-jwEACAAJ

Robbins, S. P., Judge, T. A., Odendaal, A., & Roodt, G. (2013). *Organisational behaviour: Global and Southern African perspectives*. Pearson.

Robertson, C. (2016). *Cost reduction: Bridging the gap Making cost savings real and making them stick*. Deloitte Canada. https://www2.deloitte.com/content/dam/Deloitte/ca/Documents/finance/ca-en-FA-cost-reduction-POV.PDF

Robinson, S. L., Wang, W., & Kiewitz, C. (2014). Coworkers behaving badly: The impact of coworker deviant behavior upon individual employees. *Annual Review of Organizational Psychology and Organizational Behavior, 1*, 123-143. https://doi.org/10.1146/annurev-orgpsych-031413-091225

Rogers, P. (2014). *Theory of change* (Methodological Briefs: Impact Evaluation, Issue 2). UNICEF Office of Research.

Rogers, P., & Weiss, C. (2007). Theory-based evaluation: Reflections ten years on: Theory-based evaluation: Past, present, and future. *New Directions for Evaluation, 2007*, 63-81. https://doi.org/10.1002/ev.225

Roongrerngsukea, S., & Liefooghe, A. (2013). Attracting gold-collar workers: Comparing organizationalattractiveness and work-related values across generations in China, India and Thailand. *Asia Pacific Business Review, 19*(3), 337-335. https://doi.org/http://dx.doi.org/10.1080/13602381.2012.747784

Rothbard, N. P. (2001). Enriching or depleting? The dynamics of engagement in work and family roles. *Administrative Science Quarterly, 46*, 655-684. https://doi.org/10.2307/3094827

Rousseau, D. (2012). *The Oxford handbook of evidence-based management*. Oxford University Press.

Rousseau, D., & Barends, E. (2011). Becoming an evidence-based HR practitioner. *Human Resource Management Journal, 21*, 221-235.

Rousseau, D., & Boudreau, J. (2011). Sticky findings: Research evidence practitioners find useful. In S. A. Mohrman & E. E. Lawler III (Eds.), *Useful research: Advancing theory and practice* (pp. 269-288). Berrett-Koehler.

Rousseau, D., & McCarthy, S. (2007). Educating managers from an evidence-based perspective. *Academy of Management Learning & Education, 6*, 84-101. https://doi.org/10.5465/AMLE.2007.24401705

Rynes, S. L., & Bartunek, J. M. (2017). Evidence-Based management: Foundations, development, controversies and future. *Annual Review of Organizational Psychology and Organizational Behavior, 4*, 235-261. https://doi.org/10.1146/annurev-orgpsych-032516-113306

Rynes, S. L., Brown, K. G., & Colbert, A. E. (2002). Seven common misconceptions about human resource practices: Research findings versus practitioner beliefs. *Academy of Management Executive, 16*(3), 92-103. https://doi.org/10.5465/AME.2002.8540341

Rynes, S. L., Giluk, T. L., & Brown, K. G. (2007). The very separate worlds of academic and practitioner periodicals in human resource management: Implications for evidence-based management *Academy of Management Journal, 50*, 987-1008. https://doi.org/10.5465/AMJ.2007.27151939

Sachane, M., Bezuidenhout, A., & Botha, C. (2018). Factors that influence employee perceptions about performance management at statistics South Africa. *South African Journal of Human Resource Management, 16*, e1-e8. https://doi.org/10.4102/sajhrm.v16i0.986

Sackett, D. L., Rosenberg, W. M., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. *British Medical Journal, 312, 71–72., 312*, 2.

Sagie, A., Birati, A., & Tziner, A. (2002). Assessing the costs of behavioral and psychological withdrawal: A new model and an empirical illustration. *Applied Psychology: An International Review, 51*(1), 67.

Salleh, M., Amin, A., Muda, S., & Halim, M. A. S. A. (2013). Fairness of performance appraisal and organizational commitment. *Asian Social Science, 9*, 121-128.

Sanders, K., Dorenbosch, L., & Reuver, R. d. (2008). The impact of individual and shared employee perceptions of HRM on affective commitment: Considering climate strength. *37*(4), 412-425. https://doi.org/10.1108/00483480810877589

Sarangi, S., & Srivastava, R. K. (2012). Impact of organizational culture and communication on employee engagement: An investigation of Indian private banks. *South Asian Journal of Management, 19*, 18-33.

Schaubroeck, J., Ganster, D., Sime, W. E., & Ditman, D. (1993). A field experiment testing supervisor role clarification. *Personnel Psychology, 46*(1), 1-25.

Schaufeli, W. B., & Bakker, A. B. (2010). Defining and measuring work engagement: Bringing clarity to the concept. In A. B. Bakker & M. P. Leiter (Eds.), *Work engagement: A handbook of essential theory and research* (pp. 10-24). Taylor & Francis Ltd.

Schaufeli, W. B., Salanova, M., González-Romá, V., & Bakker, A. B. (2002). The measurement of engagement and burnout: A two sample confirmatory factor analytic approach. *Journal of Happiness Studies, 3*, 71-92.

Schermelleh-Engel, K., Moosbrugger, H., & Muller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8*, 23-74.

Schminke, M., Ambrose, A. L., & Cropanzano, R. S. (2000). The effect of organizational structure on perceptions of procedural fairness. *Journal of Applied Psychology, 85*, 294–304.

Schwarzman, J., Bauman, A., Gabbe, B., Rissel, C., Shilton, T., & Smith, B. J. (2018). Organizational determinants of evaluation practice in Australian prevention agencies. *Health Education Research, 33*(3), 243-255. https://doi.org/10.1093/her/cyy015

Scriven, M. (1974). Evaluation perspectives and procedures. Evaluation in Education, California.

Seiden, S., & Sowa, J. E. (2011). Performance management and appraisal in human service organizations: Management and staff perspectives. *Public Personnel Management, 40*(3), 251-264.

Shadish, W. R., & Luellen, J. K. (2011). History of evaluation. In S. Mathison (Ed.), *Encyclopedia of Evaluation* (pp. 184-186). Sage Publications, Inc.

Sharma, N. P., Sharma, T., & Agarwal, M. N. (2016). Measuring employee perception of performance management system effectiveness: Conceptualization and scale development. *Employee Relations, 38*, 224-247. https://doi.org/10.1108/ER-01-2015-0006

Sharma, P., Lampley, J., & Good, D. (2015). Organizational communication: Perceptions of staff members' level of communication satisfaction and job satisfaction. *Journal of Academic Administration in Higher Education, 11*, 43-54.

Shields, J. (2007). *Managing employee performance and reward: Concepts, practices, strategies.* Cambridge University Press.

Shih, H. A., Chiang, Y. H., & Hsu, C. C. (2010). High involvement work system, work-family conflict, and expatriate performance - Examining Taiwanese expatriates in China. *International Journal of Human Resource Management, 21*(11), 2013-2030. https://doi.org/10.1080/09585192.2010.505101

Shkoler, O., Rabenu, E., Vasiliu, C., Sharoni, G., & Tziner, A. (2017). Organizing the confusion surrounding workaholism: New structure, measure, and validation. *Frontiers in Psychology, 8*(1803). https://www.frontiersin.org/article/10.3389/fpsyg.2017.01803

Shulha, L. (2010, November). *Participatory evaluation essentials: Highlighting the essential features of participatory evaluation.* American Evaluation Association, San Antonio, Texas.

Singh, S. (2000). Relationship between managers' authority power and perception of their subordinates' behaviour. *Indian Journal of Industrial Relations, 35*(3), 275-300.

Sivo, S. A., Fan, X., Witta, E. L., & Willse, J. T. (2006). The search for "optimal" cutoff properties: Fit index criteria in structural equation modeling. *The Journal of Experimental Education, 74*(3), 267-288. http://www.jstor.org.ezp.lib.unimelb.edu.au/stable/20157428

Skinner, D. O. (2001). *Barriers to the evaluation of human resource management initiatives: Three public sector case studies* [Open University].

Smet, K., Vander Elst, T., Griep, Y., & De Witte, H. (2016). *The explanatory role of rumours in the reciprocal relationship between organizational change communication and job insecurity: A within-person approach* Great Britain, Taylor & Francis.

Smith, J. R., Micich, L. A., & McWilliams, D. L. (2016). Organization citizenship and employee withdrawal behavior in the workplace. *International Journal of Human Resource Studies, 6*(3), 43-62.

Smith, S. D., & Galbraith, Q. (2012). Motivating millennials: Improving practices in recruiting, retaining, and motivating younger library staff. *Journal of Academic Librarianship, 38*(3), 135-144.

Snyder, M., & Stukas, J. A. (2007). Interpersonal processes in context: understanding the influence of settings and situations on social interaction. In K. Fiedler (Ed.), *Frontiers of social psychology: Social communication* (pp. 363 - 388). Psychology Press.

