The effect of coaching frequency and tenure on engagement and turnover in the New Zealand Police Service: a conditional process analysis

A thesis presented in partial fulfilment of the requirements for the degree of

Master of Business Studies

At Massey University, Albany

New Zealand

Jesse William Mowat

2022

ABSTRACT

To realise the vision of making New Zealand the safest country in the world, in 2017 the New Zealand Police Service (NZ Police) introduced the Police High Performance Framework (PHPF) to equip and enable the organisation's 14,500 strong workforce to perform to its best. As a part of the framework, the organisation encouraged its supervisors to deliver monthly one-on-one job-related coaching to each of their direct reports. Notwithstanding the allure of the 'supervisor-as-coach' model, there is debate about how and when it delivers the favourable outcomes that organisations anticipate. Via the post-positivist research tradition and through the lens of leader-member exchange theory, this study helps fill this research-gap by exploring the direct and mediated effect of coaching frequency on employee engagement, and coaching frequency on intention to turnover, and exploring whether employee tenure (length-of-service) has a moderating influence. Using conditional process analysis, three linear regression models were fitted to survey-data collected from 150 NZ Police employees. It was found that coaching frequency has a tiny but statistically significant effect on employee engagement and intention to turnover both directly, and when mediated by the perceived quality of the employee coaching relationship (PQECR). Additionally, the study found that the PQECR has a small, positive, direct effect on employee engagement, and a moderate, negative, direct effect intention to turnover. Surprisingly, it was found that tenure has no moderating effect whatsoever. Therefore, to fully benefit from the supervisor-as-coach model, NZ Police should focus on two main things. The first is to equip the organisation's supervisors by providing them with ongoing training on how to strengthen their direct reports' PQECR. The second is to make structural changes that enable the organisation's supervisors to coach each of their direct reports (regardless of tenure) on at least a monthly basis as prescribed.

ACKNOWLEDGEMENTS

Accepting the challenge of a project of this magnitude is not without trepidation. Indeed, without the unwavering faith, encouragement, and ongoing support from Dr Darryl Forsyth (Massey University Business School, Albany), it is a certainty that this journey would not have progressed beyond its first tentative steps. I would also like to acknowledge the support of Associate Professor David Brougham (Massey University Palmerston North) who, despite the distance separating us, was always ready and willing to provide sage advice, read drafts, and provide encouraging feedback.

Collecting data for this project proved to be the most significant challenge, and I could not have achieved this without the support of Ms Vanessa Merrett, the NZ Police Research Liaison Officer assigned to this project. Vanessa always had time for me, made introductions, and shared her networks. Simply put, she was the key that opened doors.

Last, but definitely not least, I would like to thank my wife, Bonita, who encouraged me to 'go for it' from the beginning, and never once complained about the hours I spent toiling away in our home office during evenings, weekends, and holidays. It has not been easy being the spouse of this part-time student researcher; we are both fortunate that Bonita has the patience of a Saint.

Finally, I would like to thank all of NZ Police employees who took the time and effort to respond to my survey. Without you (and your data), there would be no study, and no opportunity to seriously think about how workplace coaching might be done better in the NZ Police.

DEDICATION

This study is dedicated to two people. To my wife, Bonita. No one could have been more supportive or tolerant than you. I know it has not been easy and I can hardly believe we are at the finish line. Thank you for all the times you cooked meals, did chores, and organised our kids, making space for me to get it all done. Credit for this piece of work is as much yours as it is mine.

To Dr Julie Lynette Carr, my mother, who passed away during this project. Mum, you were so proud of me for extending my education, and I know you would have been in the front-row at graduation. I have thought about you lots during this journey and miss you heaps.

DISCLAIMER

The results presented in this thesis are the work of the researcher, Jesse William Mowat, supervised by Dr Darryl Forsyth (Massey University) and Associate Professor David Brougham (Massey University). Jesse William Mowat takes full responsibility for the results.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

The topic of this research is the phenomenon of employee coaching, a leadership practice that takes place in the context of the workplace between a supervisor and each of their direct reports. Notwithstanding a minority of detractors such as Deming (2000), most organisational behaviour commentators support the idea that performance management at the individual level is a logical means for achieving organisational goals (DeNisi & Murphy, 2017; Du Plessis & Van Niekerk, 2017; Stone, 2017). Therefore, organisations passing-up such an opportunity risk fading away. Indeed, the NZ Police mobilised in 2017 when Commissioner of the day, Mike Bush, launched the Police High Performance Framework (PHPF) (New Zealand Police, 2017). PHPF is a performance management system designed to equip and enable NZ Police's circa 14,500 employees to deliver the outcomes desired by the organisation (New Zealand Police, 2017).

The investment in PHPF makes sense because according to NZ Police's strategic document, Our Business, the organisation's vision is to make New Zealand the safest country in the world (New Zealand Police, 2018). However, since the inception of PHPF, Iceland has occupied the top spot, consistently ranked as the most peaceful country in the world (Institute for economics & peace, 2021), so is the PHPF a flawed system, or is something else going on?

1.1 Research background

PHPF prescribes that progress reviews (coaching conversations), should take place between every police supervisor (coach) and each direct report (coachee) at least once per month (New Zealand Police, 2017). The purpose of these one-on-one coaching conversations is threefold: firstly, to ensure the coachee understands that by performing specific behaviours, they will help the organisation realise its vision. Secondly, the coach provides feedback to the coachee on their individual performance, celebrating what the coachee has done well and exploring areas for improvement. Finally, the coach and the coachee work together to produce an individual development plan that focuses on closing performance gaps so the coachee can maximise their potential (New Zealand Police, 2017).