Society for Human Resource Management. (2016). *Advanced analytics: Using data to drive HR excellence* (SHRM HR Metrics, Issue 2016). https://www.shrm.org/LearningAndCareer/learning/Documents/SHRM%20HR%20Metrics_AA.pdf

Society for Human Resource Management. (2017). *Performance management that makes a difference: An evidence-based approach* (SHRM Science-to-Practice) https://www.shrm.org/hr-today/trends-and-forecasting/special-reports-and-expert-views/documents/performance%20management.pdf

Society for Human Resource Management. (n.d.). *About*. https://www.shrm.org/about-shrm/Pages/default.aspx

Society for Human Resource Management, & Kronos. (2014). *Total financial impact of employee absences across the United States, China, Australia, Europe, India and Mexico*. https://www.shrm.org/hr-today/trends-and-forecasting/research-and-surveys/Documents/Total%20Financial%20Impact%20of%20Employee%20Absences%20Report.pdf

Solansky, S. T. (2010). The evaluation of two key leadership development program components: Leadership skills assessment and leadership mentoring. *The Leadership Quarterly, 21*, 675-681. https://doi.org/http://dx.doi.org/10.1016/j.leaqua.2010.06.009

Solberg, E., & Dysvik, A. (2016). Employees' perceptions of HR investment and their efforts to remain internally employable: Testing the exchange-based mechanisms of the 'new psychological contract'. *International Journal of Human Resource Management, 27*(9), 909-927.

Soldan, Z., & Nankervis, A. (2014). Employee perceptions of the effectiveness of diversity management in the Australian public service: Rhetoric and reality. *Public Personnel Management, 43*, 543-564. https://doi.org/10.1177/0091026014533093

Soltani, E. (2005). Conflict between theory and practice: TQM and performance appraisal. *International Journal of Quality and Reliability Management, 22*, 796-818.

Songstad, N. G., Lindkvist, I., Moland, K. M., Chimhutu, V., & Blystad, A. (2012). Assessing performance enhancing tools: experiences with the open performance review and appraisal system (OPRAS) and expectations towards payment for performance (P4P) in the public health sector in Tanzania. *Globalization and Health, 8*.

Southall, A. (2014). *The discretionary role of the line manager in inspiring work engagement and innovative behaviour: A study of social exchange and job resources in the public sector* [Aston University]. edsble.

Srinivasan, V. (2012, 2012/03/01/). Multi generations in the workforce: Building collaboration. *IIMB Management Review, 24*(1), 48-66. https://doi.org/https://doi.org/10.1016/j.iimb.2012.01.004

Stake, R. E. (2011). Program evaluation, particularly responsive evaluation. *Journal of MultiDisciplinary Evaluation, 7*, 287-310.

Steele, G. A., & Plenty, D. (2015). Supervisor—Subordinate communication competence and job and communication satisfaction. *International Journal of Business Communication, 52*(3), 294-318. https://doi.org/https://doi.org/10.1177/2329488414525450

Stetler, C. B., Ritchie, J. A., Rycroft-Malone, J., & Charns, M. P. (2014). Leadership for evidence-based practice: Strategic and functional behaviors for institutionalizing EBP. *Worldviews on Evidence-Based Nursing*, 219. https://doi.org/10.1111/wvn.12044

Stevahn, L., King, J. A., Ghere, G., & Minnema, J. (2005, 01/01/). Establishing essential competencies for program evaluators. *American Journal of Evaluation, 26*(1), 43-59.

Storey, J. (2001). Human resource management today: An assessment. In J. Storey (Ed.), *Human Resource Management: A critical text* (2nd ed., pp. 3-20). Thompson Learning.

Strebler, M. (2001). Overloaded performance appraisals failing. *Manager: British Journal of Administrative Management*, 8.

Strebler, M. T., Bevan, S., & Robinson, D. (2001). *Performance review: Balancing objectives and content* (Report 370). Institute for Employment Studies.

Stufflebeam, D. L., Madaus, G. F., & Kellaghan, T. (2002). *Evaluation models: Viewpoints on educational and human services evaluation* (2nd ed.). Kluwer Academic Publishers.

Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology.* (1st ed.). Jossey-Bass Publishers.

Sun, L.-Y., Aryee, S., & Law, K. S. (2007). High-Performance human resource practices, citizenship behavior, and organizational performance: A relational perspective. *The Academy of Management Journal, 50*(3), 558-577. https://doi.org/10.2307/20159873

Switzer, G. E., Wisniewski, S. R., Belle, S. H., Dew, M. A., & Schultz, R. (1999). Selecting, developing, and evaluating research instruments. *Social Psychiatry & Psychiatric Epidemiology, 34*, 399.

Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics* (6th, Ed.). Harlow, Essex: Pearson New International Edition.

Taghipour, A., & Dezfuli, Z. K. (2013). Designing and testing a model of antecedents of work engagement. *Procedia - Social and Behavioral Sciences, 84*, 149-154.

Takeuchi, N., & Takeuchi, T. (2013). Committed to the organization or the job? Effects of perceived HRM practices on employees' behavioral outcomes in the Japanese healthcare industry. *International Journal of Human Resource Management, 24*(11), 2089-2106. https://doi.org/10.1080/09585192.2013.767059

Takeuchi, R., Lepak, D. P., Wang, H., & Takeuchi, K. (2007). An empirical examination of the mechanisms mediating between high-performance work systems and the performance of Japanese organisations. *Journal of Applied Psychology, 92*(4), 1069–1083. https://doi.org/10.1037/0021-9010.92.4.1069

Teclaw, R., Osatuke, K., Fishman, J., Moore, S. C., & Dyrenforth, S. (2014). Employee age and tenure within organizations: Relationship to workplace satisfaction and workplace

climate perceptions. *Health Care Manager, 33*(1), 4-19. https://doi.org/10.1097/01.HCM.0000440616.31891.2d.

Terpstra, D. E., & Rozell, E. J. (1997). Sources of human resource information and the link to organizational profitability. *The Journal of Applied Behavioral Science, 33*(1), 66.

Thakur, S. K. (2011). Fundamental economic structure and structural change in regional economies: A methodological approach. *Region et Developpement, 33*, 9-38. https://doi.org/http://regionetdeveloppement.org/

The Evaluation Center. (2014). *Connecting logic models to data you can use.* Youtube. https://www.youtube.com/watch?v=Cjpm06n07Tg

Thompson, A. A., & Strickland, A. J. (1999). *Strategic management:Concepts and cases.* (11th ed.). McGraw Hill.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications.* (1st ed.). American Psychological Association,. https://doi.org/https://doi.org/10.1037/10694-000

Thompson, P., & Harley, B. (2007). HRM and the worker: Labour process perspective. In P. Boxall, P. Purcell, & P. Wright (Eds.), *Oxford handbook of human resource management* (pp. 147-165). Oxford University Press.

Tomás, J. M., Santos, S. d. l., Georgieva, S., & Enrique, S. (2018). Utrecht work engagement scale in Dominican teachers: Dimensionality, reliability, and validity. *Escala de Utrecht de implicación en el trabajo en los profesores dominicanos: dimensionalidad, fiabilidad y validez., 34*, 89-93.

Tracey, J. B. (2010). Human resources strategy in the hospitality industry: Where do we go from here? *School of Hotel Administration Collection*, 16. Retrieved 2019, from http://scholarship.sha.cornell.edu/articles/957

Tracey, J. B., & Hinkin, T. R. (2008). Contextual factors and cost profiles associated with employee turnover. *Cornell Hospitality Quarterly, 49*(1), 12-27. https://doi.org/10.1177/0010880407310191

Travis, D. J., Lizano, E. L., & Mor Barak, M. E. (2016). 'I'm So Stressed!': A longitudinal model of stress, burnout and engagement among social workers in child welfare settings. *British Journal of Social Work, 46*(4), 1076-1095. https://doi.org/10.1093/bjsw/bct205

Tremblay, M., Cloutier, J., Simard, G., Chênevert, D., & Vandenberghe, C. (2010). The role of HRM practices, procedural justice, organizational support and trust in organizational commitment and in-role and extra-role performance. *International Journal of Human Resource Management, 21*, 405-433.

Tsai, M.-T., Chuang, S.-S., & Hsieh, W.-P. (2009). An integrated process model of communication satisfaction and organizational outcomes. *Social Behavior & Personality: an international journal, 37*, 825-834.

Tsai, Y., & Wu, S.-W. (2010). The relationships between organisational citizenship behaviour, job satisfaction and turnover intention. *Journal of Clinical Nursing, 19*, 3564-3574. https://doi.org/10.1111/j.1365-2702.2010.03375.x

Tubre, T. C., & Collins, J. M. (2000). Jackson and Schüler (1985) revisited: A meta-analysis of the relationships between role ambiguity, role conflict, and job performance. *Journal of Management, 26*(1), 155-169.

Turunc, Ö., & Altuntepe, N. (2015). How do the perceived organizational support and the leadership support affect entrepreneurship behaviours? The mediating role of LMX. *Algılanan Örgütsel Destek ve Lider desteği Girişimci Davranışları nasıl Etkiler? Lider - Üye Etkileşiminin Aracılık Rolü., 7*, 33-47.

Twenge, J. M. (2010). A review of the empirical evidence on generational differences in work attitudes. *Journal of Business and Psychology, 25*, 201-210. https://doi.org/https://doi.org/10.1007/s10869-010-9165-6

U.S. Bureau of Labor Statistics. (2018). *Table 5. Median years of tenure with current employer for employed wage and salary workers by industry, selected years, 2008-2018*. U.S. Bureau of Labor Statistics. https://www.bls.gov/news.release/tenure.t05.htm

Ullman, J. B. (2006). Structural equation modeling: Reviewing the basics and moving forward. *Journal of Personality Assessment, 87*(1), 35-50. https://doi.org/10.1207/s15327752jpa8701_03

Uslu, T., & Çubuk, D. (2015). The effects of knowledge management and self-organization on organizational creativity: The mediating roles of corporate innovativeness and organizational communication. *International Journal of Organizational Leadership*, 403.

Üstüner, M., & Kis, A. (2014). The relationship between communication competence and organizational conflict: A study on heads of educational supervisors. *Eurasian Journal of Educational Research, 14*(56), 23-44.

van Aken, J., & Romme, A. (2012). A design science approach to evidence-based management. In D. M. Rousseau (Ed.), *The Oxford Handbook of Evidence-Based Management*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199763986.013.0003

van Dijk, H. G., & Legalatladi, N. M. V. (2015). Determining employee perceptions towards performance management and development. *African Journal of Public Affairs, 8*(2), 16. https://pdfs.semanticscholar.org/5630/b8139f55c0ed2d5dc431bfa5363c47558201.pdf

van Dijke, M., Langendijk, G., De Cremer, D., & Anderson, C. (2018). Ranking low, feeling high: How hierarchical position and experienced power promote prosocial behavior in response to procedural justice. *Journal of Applied Psychology, 103*, 164-181.

van Koperen, T. M., Renders, C. M., Spierings, E. J. M., Hendriks, A.-M., Westerman, M. J., Seidell, J. C., & Schuit, A. J. (2016). Recommendations and improvements for the evaluation of integrated community-wide anterventions approaches. *Journal of Obesity*, 1-13.

Veld, M., & Alfes, K. (2017). HRM, climate and employee well-being: Comparing an optimistic and critical perspective. *International Journal of Human Resource Management, 28*(16), 2299-2318. https://doi.org/10.1080/09585192.2017.1314313

Venz, L., & Sonnentag, S. (2015, 12/1/December 2015). Being engaged when resources are low: A multi-source study of selective optimization with compensation at work. *Journal of Vocational Behavior, 91*, 97-105. https://doi.org/10.1016/j.jvb.2015.09.008

Vera, M., Martínez, I., Lorente, L., & Chambel, M. (2016). The role of co-worker and supervisor support in the relationship between job autonomy and work engagement among portuguese nurses: A multilevel study. *Social Indicators Research, 126*, 1143-1156.

Vermeeren, B. (2014). Variability in HRM implementation among line managers and its effect on performance: A 2-1-2 mediational multilevel approach. *International Journal of Human Resource Management, 25*(22), 3039-3059. https://doi.org/10.1080/09585192.2014.934891

Vermeeren, B., Steijn, B., Tummers, L., Lankhaar, M., Poerstamper, R.-J., & van Beek, S. (2014, 2014/06/17). HRM and its effect on employee, organizational and financial outcomes in health care organizations. *Human Resources for Health, 12*(1), 35. https://doi.org/10.1186/1478-4491-12-35

Vermeulen, F. (2007). "I shall not remain insignificant": Adding a second loop to matter more. *Academy of Management Journal, 50*(4), 754-761.

Viljevac, A., Cooper-Thomas, H., & Saks, A. (2012). An investigation into the validity of two measures of work engagement. *International Journal of Human Resource Management, 23*, 3692-3709.

Villajos, E., Tordera, N., Peiró, J. M., & van Veldhoven, M. (2018). Refinement and validation of a comprehensive scale for measuring HR practices aimed at performance-enhancement and employee-support. *European Management Journal*. https://doi.org/10.1016/j.emj.2018.10.003

Vlastarakos, P. V., & Nikolopoulos, T. P. (2008). The interdisciplinary model of hospital administration: Do health professionals and managers look at it in the same way? *European Journal of Public Health, 18*, 71-76. https://doi.org/10.1093/eurpub/ckm042

Vnouckova, L., & Urbancova, H. (2016). Staff turnover costs in the Czech Republic. *Ekonomicka Revue: Central European Review of Economic Issues, 19*(2), 37-51. https://doi.org/http://www.ekf.vsb.cz/cerei/en/

Vosko, L. F., MacDonald, M., & Campbell, I. (2009). *Gender and the contours of precarious employment* (L. F. Vosko, M. MacDonald, & I. Campbell, Eds. Vol. 8). Routledge.

Vroom, V. H. (1964). *Work and motivation*. Wiley.

W.K. Kellogg Foundation. (2004). *Logic model development guide*. W.K. Kellogg Foundation. https://www.bttop.org/sites/default/files/public/W.K.%20Kellogg%20LogicModel.pdf

Wagenheim, M., & Rood, A. S. (2010). The relationship between employee satisfaction with organizational communication and customer orientation. *Managing Leisure. 15*, 83.

Walden, J., Jung, E. H., & Westerman, C. Y. K. (2017). *Employee communication, job engagement, and organizational commitment: A study of members of the Millennial Generation.* Great Britain, Taylor & Francis.

Walk, M., Handy, F., & Schinnenburg, H. (2013). Expectations and experiences of young employees: The case of German nonprofits. *Administration in Social Work, 37*(2), 133-146. https://doi.org/10.1080/03643107.2012.667658

Walker, S. K. (2017). *Retention Strategies for Reducing Voluntary Turnover in a Higher Education Institution* (Publication Number 10619260) [Doctor of Business Administration, Walden University]. ProQuest Dissertations & Theses Global. Ann Arbor.

Walker, W. (2005). The strengths and weaknesses of research designs involving quantitative measures. *Journal of Research in Nursing, 10*(5), 571-582.

Walton, R. E. (1985). From control to commitment in the workplace. *Harvard Business Review, 63*, 77.

Wang, Y., Kim, S., Rafferty, A., & Sanders, K. (2020). Employee perceptions of HR practices: A critical review and future directions. *International Journal of Human Resource Management, 31*(1), 128-173. https://doi.org/10.1080/09585192.2019.1674360

Weber, P. S., & Weber, J. E. (2001). Changes in employee perceptions during organizational change. *Leadership & Organization Development Journal, 22*(6), 291-300. https://doi.org/10.1108/01437730110403222

Weiss, C. H. (1993). Where politics and evaluation research meet. *Evaluation Practice, 14*(1), 93-106. https://doi.org/10.1016/0886-1633(93)90046-R

Weiss, C. H. (1997a). How can theory-based evaluation make a greater headway? *Evaluation Review, 21*, 501.

Weiss, C. H. (1997b). Theory-Based evaluation: Past, present, and future. *New Directions for Evaluation, 1997*, 41-55. https://doi.org/10.1002/ev.1086

Werner, R. (2012). Buried treasure: A business librarian's insights on finding the evidence. In D. M. Rousseau (Ed.), *The Oxford Handbook of Evidence-Based Management* Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199763986.013.0015

West, D., & Blackman, D. (2015). Performance management in the public sector. *Australian Journal of Public Administration, 74*(1), 73-81. https://doi.org/10.1111/1467-8500.12130

Wetzel, A. (2012). Factor analysis methods and validity evidence: A review of instrument development across the medical education continuum. Academic Medicine, *87*(8), 1060-1069. https://doi.org/10.1097/ACM.0b013e31825d305d

Whalley, C. E. (2000). *Issues in managing performance the manager's tale* [Doctor of Education, Brunel University].

Whetstone, M. L. (2017). *A quantitative study of U.S. federal government senior executive online learning satisfaction, Job satisfaction, and turnover intent* (Publication Number 10743472) [Doctor of Philosophy, Northcentral University]. ProQuest LLC. San Diego, California.

Wikimedia Foundation Inc. (n.d.). *WikiProject Business*. https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Business

Wilbourne, S. (2001). The expanding ocean of difference between academic research and book publishing. *Publishing Research Quarterly, 17*(3), 29-32. https://doi.org/10.1007/s12109–001-0030–3

Wildschut, L. P. (2014). *Theory-based evaluation, logic modelling and the experience of SA non-governmental organisations* [Doctor of Philosophy, Faculty of Arts and Social Sciences, Stellenbosch University].

Wilson, V. (2014). Research methods: Scoping studies. *Evidence Based Library & Information Practice, 9*, 97-99.

Winterton, J. (2004). A conceptual model of labour turnover and retention. *Human Resource Development International, 7*(3), 371-390. https://doi.org/10.1080/1367886042000201967

Witt, L. A., Andrews, M. C., & Kacmar, K. M. (2000). The role of participation in decision-making in the organizational politics-job satisfaction relationship. *Human Relations, 53*(3), 341-358.

https://ezp.lib.unimelb.edu.au/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=sih&AN=12135735&site=eds-live&scope=site

Woldearegay, A. G., & Woldearegay, A. G. (2013). Development and validation of a hybrid measure of organisational communication satisfaction. *Communicare: Journal for Communication Sciences in Southern Africa, 34*(1), 1 - 26

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement, 76*(6), 913-934. https://doi.org/10.1177/0013164413495237

Workplace Gender Equality Agency. (2016). *Unpaid care work and the labour market*. Australian Government. https://www.wgea.gov.au/sites/default/files/documents/australian-unpaid-care-work-and-the-labour-market.pdf

Workplace Gender Equality Agency. (2018). *Gender and negotiation in the workplace*. https://www.wgea.gov.au/sites/default/files/documents/Negotiation-Paper-Final.pdf

World Federation of People Management Associations. (n.d.). *Research*. https://www.wfpma.com/research

WorldatWork, & Sibson Consulting. (2010). *2010 study on the state of performance management*. https://www.worldatwork.org/docs/research-and-surveys/2010-study-on-the-state-of-performance-management.pdf

Worley, J. A. (2006). *A factor analytic study to evaluate the structure of the survey of perceived organizational support* [Graduate College of the Oklahoma State University]. Oklahoma State University.