The emphasis placed on coaching conversations confirms they are considered a crucial component of PHPF. Indeed, they are the substratum from which high-performance takes root. Theoretically then, if coaching conversations take place infrequently or not at all, the stimulus for high-performance is dilute. Imagine flying a fighter-jet with a missing engine?

Yes, it will remain airborne, but the aircraft's performance is likely to be underwhelming. Moreover, when a supervisor-coach fails to engage in frequent coaching conversations with a coachee, a crucial component is withheld, meaning the coachee is under-nurtured and less likely to perform as well as they could.

1.2 Statement of the problem

Given the importance placed of coaching conversations within the PHPF, it is reasonable to expect the NZ Police would go to considerable lengths to develop a workplace culture that encourages employee coaching, including training the organisation's supervisors and adjusting organisational settings to ensure employee coaching is delivered as prescribed, but this does not appear to be the case. During a small-scale-study conducted by the researcher in 2020 (Mowat, 2020), three years after the introduction of PHPF, data was collected from a convenience sample of 22 NZ Police Officers (6 Sergeants and 16 Constables) to find out if coaching conversations were taking place as frequently as prescribed. The main results of the small-scale study were as follows:

- 100% of coaches (Sergeants) agreed that they engage in coaching conversations with their coachees (Constables); however, only 83% agreed that this practice occurs monthly. The remainder agreed that coaching conversations take place at a frequency of once every two-three months
- 31% of coachees agreed that they had not experienced a coaching conversation with their coach in the past six months but would welcome this opportunity
- Coachees that experienced at least one coaching conversation in the previous six months reported that the average frequency of coaching conversations with their coach was once every three months

These results suggest that coaching conversations are not happening as frequently as they should, and this could pose a serious problem for NZ Police because the organisation's vision is ambitious, and it may struggle to realise the vision when a vital component is missing from the machinery. Indeed, the literature on performance feedback suggests that frequent and specific feedback has a greater positive effect on work performance than infrequent and global feedback (Park, Johnson, Moon, & Lee, 2019).

Having presented an overview of the problem, it is worthwhile considering why coaching conversations may not be happening as frequently as prescribed. Responses from those who

participated in the previously mentioned small-scale-study (Mowat, 2020) suggest there are three main barriers:

- Lack of motivation: shift-worker fatigue has a negative effect on the coach's mindset, reducing the motivation to engage in coaching conversations with coachees
- Lack of opportunity: coachees are frequently abstracted for other duties (Court, training, and operations) or on annual leave, reducing the opportunity for coaching conversations to take place
- Lack of priority: operational delivery takes priority over coaching conversations

This combination of psychological and structural barriers is unsurprising considering the nature of policing. According to the Policing Act (2008), the NZ Police are responsible for law enforcement, national security, and emergency management (among other functions) which are serious responsibilities, and one can appreciate why coaching conversations are set aside in favour of operational delivery. Interestingly, a similar trend was observed by Chen, Khan and Lin (2019) in their study of a sample of 530 hospital administrators in Taiwan. The authors found that coachees did not universally value a people-before-task approach. The explanation for this attitude was that in a hospital setting, where emergency situations are frequent, task completion must necessarily come before people (Chen et al., 2019).

In addition to the nature of police-work, pressure to conform to a prototypical police officer stereotype may impact on coaching effectiveness. Whilst police supervisors and their direct reports may enjoy positive social relations, they are also expected to be tough, resilient, and prepared to face danger, so it is not unreasonable to suggest the vocational culture may discourage the admission of personal weakness. As a result, coachees may resist their coaches' attempts to identify and correct the drivers of sub-optimal performance (Misiukonis, 2011).

Along with coachee-resistance, it is also likely that a proportion of police supervisors do not perceive themselves as coaches or consider coaching their direct reports to be a particularly valuable use of their time. For example, a Sergeant-coach who participated in the aforementioned research project (Mowat, 2020) commented that PHPF is a box-ticking exercise, of little value, done to please police managers. Anti-coaching mindsets like this are acknowledged in the literature. Drawing on the earlier work of Heslin, Vandewalle and

Latham (2006), Lawrence (2017) posits that some coaches do not believe that others can change (entity theory) and because of this, they perceive coaching as a waste of effort.

On the contrary, it is likely that the majority of police supervisors perceive employee coaching as valuable; however, they may struggle to squeeze it in around competing demands. All of the Sergeants that participated in the research project agreed that operational delivery, not coaching conversations, takes priority. This situation is a "folly" (Kerr, 1995, p. 7) which means that whilst NZ Police encourage the delivery of coaching conversations, the organisation's systems motivate supervisors to prioritise operational activity (crime prevention, triaging calls for service, and supervising in the field). That said, the perception of being 'time-poor' is not unique to policing. Lawrence (2017) suggests, in general, that being time-poor, and a lack of coaching skills, are two of the more common reasons why supervisors do not offer coaching to their direct reports.

1.3 Research aim

Utilising LMX as a lens, the aim of this research, in the context of the NZ Police, is to answer the following four questions (see Figure 1 below):

- Is coaching frequency a useful predictor of employee engagement and intention to turnover?
- Is the association between coaching frequency and employee engagement mediated by the perceived quality of the employee coaching relationship (PQECR)?
- Is the association between coaching frequency and intention to turnover mediated by the PQECR?
- Is the association between coaching frequency and the PQECR moderated by employee tenure?