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist, 34*, 806-838. https://doi.org/10.1177/0011000006288127

Wright, P., & Boswell, W. (2002). Desegregating HRM: A review and synthesis of micro and macro human resource management research. *Journal of Management, 28*, 247-276. https://doi.org/10.1016/S0149-2063(02)00128-9

Wright, P., & McMahan, G. (1992). Theoretical perspectives for strategic human resource management. *Journal of Management, 18*(2), 295-320.

Wright, P., McMahan, G., & McWilliams, A. (1994). Human resources and sustained competitive advantage: A resource-based perspective. *International Journal of Human Resource Management, 5*(301–326).

Wright, P., & Nishii, L. (2013a). HRM and performance: The role of effective implementation. In D. G. J. Paauwe, & P. Wright (Ed.), *HRM, & performance: Achievements, & challenges*. John Wiley & Sons.

Wright, P., & Nishii, L. (2013b). Strategic HRM and organizational behaviour: Integrating multiple levels of analysis. In D. Guest, J. Paauwe, & J. Wright (Eds.), *Human resource management and performance: Progress and prospects* (pp. 97-110). Wiley-Blackwell.

Xanthopoulou, D., Bakker, A. B., Kantas, A., & Demerouti, E. (2012). Measuring burnout and work engagement: Factor structure, invariance, and latent mean differences across Greece and the Netherlands. *International Journal of Business Science and Applied Management, 7*(2), 40-53.

Yakın, M., & Erdil, O. (2012). Relationships between self-efficacy and work engagement and the effects on job satisfaction: A survey on certified public accountants. *Procedia - Social and Behavioral Sciences, 58*, 370-378.

Yamaguchi, I. (2009). Influences of organizational communication tactics on trust with procedural justice effects: A cross-cultural study between Japanese and American workers. *International Journal of Intercultural Relations 33*, 21-31.

Yamaguchi, I. (2017). Instruments for organizational communication assessment for Japanese care facilities. *Corporate Communications: An International Journal, 22*, 471-485.

Yamamoto, H. (2013). The relationship between employees' perceptions of human resource management and their retention: From the viewpoint of attitudes toward job-specialties. *International Journal of Human Resource Management, 24*(4), 747-767. https://doi.org/10.1080/09585192.2012.697478

Yildiz, K. (2013). Analysis of the relation of teachers' organizational identification and organizational communication. *Educational Sciences: Theory and Practice, 13*, 264-272.

Youngcourt, S. S., Leiva, P. I., & Jones, R. G. (2007). Perceived purposes of performance appraisal: Correlates of individual-and position-focused purposes on attitudinal outcomes. *Human Resource Development Quarterly, 18*(3), 315-343. https://doi.org/10.1002/hrdq.1207

Yousaf, A., Sanders, K., & Yustantio, J. (2018). High commitment HRM and organizational and occupational turnover intentions: The role of organizational and occupational commitment. *International Journal of Human Resource Management, 29*(10), 1661-1682. https://doi.org/10.1080/09585192.2016.1256905

Yu, C., & Frenkel, S. J. (2013). Explaining task performance and creativity from perceived organizational support theory: Which mechanisms are more important? *Journal of Organizational Behavior, 34*(8), 1165-1181. https://doi.org/https://doi.org/10.1002/job.1844

Yu, T., & Richardson, J. C. (2015). An exploratory factor analysis and reliability analysis of the student online learning readiness (SOLR) instrument. *Journal of Asynchronous Learning Network, 19*(5), 120-141. https://doi.org/10.24059/olj.v19i5.593

Yunos, F., Johar, F., & Sabri, S. (2015). Planners' perception on factors that affects plan implementation in Iskandar Malaysia. *International Journal of Built Environment and Sustainability, 2*(1), 1-6. https://doi.org/10.11113/ijbes.v2.n1.51

Zanardelli, J. (2012). At the intersection of the academy and practice at asbury heights. In D. M. Rousseau (Ed.), *The Oxford Handbook of Evidence-Based Management.* Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199763986.013.0011

Zhang, J., & Liu, Y. (2010). Organizational climate and its effects on organizational variables: An empirical study. *International Journal of Psychological Studies, 2*(2), 189-201. https://doi.org/10.5539/ijps.v2n2p189

Zhong, L., Wayne, S. J., & Liden, R. C. (2016). Job engagement, perceived organizational support, high-performance human resource practices, and cultural value orientations: A cross-level investigation. *Journal of Organizational Behavior, 37*, 823-844.

Zhou, J., & George, J. M. (2001). When job dissatisfaction leads to creativity: Encouraging the expression of voice. *The Academy of Management Journal, 44*, 682.

Zhou, Y., Lu, J., Liu, X., Zhang, P., & Chen, W. (2014). Effects of core self-evaluations on the job burnout of nurses: The mediator of organizational commitment. *Plos One, 9*.

Zhu, Y. (2013). Individual behavior: In-role and extra-role. *International Journal of Business Administration*, *4*(1), 23. https://doi.org/10.5430/ijba.v4n1p23

Ziegler, R., & Schlett, C. (2016). An attitude strength and self-perception framework regarding the bi-directional relationship of job satisfaction with extra-role and in-role behavior: The doubly moderating role of work centrality. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.00235

Zopiatis, A., Krambia-Kapardis, M., & Varnavas, A. (2012). Y-ers, X-ers and Boomers: Investigating the multigenerational (mis)perceptions in the hospitality workplace. *Tourism and Hospitality Research, 12*(2), 101-121. www.jstor.org/stable/43498600

Zwijze-Koning, K., & de Jong, M. (2007). Evaluating the communication satisfaction questionnaire as communication audit tool. *Management Communication Quarterly, 20*(3), 261-282. https://doi.org/10.1177/0893318906295680

Yousaf, A., Sanders, K., & Yustantio, J. (2018). High commitment HRM and organizational and occupational turnover intentions: The role of organizational and occupational commitment. *International Journal of Human Resource Management, 29*(10), 1661-1682. https://doi.org/10.1080/09585192.2016.1256905

Appendix A: Scoping Study Results Charting

#	Author & Date	Country	Relevant Construct	Sample	Main Analysis
1	(Abu Bakar & Mustaffa, 2013)	Malaysia	COM	596	Factor
2	(Akan et al., 2016)	Turkey	COM	406	Factor, Multiple regression
3	(Alfes, Shantz, et al., 2013)	UK	HRM	297	Factor, Hierarchical multiple regression
4	(Alfes, Shantz, et al., 2013)	UK	HRM, ENG, POS	1,796	Hierarchical multiple regression
5	(Ang et al., 2017)	Australia	HRM	251	Multiple linear regression
6	(Balducci et al., 2010)	Italy & Netherlands	ENG	2881	Factor
7	(Baluch et al., 2013)	UK	HRM	167	Structural equation modelling
8	(Barbars, 2016)	Latvia	ENG	393	Correlation, Regression
9	(Bednall et al., 2013)	Netherlands	HRM	238	Factor, Hierarchical linear regression
10	(Behson, 2005)	USA	PSS	2248	Budescu's dominance analysis and Johnson's relative weight procedure
11	(Biswas et al., 2013)	India	ENG, POS	238	Recursive path model
12	(Bjørnstad et al., 2016)	Norway	POS	565	Multiple linear and logistic regression
13	(Boon et al., 2011)	Netherlands	HRM	412	Factor, Hierarchical linear regression
14	(Byrne et al., 2016)	US	ENG	633	Factor
15	(Caesens et al., 2016)	Belgium	POS	1963	Factor, Regression
16	(Carriere & Bourque, 2009)	Canada	COM	468	Multiple regression
17	(Castanheira & Story, 2016)	Portugal	HRM	117	Path
18	(Chan & Lai, 2017)	China	COM	294	Factor, Regression
19	(Chen et al., 2011)	China	HRM	984	Regression
20	(Conway & Monks, 2009)	Ireland	HRM	288	Hierarchical linear regression
21	(Dalal et al., 2012)	US	ENG	191	Univariate and Multivariate relative-weight
22	(De Nobile & McCormick, 2008)	AUS	COM	356	Factor, Multiple regression
23	(De Nobile, 2017)	AUS	COM	358	Factor, Multiple regression
24	(De Winne et al., 2013)	Belgium	HRM	1821	Multiple linear regression
25	(Delmotte et al., 2012)	Belgium	HRM	1385	Structural equation modelling
26	(Demir, 2015)	Turkey	POS	169	Factor

#	Author & Date	Country	Relevant Construct	Sample	Main Analysis
27	(Den Hartog et al., 2013)	Netherlands	HRM	2512	Multilevel structural equation modeling
28	(Dorenbosch et al., 2006)	Netherlands	HRM	671	Hierarchical linear regression
29	(Downs & Hazen, 1977)	USA	COM	691	Factor
30	(Edgar & Geare, 2005)	New Zealand	HRM	626	Multiple regression
31	(Edgar & Geare, 2014)	New Zealand	HRM	102	Multiple linear regression
32	(Ehrnrooth & Bjrkman, 2012)	Sweden	HRM	152	Structural equation modelling
33	(Eisenberger et al., 1986)	USA	POS	458	Factor
34	(Eldor & Harpaz, 2016)	Israel	ENG	625	Factor, Multilevel regression
35	(Erdogan et al., 2015)	Turkey	POS	219	Factor
36	(Frenkel & Bednall, 2016)	Australia	HRM	201	Structural equation modelling
37	(Frenkel, Li, et al., 2012)	Australia	HRM	618	Structural equation modelling
38	(Frenkel, Restubog, et al., 2012)	China	HRM	199	Structural equation modelling
39	(Gavino et al., 2012)	US	HRM	344	Factor
40	(Gillet et al., 2013)	France	POS	735	Structural equation modelling
41	(Guglielmi et al., 2016)	Italy	ENG	556	Hierarchical multiple regression
42	(Gyekye & Salminen, 2009)	Ghana	POS	320	Pearson correlation
43	(Harris et al., 2013)	US	POS	219	Hierarchical moderated regression
44	(Hassan et al., 2011)	Pakistan	COM	160	Pearson correlation
45	(He et al., 2014)	Vietnam	POS	502	Factor, Moderated structural equation modelling
46	(Heritage et al., 2016)	AUS	ENG	246	Hierarchical multiple regression
47	(Herrbach et al., 2009)	France	POS	514	Hierarchical linear regression
48	(Iqbal & Hashmi, 2015)	Pakistan	POS	170	Factor, Regression
49	(Jayawardana et al., 2013)	Sri Lanka	POS	155	Factor, Path, Multiple regression
50	(Jensen et al., 2013)	UK	HRM	1,592	Hierarchical linear regression
51	(K. Jiang et al., 2017)	China	HRM	620	Hierarchical linear regression
52	(Joo et al., 2015)	South Korea	POS	291	Factor, Multiple regression
53	(Karanika-Murray et al., 2015)	UK	ENG	158	Mediation, Regression
54	(Kasalak & Bilgin Aksu, 2014)	Turkey	POS	214	Factor, Multiple linear regression
55	(Kaše et al., 2009)	Slovenia	HRM	343	Multiple regression
56	(Katou et al., 2014)	Greece	HRM	1,250	Structural equation modelling
57	(Kehoe & Wright, 2013)	USA	HRM	Undefined	Hierarchical linear regression

#	Author & Date	Country	Relevant Construct	Sample	Main Analysis
58	(Kilroy et al., 2016)	Canada	HRM	545	Structural equation modelling
59	(Kilroy et al., 2017)	Canada	HRM	185	Structural equation modelling
60	(Köse, 2016)	Turkey	ENG, POS	433	Structural equation modelling
61	(Koster, 2011)	Multiple	HRM	18,522	Hierarchical linear regression
62	(Kottke & Sharafinski, 1988)	USA	POS, PSS	216	Factor
63	(Kuvaas & Dysvik, 2009)	Norway	HRM	826	Hierarchical linear regression
64	(Lamm et al., 2015)	US	POS	733	Factor, Multiple regression
65	(Lee & Peccei, 2007)	South Korea	POS	910	Structural equation modelling
66	(Li & Frenkel, 2017)	China	HRM	352	Hierarchical linear regression
67	(Li et al., 2017)	China	HRM	196	Hierarchical linear regression
68	(Liao et al., 2009)	Japan	HRM	2894	Factor, Hierarchical linear regression
69	(Ma et al., 2016)	China	HRM	311	Hierarchical linear regression
70	(Mache et al., 2014)	Germany	ENG	223	Correlation, Regression
71	(Macky & Boxall, 2007)	New Zealand	HRM	424	MANCOVA
72	(Maden & Kabasakal, 2014)	Turkey	POS	213	Structural equation modelling, Regression
73	(Malach-Pines, 2005)	Israel	ENG (Burnout)	869	Factor
74	(Mao & Hale, 2015)	China	COM	154	Independent-samples t-tests, Pearson correlations
75	(McCarthy et al., 2013)	Ireland	POS, PSS	862	Hierarchical linear regression
76	(Mitic et al., 2017)	Serbia	COM	380	Pearson correlation, Hierarchical regression
77	(Monks et al., 2016)	Ireland	HRM	2348	Structural equation modelling
78	(Monks et al., 2016)	UK & Ireland	HRM	498	Structural equation modelling
79	(Mukkavilli et al., 2017)	India	ENG	371	ANOVA
80	(Musah et al., 2017)	Malaysia	COM	98	Factor, Regression
81	(Netemeyer et al., 1997)	USA	PSS	297	Structural equation modelling
82	(Neves & Eisenberger, 2012)	US	COM, POS	236	Factor
83	(Newman et al., 2015)	China	POS, PSS	393	Structural equation modelling
84	(Ngo et al., 2014)	China	POS	591	Factor, Hierarchical regression
85	(Nordin et al., 2014)	Malaysia	COM	125	Pearson correlation
86	(Park et al., 2016)	China	POS	176	Factor, Hierarchical regression

#	Author & Date	Country	Relevant Construct	Sample	Main Analysis
87	(Petrović et al., 2017)	Serbia	ENG	860	Factor, Hierarchical regression
88	(Piening et al., 2013)	UK	HRM	169	General method of moments estimator
89	(Raina & Roebuck, 2016)	India	COM	105	Factor, Multiple regression
90	(Rayton & Yalabik, 2014)	UK	ENG	191	Structural equation modelling
91	(Rich et al., 2010)	USA	ENG	245	Structural equation modelling
92	(Sanders et al., 2008)	Netherlands	HRM	959	Hierarchical linear regression
93	(Sarangi & Srivastava, 2012)	India	COM, ENG	247	Correlation, Regression
94	(Schaufeli et al., 2002)	Spain	ENG	933	Factor
95	(Sharma et al., 2015)	US	COM	2,600	ANOVA, Pearson correlation
96	(Shih et al., 2010)	Taiwan	HRM	174	Structural equation modelling
97	(Smet et al., 2016)	Belgium	COM	1994	Intra-class correlation coefficients, mediation
98	(Solberg & Dysvik, 2016)	Norway	HRM	238	Path
99	(Steele & Plenty, 2015)	Trinidad & Tobago	COM	204	Correlation, Regression
100	(Taghipour & Dezfuli, 2013)	Iran	ENG	226	Structural equation modelling
101	(Takeuchi & Takeuchi, 2013)	Japan	HRM	1052	Structural equation modelling
102	(Takeuchi et al., 2007)	Japan	HRM	849	Hierarchical linear regression
103	(Tremblay et al., 2010)	Canada	HRM, POS	1219	Structural equation modelling
104	(Tsai et al., 2009)	Taiwan	COM	467	Multiple regression
105	(Turunc & Altuntepe, 2015)	Turkey	POS, PSS	480	Factor, Hierarchical regression
106	(Uslu & Çubuk, 2015)	Turkey	COM	227	Factor, Regression
107	(Üstüner & Kis, 2014)	Turkey	COM	217	Correlation, Regression
108	(Veld & Alfes, 2017)	Netherlands	HRM	415	Hierarchical linear regression
109	(Vera et al., 2016)	Portugal	ENG	313	Hierarchical linear regression
110	(Vermeeren, 2014)	Netherlands	HRM	356	Hierarchical linear regression
111	(Viljevac et al., 2012)	Canada	ENG	139	Structural equation modelling
112	(Wagenheim & Rood, 2010)	US	COM	97	Simple linear regression
113	(Walden et al., 2017)	Singapore	COM, ENG	539	Structural equation modelling
114	(Yakın & Erdil, 2012)	Turkey	ENG	161	Correlation, Regression
115	(Yamaguchi, 2009)	Japan & US	COM	1008	Structural equation modelling
116	(Yamaguchi, 2017)	Japan	COM	742	Factor

#	Author & Date	Country	Relevant Construct	Sample	Main Analysis
117	(Yamamoto, 2013)	Japan	HRM	400	Multiple regression
118	(Yildiz, 2013)	Turkey	COM	362	Path, Relational scanning model
119	(Yu & Frenkel, 2013)	AUS	POS	206	Structural equation modelling
120	(Zhong et al., 2016)	China	ENG, HRM, POS	605	Factor, Hierarchical linear regression

Note. HR – perception of Human Resource (management), COM – Perception of Organizational Communication, POS – Perception of Organizational Support, PSS – Perception of Supervisory Support, ENG – perception of Work Engagement

Appendix B: Survey One

Welcome and thank you for partaking in our research

We compiled a survey that aims to help staff like you assess the performance management system that their organization uses to monitor and improve their everyday performance.

We need your help to test it.
Please answer the following questions about your experiences with performance management in your organization.
Remember, there are no right or wrong answers

The survey should take around 20 – 30 minutes maximum to complete.

Your voice will make this research relevant to workers like you

Please read the PLS and Consent form the before proceeding.

Please indicate your consent to complete this survey

◯ I consent
◯ I do not consent

Were you assigned an ID number from Prolific?

If Yes - Please provide your Prolific ID number

If No - To thank you for your time, we'd like to send you a book voucher upon completion of this survey. Please enter a secure email address below to receive the voucher.