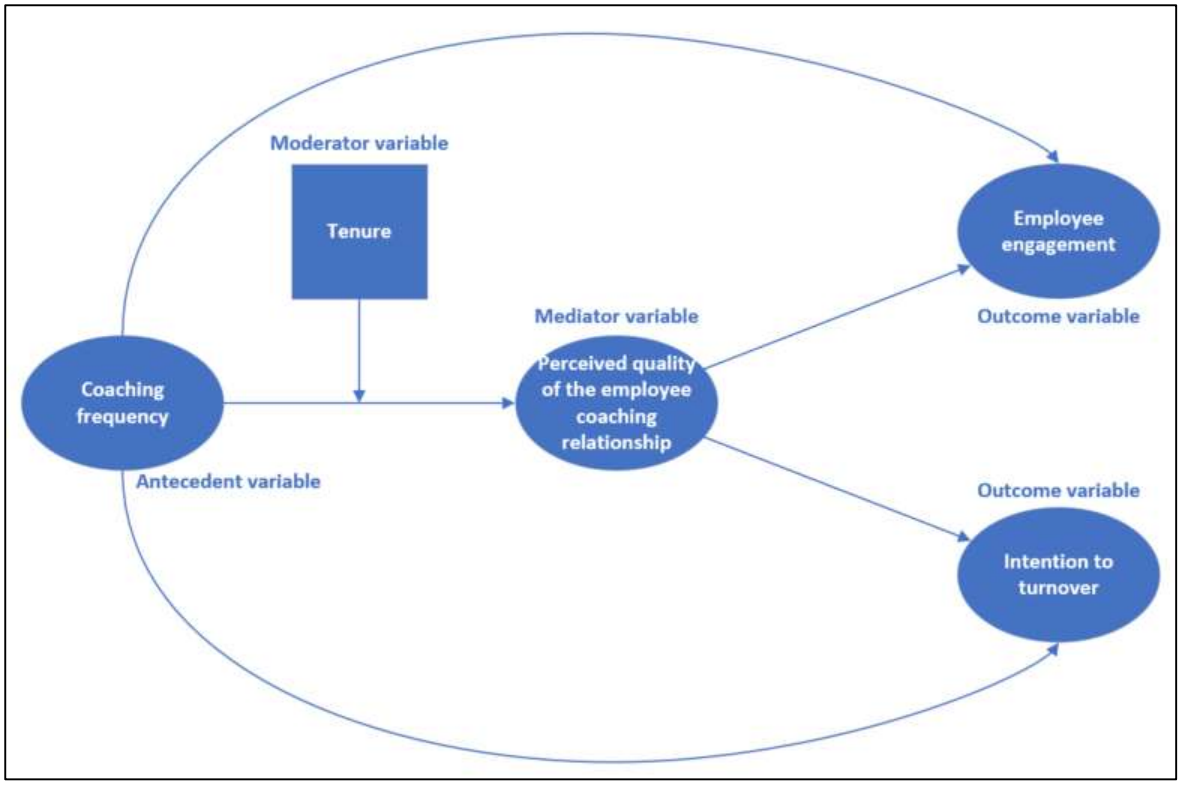

Figure 1: The proposed conceptual model of employee coaching

1.4 Research contributions

This research does not revolutionise how organisations coach their employees; however, it does make a modest contribution for several reasons. At the very least, it answers calls from within the scholarship for further empirical investigation into the practice of employee coaching (Hagen & Peterson, 2014; Lawrence, 2017; Matsuo, 2018). Moreover, it adds to a growing literature that attempts to explain the phenomenon (Hagen & Peterson, 2014; Lawrence, 2017; Matsuo, 2018).

1.4.1 Contribution to theory

This research explains employee coaching through the lens of LMX, developed by Dansereau, Graen and Haga (1975). LMX was selected because it emphasises the importance of the quality of the interpersonal relationship between a leader (coach) and a member (coachee) (Lawrence, 2017; Ribeiro, Nguyen, Duarte, Torres de Oliveira, & Faustino, 2020; Tanskanen, Mäkelä, & Viitala, 2019; Weer, DiRenzo, & Shipper, 2016; Ye, Wang, Wendt, Wu, & Euwema, 2016). Supported by the results of this study (Chapter 4) and the discussion of the results (Chapter 6), LMX has been further validated as a useful theoretical framework for explaining the relationship between coaching frequency, employee engagement, and intention to turnover.

1.4.2 Contribution to method

This study was conducted entirely within the boundaries of the post-positivist paradigm. Whilst no complex, unique, or cutting-edge statistical techniques were employed, it is worth noting that data analysis was conducted using conditional process analysis, developed by Professor Andrew Hayes Hayes (2017). Additionally, three of the study variables were measured using recently developed psychometric scales: the PQECR (Gregory & Levy, 2010) was used to measure the perceived quality of the employee coaching relationship. The EES (Shuck, Adelson, & Reio Jr, 2017) was used to measure employee engagement, and the TISCALE (Dwivedi, 2015) was used to measure intention to turnover. As per the results (Chapter 4), in the context of the data collected for this study, which is independent of the data collected for scale development, all three instruments were found to be valid and reliable. Therefore, this study makes a modest methodological contribution.

1.4.3 Contribution to practice

This research has the potential to make a significant contribution to practice because it is has produced knowledge that can benefit the NZ Police, its employees, and the wider community:

- It has produced empirical evidence of the nature and quality of employee coaching in the context of NZ Police
- It presents the NZ Police Executive with an opportunity to adjust current organisational settings so that employee coaching practice better aligns with the PHPF
- It presents supervisors at all levels of NZ Police with the opportunity to develop a greater appreciation for the benefits of employee coaching, and is likely to motivate many supervisors to engage in the practice more frequently
- The NZ Police executive may adopt the PQECR scale to deselect low-scoring supervisors and support them to develop the mind-set and coaching skills to engender high-quality interpersonal relationships with their direct reports. A targeted approach such as this may be more cost effective than standardised (across-the-board) training (McCarthy & Milner, 2020; Weer et al., 2016)
- The benefits associated with improved employee coaching, such as improved engagement, job-satisfaction, performance, and reduced intention to turnover, will be enjoyed by NZ Police and its employees

1.5 Thesis structure

Looking forward, Chapter 2 reviews the extant literature and provides a concise account of what is currently understood about the practice of employee coaching. To make sense of the phenomena, it is unpacked through the theoretical lens of LMX theory. The literature review concludes with the presentation of a hypothetical model of employee coaching, comprising six proposed (alternative) hypothesis. Chapter 3 describes how this study was designed and the methodology used to test the hypotheses, and the results of the data analysis are presented in Chapter 4. This is followed by a discussion on the implications for the proposed hypotheses in Chapter 5, and in Chapter 6, the thesis is brought to a logical conclusion by concisely synthesising the initial problem, research questions, results, and considering the wider implications.

CHAPTER 2: LITERATURE REVIEW

2.1 Overview

The topic at the centre of this research is the leadership practice of employee coaching, an activity that takes place between a supervisor-coach and each of their coachees. As outlined in the research aim, the study incorporates additional constructs (i.e. engagement, and intention to turnover) and these constructs are also explored in the review.

In the paragraphs that follow, the review synthesises many ideas that have been drawn from the employee coaching scholarship. In particular, it outlines a case for the empirical investigation into the relationship between coaching frequency and employee engagement, and coaching frequency and intention to turnover. Moreover, it posits that these associations are mediated by the perceived quality of the employee coaching relationship (PQECR), and the relationship between coaching frequency and PQECR is moderated by tenure (see Figure 1: the proposed conceptual model of employee coaching).

Following advice from Aguinis, Ramani and Alabduljader (2018), the literature pertaining to the topic of employee coaching was systematically selected and is based on the content of scholarly peer-reviewed journal articles that were located using Massey University's Discover search engine. To keep the search on-topic, search terms were restricted to include articles containing the phrase 'managerial coaching' or 'employee coaching' in the article title. The search returned 406 results, so the following limitations were applied:

- Database providers: Business Source Complete, and Scopus®
- Language of publication: English
- Publication date range: 2016-2021 (to capture the past five years)

With the assistance of these limitations, the search returned 30 articles. One article was a duplicate and two others were excluded because they comprised a basic preview for an upcoming journal publication. Within the remaining 28 articles, 20 captured the core conversation and central thinking on employee coaching, and these articles formed the cornerstone of the managerial/employee coaching aspect of this literature review (refer Appendix B).

2.2 Theoretical foundation

A significant proportion of research within the social sciences is conducted within a particular theoretical framework because it helps to explain the phenomenon being

investigated (Bell, Bryman, & Harley, 2019). In particular, it is expected for those conducting quantitative research (Appelbaum, Cooper, Kline, Nezu, Mayo-Wilson et al., 2018). The theoretical foundation for this research, LMX, is discussed at the beginning of the literature review to provide relevant insight to how employee coaching is understood within this research.

The vagueness surrounding the employee coaching phenomena is acknowledged in the literature (Dahling, Taylor, Chau, & Dwight, 2016; Ismail, Ahmad, & Zainol, 2016; Weer et al., 2016), and the range of theories used to explain it is broad. For example: goal setting theory (Dahling et al., 2016); motivation language based theory (Ismail et al., 2016); regulatory focus theory (Weer et al., 2016); social role theory (Ye et al., 2016); trait activation theory (Kunst, van Woerkom, Poell, & van Kollenburg, 2018); achievement goal theory (Kunst et al., 2018)…and the list continues. This suggests that employee coaching is not well defined because the scholarship is yet to settle on any particular theory that explains it. That said, within the literature reviewed for this study, the most popular theory for explaining employee coaching was LMX theory, adopted in seven of the studies (Ali, Lodhi, Orangzab, Raza, & Ali, 2018; Hsieh & Huang, 2018; McCarthy & Milner, 2020; Pousa, Richards, & Trépanier, 2018; Raza & Ahmed, 2020; Ribeiro et al., 2020; Tanskanen et al., 2019).

Despite the absence of a dominant theory that explains employee coaching, in the post-positivist tradition, it is considered good practice to approach an empirical inquiry from a theoretical angle (Davies & Fisher, 2018). Moreover, citing Liu and Batt (2010), Dahling et al. (2016) argue that the theoretical mechanisms explaining the link between employee coaching behaviour and organisational outcomes are underexplored. Therefore, for the purpose of this study, LMX theory is the lens through which the phenomenon of employee coaching is explained.