Please indicate the extent to which you agree with the following statements.	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
Job descriptions are clearly written & understood	◯	◯	◯	◯	◯	◯
Job descriptions are aligned with organisational goals & good practice.	◯	◯	◯	◯	◯	◯
Performance evaluation policies are written & available to all staff & relevant stakeholders.	◯	◯	◯	◯	◯	◯
There are safeguards of the process to ensure consistency & the manager's fairness of judgement	◯	◯	◯	◯	◯	◯
The information gathered during an evaluation is protected & held confidentially.	◯	◯	◯	◯	◯	◯
Only those with a legitimate purpose have access to personnel evaluations	◯	◯	◯	◯	◯	◯
Managers conduct all interactions (written & verbal) in a professional, constructive manner	◯	◯	◯	◯	◯	◯
There is a process in place to address incidences of unprofessional interactions in performance evaluations	◯	◯	◯	◯	◯	◯
Both strengths & weaknesses are examined	◯	◯	◯	◯	◯	◯
Performance ratings are conducive to differentiation among levels of performance	◯	◯	◯	◯	◯	◯
There are safeguards to ensure that pre-existing conditions or events will not compromise the manager's ability to be fair & unbiased.	◯	◯	◯	◯	◯	◯
The evaluation process meets all laws & guidelines, including those established through collective bargaining.	◯	◯	◯	◯	◯	◯
Those involved generally agree that the evaluations are fair and efficient.	◯	◯	◯	◯	◯	◯
The evaluation process reflects the organisation's goals & mission	◯	◯	◯	◯	◯	◯
Feedback & professional development based on evaluation are aligned with the organisation's goals & missions	◯	◯	◯	◯	◯	◯
All users (staff, managers, etc) of the evaluation process have been clearly identified from the beginning of the evaluation cycle.	◯	◯	◯	◯	◯	◯
The uses for the information (dismissal, tenure, merit pay, etc) have been clearly identified	◯	◯	◯	◯	◯	◯
All the performance managers have received appropriate training in the evaluation process	◯	◯	◯	◯	◯	◯

Please indicate the extent to which you agree with the following statements.	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
Those who manage the records have received appropriate training, & hold appropriate credentials	○	○	○	○	○	○
Performance criteria reflect only the job expectations of those evaluated	○	○	○	○	○	○
The criteria for one group are not used for another group with unrelated job expectations (e.g. an evaluation form for teachers used for guidance counsellors)	○	○	○	○	○	○
All evaluation reports generated by the managers meet deadlines & provide useful, accurate information.	○	○	○	○	○	○
Data generated by evaluations are used in developing professional development plans.	○	○	○	○	○	○
There is appropriate follow-up of evaluation results	○	○	○	○	○	○
Procedures for collecting data are as simple & job-embedded as possible to prevent undue staff overburdening	○	○	○	○	○	○
All stakeholders have the opportunity to question the procedures or results of an evaluation	○	○	○	○	○	○
There is a process in place to determine the outcome of questions asked concerning an evaluation	○	○	○	○	○	○
The organisation can afford the resources to conduct performance evaluation in the way that will maximise its effect	○	○	○	○	○	○
There is an adequate number of data sources to provide a comprehensive view of performance.	○	○	○	○	○	○
There is a system in place ensuring that the managers follow procedures, analyse all appropriate data, & report judgements based only on the criteria set forth in the system	○	○	○	○	○	○
The expectations & scope of work for staff are clearly defined & understood	○	○	○	○	○	○
The expectations have been provided to all staff in both written and verbal formats	○	○	○	○	○	○
The expectations are reasonable and directly related to job descriptions	○	○	○	○	○	○
Whenever data are collected, there is an expectation that the details of the circumstances are also be recorded (e.g. Notation on observation forms)	○	○	○	○	○	○

Please indicate the extent to which you agree with the following statements.	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
All stakeholders clearly understand the purposes & procedures to be followed	◯	◯	◯	◯	◯	◯
There are safeguards in place to ensure that the results of any given evaluation would be the same regardless of manager	◯	◯	◯	◯	◯	◯
There are safeguards in place to ensure that the evaluation procedures are the same regardless of manager	◯	◯	◯	◯	◯	◯
There is a structure in place ensuring that all evaluation information is held in a secure location	◯	◯	◯	◯	◯	◯
There is a system in place to record person, time, date, & purpose of access to records	◯	◯	◯	◯	◯	◯
There are safeguards to ensure that the results of an evaluation are not influenced by preconceived ideas of the manager that may be unrelated to the actual job performance	◯	◯	◯	◯	◯	◯
Manager training includes bias control & diversity awareness training	◯	◯	◯	◯	◯	◯
There is a grievance process in place to offer protection to staff	◯	◯	◯	◯	◯	◯
There are safeguards of the manager's final reports & disposition to ensure accuracy & use of data	◯	◯	◯	◯	◯	◯
Conclusions drawn are consistent with the data gathered	◯	◯	◯	◯	◯	◯
Managers are required to justify the disposition of an evaluation based on documentation of performance	◯	◯	◯	◯	◯	◯
A system is in place to allow the periodic review of the performance evaluation system to ensure its continued usefulness	◯	◯	◯	◯	◯	◯
Staff are not actively encouraged to question policies & processes that affect their work ®	◯	◯	◯	◯	◯	◯
Using evidence (e.g. survey results or articles) as a basis for problem solving & decision-making is encouraged	◯	◯	◯	◯	◯	◯
Allocation of HR resources is based on relevant needs assessment.	◯	◯	◯	◯	◯	◯
Our performance management system is not regularly evaluated to ensure its continued usefulness & positive impact on staff performance ®	◯	◯	◯	◯	◯	◯

Please indicate the extent to which you agree with the following statements.	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
At work, I'm encouraged to use …						
• academic research to inform important work decisions	◯	◯	◯	◯	◯	◯
• business magazines to inform important work decisions	◯	◯	◯	◯	◯	◯
• past experiences to inform important work decisions	◯	◯	◯	◯	◯	◯
• manager / colleague advice to inform important work decisions	◯	◯	◯	◯	◯	◯
• professional development to inform important work decisions	◯	◯	◯	◯	◯	◯
• staff survey results to inform important work decisions	◯	◯	◯	◯	◯	◯
• staff and / or organisation performance results to inform important work decisions	◯	◯	◯	◯	◯	◯
In general, there is a high level of performance among …						
• new staff	◯	◯	◯	◯	◯	◯
• recent graduates	◯	◯	◯	◯	◯	◯
• tenured staff	◯	◯	◯	◯	◯	◯
• casual staff	◯	◯	◯	◯	◯	◯
• staff on a fixed-term contract	◯	◯	◯	◯	◯	◯
• permanent staff	◯	◯	◯	◯	◯	◯
In general, unplanned lateness is frequent among ® …						
• new staff	◯	◯	◯	◯	◯	◯
• recent graduates	◯	◯	◯	◯	◯	◯
• tenured staff	◯	◯	◯	◯	◯	◯
• among casual staff	◯	◯	◯	◯	◯	◯
• staff on a fixed-term contract	◯	◯	◯	◯	◯	◯
• permanent staff	◯	◯	◯	◯	◯	◯
• low performers	◯	◯	◯	◯	◯	◯
• high performers	◯	◯	◯	◯	◯	◯
In general, unplanned absence is frequent among ® …						
• new staff	◯	◯	◯	◯	◯	◯

Please indicate the extent to which you agree with the following statements.	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
• recent graduates	◯	◯	◯	◯	◯	◯
• tenured staff	◯	◯	◯	◯	◯	◯
• casual staff	◯	◯	◯	◯	◯	◯
• staff on a fixed-term contract	◯	◯	◯	◯	◯	◯
• permanent staff	◯	◯	◯	◯	◯	◯
• low performers	◯	◯	◯	◯	◯	◯
• high performers	◯	◯	◯	◯	◯	◯
In general, turnover is high among ® …						
• new staff	◯	◯	◯	◯	◯	◯
• recent graduates	◯	◯	◯	◯	◯	◯
• tenured staff	◯	◯	◯	◯	◯	◯
• casual staff	◯	◯	◯	◯	◯	◯
• staff on a fixed-term contract	◯	◯	◯	◯	◯	◯
• permanent staff	◯	◯	◯	◯	◯	◯
• low performers	◯	◯	◯	◯	◯	◯
• high performers	◯	◯	◯	◯	◯	◯
At my workplace, the communication climate is …						
• Motivating	◯	◯	◯	◯	◯	◯
• Positive	◯	◯	◯	◯	◯	◯
• multi-directional (flows between & within teams & leadership)	◯	◯	◯	◯	◯	◯
• dominated by gossip ®	◯	◯	◯	◯	◯	◯
• dominated by complaints ®	◯	◯	◯	◯	◯	◯
Creative solutions & new ideas are not welcomed at work ®	◯	◯	◯	◯	◯	◯
Communication with my supervisor is …						
• Difficult ®	◯	◯	◯	◯	◯	◯
• welcoming of questions	◯	◯	◯	◯	◯	◯
• informative	◯	◯	◯	◯	◯	◯