The rational for choosing LMX theory is that it focuses on the social exchange relationship between a leader (coach) and a member (coachee) which is considered a crucial element for the success of employee coaching (Gregory & Levy, 2010; Lawrence, 2017; Tanskanen et al., 2019). In their study comparing the effects of pressure-based versus facilitative coaching behaviour, Weer et al. (2016) posit that effective coaching is a reciprocal relationship between a coach and coachee, and success for both parties depends on the quality of the coaching relationship. According to others, the interpersonal relationship is the foundation upon which the coaching relationship is built (Gyllensten & Palmer, 2007); is the catalyst

for driving positive change (Ye et al., 2016); and a crucial element for success in supervisor-employee relations (Graham, Wedman, & Garvin–Kester, 1993). This suggests that for employee coaching to be effective, it relies upon, and must contribute to, a high-quality interpersonal relationship between coach and coachee, and this is something that LMX theory helps to explain.

2.2.1 Leader-member exchange theory

Developed by Dansereau et al. (1975), LMX provides an interpretation of the interpersonal exchange relationship between a coach (leader) and each of their coachees (member) (Ali et al., 2018; Griffin, Phillips, & Gully, 2017; Hsieh & Huang, 2018). According to the theory, coaches tend to split their teams dichotomously into an 'in-group' and an 'out-group' (Griffin et al., 2017) with the in-group members receiving greater negotiating latitude, leadership support, and leadership attention (Dansereau et al., 1975). It is unknown exactly how the coach decides into which group each coachee is placed; however, it is suspected that personality, compatibility, and competence play a role (Griffin et al., 2017). As opposed to out-group coachees, in-group coachees are offered inducements (i.e. work-place coaching) in exchange for their loyalty to the coach and/or increased commitment toward the team/organisation's goals (Dansereau et al., 1975; Raza, Ali, Ahmed, & Moueed, 2017). As suggested by Pousa et al. (2018), LMX it is a sub-category of social exchange theory (Blau, 1964) albeit in the context of the workplace.

When it comes to determining the quality of LMX, subjective perception is what counts. According to (Ali et al., 2018), the ingredients of a high-quality LMX are threefold: attention from the coach; the offer of employee coaching (support); and reciprocation (member exchange) from the coachee. Others describe LMX as both a social and economic exchange (Tanskanen et al., 2019); however, on their own, these explanations fall short because they fail to make clear that the ultimate value of any interpersonal exchange relationship is a subjective judgement (Dansereau et al., 1975). Indeed, LMX can only be considered high-quality when both parties perceive it to be that way. Weer et al. (2016, p. 193) describe LMX as "positive relations and rapport between supervisors and subordinates." This is helpful because it emphasises the subjective perception of the exchange relationship. Others agree, suggesting that a high-quality LMX is characterised by honesty, trust, and mutual respect between the coach and the coachee (Ali et al., 2018; Hsieh & Huang, 2018; Pousa et al., 2018; Raza & Ahmed, 2020; Ribeiro et al., 2020).

There appears to be a close connection between LMX and employee coaching. When investigating the predictive ability of employee coaching on performance among a sample of 879 multi-organisational employees in Finland, Tanskanen et al. (2019) found the correlation between employee coaching skill and LMX to be strong and statistically significant, r=0.830, p<0.001. A similar association was drawn by Barry, Gloeckner and Kaiser (2021) in their qualitative study of coaching competencies. Barry et al. (2021, p. 211) suggest that effective coaches invest time in their coachees, really getting to know them, and the ability to provide a "supportive environment" for the coachee is absolutely essential for employee coaching to be effective. In their study, Dansereau et al. (1975) reported that the turnover rate for study participants was 38% less for those deemed part of the in-group compared with out-group participants. Unsurprisingly, a similar pattern was observed in the sample data collected for this study. Indeed, the negative effect of PQECR on intention to turnover was found to be moderate (Chapter 4), suggesting that those who perceive the quality of the coaching relationship they share with their supervisor to be strong are less likely to resign when compared to those who perceive it to be weak.

These above examples support the proposition that LMX is a useful lens for explaining the association between coaching frequency, employee engagement, and intention to turnover.

2.3 Employee coaching

Whilst it is tempting to begin with an unequivocal definition of employee coaching, the approach taken here is to start with context because it will help the reader to gain an appreciation for the complexity of employee coaching and the rationale for selecting the particular definition adopted for this study. Indeed, the same strategy is applied to the discussion for each of the variables comprising the conceptual model of employee coaching.

Looking back, the wave of interest in employee coaching (in the west at least) began to form in the late 1970s, following the publication of Ferdinand Fournies book, Coaching for Improved Work Performance (Fournies, 1978). Since then, many organisations have adopted a 'supervisor-as-coach' model (Ali et al., 2018; Barry et al., 2021; Dahling et al., 2016; Raza & Ahmed, 2020; Weer et al., 2016; Ye et al., 2016; Zhao & Liu, 2020; Zuñiga-Collazos, Castillo-Palacio, Montaña-Narváez, & Castillo-Arévalo, 2020) and coaching employees has become an essential part of the supervisor's day-to-day role (Ali & Aziz, 2018; Barry et al., 2021; Gregory & Levy, 2010; Hagen & Peterson, 2014; Lawrence, 2017; Liu & Batt, 2010; Ribeiro et al., 2020; Weer et al., 2016). Citing the Chatered institute of

personnel and development (2015), Lawrence (2017) says that 80% of organisations in the United Kingdom have an expectation that supervisors will coach their direct reports. This figure has support in the literature. In a study investigating the influence of employee coaching on coachee goal-orientation-personality-traits among a sample of 521 vocational educators based in the Netherlands, Kunst et al. (2018) found that 75.5% of the coachee-participants had informal meetings with their coaches at least once a week, and 63.5% reported having a formal meeting with their coach at least once a month. In another example, a study of large Australian Organisations found that of the 580 Human Resource managers and General managers sampled, 94% coached their direct reports and 73% of those that coached did so at least once per week (McCarthy & Milner, 2020).

Academic interest in employee coaching accelerated in the 1990s after employee coaching was identified as a new approach to leadership in organisations (Gregory & Levy, 2010; Hagen & Peterson, 2014; Matsuo, 2018). For example, institutions like the Coaching Psychology Unit, University of Sydney (University of Sydney, n.d.), and the International Centre for Coaching and Mentoring Studies (Oxford Brookes University) have been set up specifically for coaching teaching and research.

Arguably, the main driver behind the coaching movement is the link between employee coaching and effective management (Barry et al., 2021; Ribeiro et al., 2020) with some commenting that employee coaching is a human resource expert's single most important function (Ali et al., 2018). Along with the positive association with effective management, other drivers behind the wave include technology innovation (Ye et al., 2016), and competitive pressure brought on by globalisation (Ismail et al., 2016; Ribeiro et al., 2020; Tanskanen et al., 2019). The perception here is that the supervisor-as-coach model is more cost effective than the traditional training-model because supervisors typically have day-to-day contact with each of their direct reports and can deliver personalised training to meet the individual's specific competency needs (McCarthy & Milner, 2020; Weer et al., 2016). Others agree, suggesting that the training-model is closed-ended and typically delivered to groups whereas employee coaching is informal, individualised, and continuous (Dahling et al., 2016). It has also been posited that employee coaching is less susceptible to knowledge-transfer-loss because it takes place within the context of the job (Weer et al., 2016). Also, organisations are becoming more conscious of being able to sustain employees through periods of change, and employee coaching is recognised as an effective tool to achieve this (Zuñiga-Collazos et al., 2020). As the above examples demonstrate, it is not unreasonable

to suggest that what started as a gentle wave of employee coaching is fast becoming a tsunami.

2.3.1 Types of workplace coaching

Within the literature, three types of workplace coaching relationships appear most frequently: executive coaching, peer coaching, and employee coaching (Dahling et al., 2016; Gregory & Levy, 2010; Hagen & Peterson, 2014; Weer et al., 2016), but how do they differ? Executive coaching refers to the relationship between an executive employee (coachee) and an external consultant (coach) (Gregory & Levy, 2010; Weer et al., 2016); peer coaching refers to the relationship between two employees at the same level where one performs the role of coach and the other the coachee (Gregory & Levy, 2010); and employee coaching refers to the relationship between an employee (coachee) and their direct supervisor/manager (coach) (Gregory & Levy, 2010; Weer et al., 2016; Ye et al., 2016). The main differences between these types of coaching relationships are selection and power differential. Executive coachees typically select their own coach and the power rests with the executive (Lawrence, 2017). In the case of peer-coaching, the coach and the coachee may select each other and power is shared because coach and coachee are at the same level of position and do not report to each other (Gregory & Levy, 2010). In the context of employee coaching, it is less likely that the either party selects the other (Lawrence, 2017), and there is a definite power imbalance (Lawrence, 2017) because the coachee also formally reports to their coach (Gregory & Levy, 2010). These features distinguish employee coaching from the other main types of workplace coaching and confirms it has a unique set of challenges.

2.3.2 Is employee coaching effective?

Employee coaching is effective. Within the literature, numerous studies advocate for employee coaching as a means of enhancing organisational performance (Hicks & McCracken, 2013; Hunt & Weintraub, 2016; Longenecker, 2010), improve personal gratification (McCarthy & Milner, 2020), and job satisfaction (Hagen & Peterson, 2014). Indeed, citing Hamlin, Ellinger and Beattie (2006), it has been suggested that “coaching is the heart of managerial effectiveness.” (Matsuo, 2018, p. 128), and “organizational success can be deeply influenced by how well managers coach subordinates.” (Ribeiro et al., 2020, p. 2). Moreover, from the perspective of the coachee, attention from the coach is considered a valuable job resource (Tanskanen et al., 2019) and can have a positive effect the coachee’s

perception of belonging to their organisation (Chen et al., 2019; Raza & Ahmed, 2020). In support of the above, data obtained from 1,246 pharmaceutical sales employees based in the USA revealed a positive association between employee coaching skill and sales goal attainment. The correlation between these two variables was moderate (r=0.42) with employee coaching skill having a meaningful influence on sales goal attainment (Dahling et al., 2016). In another example, employee coaching behaviour was found to positively correlate with workplace wellbeing for a randomly selected sample of 279 employees belonging to a large state-owned enterprise in China (r=0.63, p<0.01). Moreover, a path analysis revealed that employee coaching had a sizable effect on workplace wellbeing for this sample (r=0.57, p<0.01) (Zhao & Liu, 2020). Such evidence confirms that employee coaching can be effective and helps to explain why the practice has grown.

2.3.3 Barriers to employee coaching

Despite the purported benefits associated with employee coaching, a number of barriers have been identified. A survey of 580 coaches based in large Australian organisations identified the top six (McCarthy & Milner, 2020):

- Lack of opportunity: some coaches struggle to balance the demands of their dual role as coach and manager
- Hesitancy to coach: some coaches are fearful of a backlash from their coachees, particularly if the feedback is perceived to be critical rather than constructive
- Reluctant coachees: some coachees prefer to be given direct instructions rather than to be coached
- Lack of coaching ability (skill): some coaches lack the necessary coaching competencies
- Lack of motivation: some coaches resist devolving power and responsibility to their coachees
- Lack of organisational support: the structures and systems of some organisations restrain coaching practice

The above list of barriers is non-exhaustive. It confirms that employee coaching is a complex leadership practice that requires planning, courage, patience, and persistence from the coach to be successful.