Please indicate the extent to which you agree with the following statements.	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
• irrelevant ®	○	○	○	○	○	○
The organisation provides me with …						
• easily accessible information relevant to my work (e.g. my role, pay, new staff etc)	○	○	○	○	○	○
• encouragement to provide feedback about information relevant to my work (e.g. my role, pay, new staff etc)	○	○	○	○	○	○
Feedback about my performance is ...						
• formally done annually & informally given throughout the year	○	○	○	○	○	○
• formally done once a year	○	○	○	○	○	○
• constructive	○	○	○	○	○	○
• focused on strengths and weaknesses	○	○	○	○	○	○
• rarely followed-up / actioned ®	○	○	○	○	○	○
• respectful	○	○	○	○	○	○
• confidential to me and relevant senior management	○	○	○	○	○	○
• disengaging ®	○	○	○	○	○	○
My job description accurately & clearly describes my tasks, responsibilities, timeframes & expectations.	○	○	○	○	○	○
Some of my responsibilities overlap so much with other teammates leading to lack of role clarity & even conflict at times ®	○	○	○	○	○	○
When it comes to opportunities at work…						
• favouritism plays a big role ®	○	○	○	○	○	○
• organisational politics play a big role ®	○	○	○	○	○	○
• learning needs are the basis of professional development	○	○	○	○	○	○
• performance is the basis of reward or disciplinary action	○	○	○	○	○	○
My supervisor shows me support through …						
• fair conflict resolution	○	○	○	○	○	○
• mentoring me	○	○	○	○	○	○
• giving me appropriate leave when needed	○	○	○	○	○	○

Please indicate the extent to which you agree with the following statements.	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
• advocating for my work needs with senior management	◯	◯	◯	◯	◯	◯
My organisation shows me support through …						
• fair conflict resolution policies	◯	◯	◯	◯	◯	◯
• job security	◯	◯	◯	◯	◯	◯
• fair job rewards	◯	◯	◯	◯	◯	◯
• giving me appropriate leave when needed	◯	◯	◯	◯	◯	◯
• responding appropriately to any feedback or questions I have	◯	◯	◯	◯	◯	◯
I don't enjoy attending social gatherings at work (e.g. annual party) ®	◯	◯	◯	◯	◯	◯
I enjoy attending formal gatherings at work (e.g. seminars)	◯	◯	◯	◯	◯	◯
I enjoy contributing to gatherings (e.g. help with setting up)	◯	◯	◯	◯	◯	◯
I find it relatively easy to be motivated to come to work	◯	◯	◯	◯	◯	◯
Generally, I feel productive at work	◯	◯	◯	◯	◯	◯
Generally, my work is challenging but not overwhelming	◯	◯	◯	◯	◯	◯
At the end of each work day, I can't wait to leave ®	◯	◯	◯	◯	◯	◯
I dislike getting new tasks ®	◯	◯	◯	◯	◯	◯

Note. ® indicates reverse coding applied at analysis.

Demographic Items

1. Which of the following best describes your job role?
    - HR staff
    - Team leader
2. What is your work location?
3. What is the nature of your current work contract?
    - Part-time
    - Full-time
    - Fixed-term
    - Permanent
    - Casual
4. How many people work in your organization?
    - 1 to 50
    - 51 to 200
    - 201 to 2000
    - 2001 to 4000
    - over 4000
5. To which industry does your organization belong? (e.g. Education, Healthcare, Aviation)

Thank you for your invaluable input

* If the participant were from Prolific Academic, they were provided with a URL to return to the Prolific Academic platform and a survey completion code.

Appendix C: Survey Two

Welcome and thank you for partaking in our research

We compiled a survey that aims to help staff like you assess the performance management system that their organization uses to monitor and improve their everyday performance.

We need your help to test it.
Please answer the following questions about your experiences with performance management in your organization.
Remember, there are no right or wrong answers

The survey should take around 10 - 15 minutes maximum to complete.

Your voice will make this research relevant to workers like you

Please read the PLS and Consent form the before proceeding.

Please indicate your consent to complete this survey

◯ I consent
◯ I do not consent

Were you assigned an ID number from Prolific?

If Yes - Please provide your Prolific ID number
If No - To thank you for your time, we'd like to send you a book voucher upon completion of this survey. Please enter a secure email address below to receive the voucher.

Demographic Items

1. Age
2. Gender
3. Which of the following best describes your job role in performance management?
    - Administrative Staff
    - Junior Management
    - Middle Management
    - Senior Management
4. What is your work location?
5. What is the nature of your current work contract? Please tick all that apply
    - Part-time
    - Full-time
    - Fixed-term
    - Permanent
    - Casual
6. How long have you been with your current employer?
7. How long have you been in the workforce (in general)?
8. How many people work in your organization?
    - 1-50
    - 51-1000
    - 1001 - 4000
    - Over 4000
9. To which industry does your organization belong? (e.g. Education, Healthcare, Aviation)

Please indicate the extent to which you agree with the following statements.	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
A system is in place to allow the periodic review of the performance evaluation system to ensure its continued usefulness	○	○	○	○	○	○
Data generated by evaluations are used in developing professional development plans.	○	○	○	○	○	○
Conclusions drawn are consistent with the data gathered	○	○	○	○	○	○
There is a system in place ensuring that the managers follow procedures, analyse all appropriate data, & report judgements based only on the criteria set forth in the system	○	○	○	○	○	○
There is an adequate number of data sources to provide a comprehensive view of performance.	○	○	○	○	○	○
There is appropriate follow-up of evaluation results	○	○	○	○	○	○
All the performance managers have received appropriate training in the evaluation process	○	○	○	○	○	○
The expectations & scope of work for staff are clearly defined & understood	○	○	○	○	○	○
Feedback & professional development based on evaluation are aligned with the organisation's goals & missions	○	○	○	○	○	○
Procedures for collecting data are as simple & job-embedded as possible to prevent undue staff overburdening	○	○	○	○	○	○
All stakeholders have the opportunity to question the procedures or results of an evaluation	○	○	○	○	○	○
There are safeguards to ensure ...						
• the results of an evaluation are not influenced by preconceived ideas of the manager that may be unrelated to the actual job performance	○	○	○	○	○	○
• the results of any given evaluation would be the same regardless of manager	○	○	○	○	○	○
• the evaluation procedures are the same regardless of manager	○	○	○	○	○	○

Please indicate the extent to which you agree with the following statements.	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
When it comes to opportunities at work favouritism plays a big role ®	○	○	○	○	○	○
Feedback about my performance is rarely followed-up / actioned ®	○	○	○	○	○	○
At my workplace, the communication climate is dominated by complaints ®	○	○	○	○	○	○
When it comes to opportunities at work organisational politics play a big role ®	○	○	○	○	○	○
feedback about my performance is disengaging ®	○	○	○	○	○	○
Creative solutions & new ideas are not welcomed at work ®	○	○	○	○	○	○
My supervisor shows me support through ...						
• giving me appropriate leave when needed	○	○	○	○	○	○
• fair conflict resolution	○	○	○	○	○	○
• mentoring me	○	○	○	○	○	○
• advocating for my work needs with senior management	○	○	○	○	○	○
My organisation shows me support through...						
• giving me appropriate leave when needed	○	○	○	○	○	○
• job security	○	○	○	○	○	○
• fair conflict resolution policies	○	○	○	○	○	○
• responding appropriately to any feedback or questions I have	○	○	○	○	○	○
• fair job rewards	○	○	○	○	○	○
Generally, my work is challenging but not overwhelming	○	○	○	○	○	○
Feedback about my performance is respectful	○	○	○	○	○	○
Communication with my supervisor is difficult ®	○	○	○	○	○	○
Job descriptions are clearly written & understood	○	○	○	○	○	○

Please indicate the extent to which you agree with the following statements.	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
Job descriptions are aligned with organisational goals & good practice.	◯	◯	◯	◯	◯	◯
My job description accurately & clearly describes my tasks, responsibilities, time-frames & expectations	◯	◯	◯	◯	◯	◯
Performance evaluation policies are written & available to all staff & relevant stakeholders.	◯	◯	◯	◯	◯	◯
In general, unplanned lateness is frequent among ®...						
• permanent staff	◯	◯	◯	◯	◯	◯
• tenured staff	◯	◯	◯	◯	◯	◯
• high performers	◯	◯	◯	◯	◯	◯
• new staff	◯	◯	◯	◯	◯	◯
• recent graduates	◯	◯	◯	◯	◯	◯
In general, unplanned absence is frequent among ®...						
• permanent staff	◯	◯	◯	◯	◯	◯
• tenured staff	◯	◯	◯	◯	◯	◯
• high performers	◯	◯	◯	◯	◯	◯
• new staff	◯	◯	◯	◯	◯	◯
• recent graduates	◯	◯	◯	◯	◯	◯
• staff on a fixed term contract	◯	◯	◯	◯	◯	◯
In general, turnover is high among ®...						
• new staff	◯	◯	◯	◯	◯	◯
• recent graduates	◯	◯	◯	◯	◯	◯
• tenured staff	◯	◯	◯	◯	◯	◯

Please indicate the extent to which you agree with the following statements.	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
• staff on a fixed-term contract	◯	◯	◯	◯	◯	◯
• casual staff	◯	◯	◯	◯	◯	◯
• low performers	◯	◯	◯	◯	◯	◯
• permanent staff	◯	◯	◯	◯	◯	◯

Note. Reverse coding is denoted with ®

Thank you for your invaluable input

* If the participant were from Prolific Academic, they were provided with a URL to return to the Prolific Academic platform and a survey completion code.