2.3.4 Tension within the scholarship

Having discussed the effectiveness of employee coaching and touched on some of the well-recognised barriers to its application, it is worthwhile mentioning the tension that exists within the employee coaching scholarship because it is important to appreciate that universal agreement has not been reached. Lawrence (2017) note the diverse opinions as to what employee coaching is, how it should be applied, and how to measure it. For example, there is disagreement on the impact of the hierarchical power imbalance between the supervisor-coach and subordinate-coachee (Matthews, 2010) with some saying it prevents trust from fully developing (Whitmore, 2002) whilst others argue that provided the coach is genuine in their support of the coachee, sufficient trust can be built regardless of any power imbalance (Hunt & Weintraub, 2016). Another area of conjecture centres on the effectiveness of different coaching skills/behaviours. According to Dahling et al. (2016), effective employee coaching consists of three distinctive behaviours: continual constructive feedback; role-modelling good performance; and collaborative goal setting. Others dispute this. In their study investigating the effect of employee coaching on individual-goal-orientation among educators, Kunst et al. (2018) found that coaches who delivered constructive feedback (defined as guidance coaching) failed to stimulate the success-focused-goal-orientation-trait in coachees in the same way as coaches who operated as sounding boards (defined as facilitative coaching). That said, tension within the scholarship can be viewed as a healthy sign. It means the topic is considered important enough to argue about and there are knowledge gaps to be filled.

2.3.5 Calls for further investigation

A consistent theme running through the workplace coaching literature is the lack of studies into employee coaching and calls for further empirical investigation (Hagen & Peterson, 2014; Lawrence, 2017). Whilst it is agreed that gaps exist, there does not appear to be dearth of empirical studies. A search for peer reviewed articles containing the phrase ‘managerial coaching’ or ‘employee coaching’ in the title returned 406 results. This tends to confirm that the concept of “manager-as-coach” is becoming increasingly popular among the scholarship (Dahling et al., 2016, p. 863).

Despite the steady flow of empirical investigations into employee coaching, gaps do remain. For example, Echeverri (2020) argues that employee coaching is a two-way process and that the scholarship has paid insufficient attention the coachee’s perception of its value.

Echeverri's qualitative study of 34 tram/bus drivers in Sweden found that employee coaching is not always perceived as valuable from the coachee's point of view. Rather, value is dependent upon the degree to which the coach and coachee understand, align, and commit to the organisation's formal rules (Echeverri, 2020). The identification of knowledge gaps is a good thing. It confirms that academic interest in employee coaching is robust and provides fruitful opportunity for future research.

2.3.6 Defining employee coaching

Having now discussed the context of employee coaching in some detail, it is time to confirm a definition of employee coaching for this study. To this end, the definition developed by Gregory and Levy (2010, p. 111) has been adopted:

> *A developmental activity in which an employee works one-on-one with his/her direct manager to improve current job performance and enhance his/her capabilities for future roles and/or challenges, the success of which is based on an effective relationship between the employee and manager, as well as the use of objective information, such as feedback, performance data, or assessments.*

This definition is preferred because it serves the aim of the research. In particular, the definition's emphasises on the importance of the interpersonal relationship supports the study's intent to measure the mediating effect of the PQECR. Secondly, the definition aligns with LMX theory which emphasises the importance of the interpersonal relationship. Thirdly, the definition is a good fit with NZ Police's PHPF policy because it also perceives employee coaching to be dyadic, meaning a one-on-one interaction between the coach and the coachee, and PHPF applies to employees at all levels of the organisation (including executives) (New Zealand Police, 2017). That said, the fit is not perfect. Whereas Gregory and Levy (2010) posit that employee coaching can enhance a coachee's capability for both current and future roles, NZ Police are clear that PHPF focuses on the coachee's current role only (New Zealand Police, 2017). Notwithstanding this minor discrepancy, the Gregory and Levy definition is fit for purpose.

2.3.7 The relevance of coaching frequency

For employee coaching to have an effect on a coachee's attitude and/or behaviour, it must necessarily take place on at least one occasion, yet there has been little investigation into the effect that coaching frequency has on organisational outcomes. For the most part, the scholarship has investigated the links between coaching competencies and organisational

outcomes (Chen et al., 2019; Hsieh & Huang, 2018; Kunst et al., 2018; Matsuo, 2018; Ribeiro et al., 2020; Weer et al., 2016) whilst largely ignoring coaching frequency. This is a gap in the literature that this study addresses by investigating the effect of coaching frequency on employee engagement and intention to turnover.

Despite the paucity of studies into the effect of coaching frequency, the grounds for investigating this effect are strong. Drawing on the earlier work of London and Smither (2002), Gregory and Levy (2010) make clear that to be effective, employee coaching must be regular and ongoing. Others describe employee coaching as a continuous process (Dahling et al., 2016). Ribeiro et al. (2020); and Ye et al. (2016) agree, saying that the provision of regular feedback is an important dimension of employee coaching. Further support for this comes from Matsuo (2018) who argues that employee coaching is part of the day-to-day relationship between coaches and their coachees. These arguments lend support to the idea that frequency does matter. Afterall, how can employee coaching be regular, ongoing, or day-to-day if it occurs infrequently?