Appendix D: The EMI Scale

Please indicate the extent to which you agree with the following statements.	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
There is appropriate follow-up of evaluation results	◯	◯	◯	◯	◯	◯
There are safeguards of the manager's final reports & disposition to ensure accuracy & use of data	◯	◯	◯	◯	◯	◯
Conclusions drawn are consistent with the data gathered	◯	◯	◯	◯	◯	◯
There is an adequate number of data sources to provide a comprehensive view of performance.	◯	◯	◯	◯	◯	◯
All the performance managers have received appropriate training in the evaluation process	◯	◯	◯	◯	◯	◯
Feedback & professional development based on evaluation are aligned with the organisation's goals & missions	◯	◯	◯	◯	◯	◯
Data generated by evaluations are used in developing professional development plans.	◯	◯	◯	◯	◯	◯
There are safeguards in place to ensure that the results of any given evaluation would be the same regardless of manager	◯	◯	◯	◯	◯	◯
Procedures for collecting data are as simple & job-embedded as possible to prevent undue staff overburdening	◯	◯	◯	◯	◯	◯
All stakeholders have the opportunity to question the procedures or results of an evaluation	◯	◯	◯	◯	◯	◯
A system is in place to allow the periodic review of the performance evaluation system to ensure its continued usefulness	◯	◯	◯	◯	◯	◯
Job descriptions are aligned with organisational goals & good practice.	◯	◯	◯	◯	◯	◯
Job descriptions are clearly written & understood	◯	◯	◯	◯	◯	◯

Please indicate the extent to which you agree with the following statements.	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
My job description accurately & clearly describes my tasks, responsibilities, timeframes & expectations	◯	◯	◯	◯	◯	◯
Performance evaluation policies are written & available to all staff & relevant stakeholders.	◯	◯	◯	◯	◯	◯
My supervisor shows me support through advocating for my work needs with senior management	◯	◯	◯	◯	◯	◯
My supervisor shows me support through fair conflict resolution	◯	◯	◯	◯	◯	◯
My supervisor shows me support through mentoring me	◯	◯	◯	◯	◯	◯
Communication with my supervisor is difficult ®	◯	◯	◯	◯	◯	◯
My organisation shows me support through fair job rewards	◯	◯	◯	◯	◯	◯
My organisation shows me support through giving me appropriate leave when needed	◯	◯	◯	◯	◯	◯
My organisation shows me support through job security	◯	◯	◯	◯	◯	◯
Generally, my work is challenging but not overwhelming	◯	◯	◯	◯	◯	◯
Feedback about my performance is disengaging ®	◯	◯	◯	◯	◯	◯
At my workplace, the communication climate is dominated by complaints ®	◯	◯	◯	◯	◯	◯
Creative solutions & new ideas are not welcomed at work ®	◯	◯	◯	◯	◯	◯
When it comes to opportunities at work favouritism plays a big role ®	◯	◯	◯	◯	◯	◯
When it comes to opportunities at work organisational politics play a big role ®	◯	◯	◯	◯	◯	◯
In general, turnover is high among recent graduates ®	◯	◯	◯	◯	◯	◯
In general, turnover is high among tenured staff ®	◯	◯	◯	◯	◯	◯
In general, turnover is high among new staff ®	◯	◯	◯	◯	◯	◯
In general, turnover is high among permanent staff ®	◯	◯	◯	◯	◯	◯

Please indicate the extent to which you agree with the following statements.	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
In general, turnover is high among casual staff ®	◯	◯	◯	◯	◯	◯
In general, unplanned absence is frequent among recent graduates ®	◯	◯	◯	◯	◯	◯
In general, unplanned absence is frequent among new staff ®	◯	◯	◯	◯	◯	◯
In general, unplanned lateness is frequent among recent graduates ®	◯	◯	◯	◯	◯	◯
In general, unplanned lateness is frequent among new staff ®	◯	◯	◯	◯	◯	◯
In general, unplanned absence is frequent among tenured staff ®	◯	◯	◯	◯	◯	◯
In general, unplanned absence is frequent among permanent staff ®	◯	◯	◯	◯	◯	◯
In general, unplanned lateness is frequent among tenured staff ®	◯	◯	◯	◯	◯	◯
In general, unplanned lateness is frequent among permanent staff ®	◯	◯	◯	◯	◯	◯

Note. Reverse coding is denoted with ®

Appendix E: EMI Scale – Item Orientation

Orientation	Item
Descriptive Observational	A system is in place to allow the periodic review of the performance evaluation system to ensure its continued usefulness
	Data generated by evaluations are used in developing professional development plans.
	There are safeguards in place to ensure that the results of any given evaluation would be the same regardless of manager
	There are safeguards of the manager's final reports & disposition to ensure accuracy & use of data
Evaluative Experiential	Communication with my supervisor is difficult ®
	Feedback about my performance is disengaging ®
	Generally, my work is challenging but not overwhelming
	My job description accurately & clearly describes my tasks, responsibilities, timeframes & expectations
	My organization shows me support through fair job rewards
	My organization shows me support through giving me appropriate leave when needed
	My organization shows me support through job security
	My supervisor shows me support through advocating for my work needs with senior management
	My supervisor shows me support through fair conflict resolution
	My supervisor shows me support through mentoring me
Evaluative Observational	All stakeholders have the opportunity to question the procedures or results of an evaluation
	All the performance managers have received appropriate training in the evaluation process
	At my workplace, the communication climate is dominated by complaints ®
	Conclusions drawn are consistent with the data gathered
	Creative solutions & new ideas are not welcomed at work ®
	Feedback & professional development based on evaluation are aligned with the organisation's goals & missions
	In general, turnover is high among casual staff ®
	In general, turnover is high among new staff ®
	In general, turnover is high among permanent staff ®

Orientation	Item
	In general, turnover is high among recent graduates ®
	In general, turnover is high among tenured staff ®
	In general, unplanned absence is frequent among new staff ®
	In general, unplanned absence is frequent among permanent staff ®
	In general, unplanned absence is frequent among recent graduates ®
	In general, unplanned absence is frequent among tenured staff ®
	In general, unplanned lateness is frequent among new staff ®
	In general, unplanned lateness is frequent among permanent staff ®
	In general, unplanned lateness is frequent among recent graduates ®
	In general, unplanned lateness is frequent among tenured staff ®
	Job descriptions are aligned with organizational goals & good practice.
	Job descriptions are clearly written & understood
	Performance evaluation policies are written & available to all staff & relevant stakeholders.
	Procedures for collecting data are as simple & job-embedded as possible to prevent undue staff overburdening
	There is an adequate number of data sources to provide a comprehensive view of performance.
	There is appropriate follow-up of evaluation results
	When it comes to opportunities at work favoritism plays a big role ®
	When it comes to opportunities at work organizational politics play a big role ®

Note. Reverse coding is denoted with ®

Appendix F: EMI Scale – Item Subject

Subject Matter	Item
Content & Motivation	A system is in place to allow the periodic review of the performance evaluation system to ensure its continued usefulness
	Feedback & professional development based on evaluation are aligned with the organisation's goals & missions
	There are safeguards in place to ensure that the results of any given evaluation would be the same regardless of manager
	There are safeguards of the manager's final reports & disposition to ensure accuracy & use of data
	There is an adequate number of data sources to provide a comprehensive view of performance.
Content & Process	Job descriptions are aligned with organizational goals & good practice.
	Job descriptions are clearly written & understood
	My job description accurately & clearly describes my tasks, responsibilities, timeframes & expectations
	Performance evaluation policies are written & available to all staff & relevant stakeholders.
Outcome	Generally, my work is challenging but not overwhelming
	In general, turnover is high among casual staff ®
	In general, turnover is high among new staff ®
	In general, turnover is high among permanent staff ®
	In general, turnover is high among recent graduates ®
	In general, turnover is high among tenured staff ®
	In general, unplanned absence is frequent among new staff ®
	In general, unplanned absence is frequent among permanent staff ®
	In general, unplanned absence is frequent among recent graduates ®
	In general, unplanned absence is frequent among tenured staff ®
	In general, unplanned lateness is frequent among new staff ®
	In general, unplanned lateness is frequent among permanent staff ®
	In general, unplanned lateness is frequent among recent graduates ®

Subject Matter	Item
	In general, unplanned lateness is frequent among tenured staff ®
Process	All the performance managers have received appropriate training in the evaluation process
	At my workplace, the communication climate is dominated by complaints ®
	Communication with my supervisor is difficult ®
	Conclusions drawn are consistent with the data gathered
	Creative solutions & new ideas are not welcomed at work ®
	Data generated by evaluations are used in developing professional development plans.
	Feedback about my performance is disengaging ®
	My organization shows me support through fair job rewards
	My organization shows me support through giving me appropriate leave when needed
	My organization shows me support through job security
	My supervisor shows me support through advocating for my work needs with senior management
	My supervisor shows me support through fair conflict resolution
	My supervisor shows me support through mentoring me
	There is appropriate follow-up of evaluation results
	When it comes to opportunities at work favoritism plays a big role ®
	When it comes to opportunities at work organizational politics play a big role ®
Process & Motivation	All stakeholders have the opportunity to question the procedures or results of an evaluation
	Procedures for collecting data are as simple & job-embedded as possible to prevent undue staff overburdening

Note. Reverse coding is denoted with ®