Assuming that coaching frequency does matter, how is it proposed to work? The rationale is that when coaching frequency increases, so do the opportunities to build and maintain a high-quality interpersonal relationship. For example, in their study that qualitatively explored the experiences of coaches, Barry et al. (2021) report that the time spent getting to know coachees, particularly in the early stages of the relationship, paid dividends. It is clear that employee coaching is more sophisticated than a "one and done" process (Barry et al., 2021, p. 212). In support, Zuñiga-Collazos et al. (2020) suggest that employee coaching aims for 'sustained' change in a coachee's attitude and behaviour, and it is difficult to accept that this can be achieved without high-frequency collaboration.

Whilst the coaching frequency hypothesis has appeal, there is empirical evidence that it makes little difference. When investigating the effect of employee coaching on sales goal performance, Dahling et al. (2016) reported that coaching frequency on its own (high and low) had no statistically significant effect on coachee performance. The authors make their position clear; it is the quality of coaching that matters, not the quantity. This is slightly contradictory given that the authors also define employee coaching in terms of "providing continual constructive, developmental feedback to subordinates…" (Dahling et al., 2016, p. 867), and it begs the question whether continual constructive feedback can also be infrequent? That said, others support the same position. Pousa et al. (2018) say that it is the way coaching is done that makes the difference, not the act alone. Both Dahling et al. (2016),

and Pousa et al. (2018) make a fair point that the quality of employee coaching certainly matters; however, it is difficult to imagine building and maintaining a high-quality interpersonal relationship with another person without frequent interaction. Indeed, most will agree that frequent collaboration is the mechanism that turns a doubtful-stranger into a trusted-friend, so whilst Dahling et al. (2016) found that coaching frequency did not a have a statistically significant effect on sales goal performance, this was not the case for the current study. Indeed, the effect of coaching frequency on PQECR, employee engagement, and intention to turnover was found to be statistically significant, albeit small (see Chapter 4).

Notwithstanding the existence of evidence suggesting coaching frequency does not matter, the current policy within NZ Police (PHPF) is for every supervisor to schedule formal coaching conversations with each of their direct reports at least once per month, supplemented by an annual performance review once per year (New Zealand Police, 2017). The rationale for the monthly cycle is not specified which makes one wonder if it is based on empirical evidence or does a monthly cycle just seem like a good fit? Regardless, it does signal that the frequency to which supervisors coach their employees does matter to NZ Police.

Within the scope of the literature review, the only study that incorporated coaching frequency as a variable was the study conducted by Dahling et al. (2016). Moreover, none of the studies investigated the effect of coaching frequency on the PQECR, employee engagement, or intention to turnover. A wider review may have captured additional studies; however, this tends to suggest that the effect of coaching frequency is underexplored, something that this study seeks to address.

2.4 Perceived quality of the employee coaching relationship (PQECR)

Like the previous section on employee coaching, the temptation to immediately define the PQECR has been resisted in favour of a context first approach.

Drawing on the work of (Kilburg, 2001), Gregory and Levy (2010), explain that for employee coaching to be effective, the coachee must perceive the coaching relationship to be genuine, built on mutual respect, trust, and empathy. In other words, the coachee must perceive the coaching relationship to be high-quality. From this perspective, the critical mechanism that determines the success of an employee coaching relationship is not an objective measure of a coach's ability, or how often coaching takes place. Rather, it depends

on the coachee's subjective perception of the quality of the coaching relationship (Gregory & Levy, 2010). Therefore, to achieve the aim of this study, a key variable of interest is the coachees' 'perception' of the quality of coaching relationship they share with their supervisor.

2.4.1 Relevance of the PQECR

The PQECR is highly relevant because the success of both coach and coachee depends upon it being strong (Chen, Sharma, Edinger, Shapiro, & Farh, 2011; Lawrence, 2017; Tanskanen et al., 2019; Weer et al., 2016; Zuñiga-Collazos et al., 2020). As per LMX theory (Section 2.2.1), when a high-quality interpersonal relationship is established, a coachee will likely reciprocate with increased commitment to the tasks at hand and less intention to turnover (McCarthy & Milner, 2020; Raza et al., 2017; Ribeiro et al., 2020; Tanskanen et al., 2019; Weer et al., 2016). Furthermore, it is argued that a coachee's perception of their coach has a strong effect on the coachee's workplace behaviour and attitude (Raza & Ahmed, 2020). Others support this view, saying that the quality of the interpersonal relationship is the catalyst for driving change (Tanskanen et al., 2019; Ye et al., 2016). Interestingly, results from a qualitative study that sampled 580 coaches based in large Australian organisations revealed that coachees enjoy it when they get attention from their coach, regardless of the nature of the conversation (McCarthy & Milner, 2020). Indeed, when a coachee perceives that their coach is genuinely interested in them as a person, the coachee is typically happier, and more engaged (McCarthy & Milner, 2020). This suggests that if a coach and a coachee do not get along on a personal level, the coach's best-practice coaching competencies are unlikely to change the coachee's attitude or behaviour. This is because employee coaching is more than an intellectual endeavour, and a coach's skill will have less impact if the interpersonal relationship with their coachee is laced with mistrust and apprehension.

It is somewhat surprising that there has been little investigation into the PQECR. Within the scope of the literature review, several studies were found that measured the coachee's perception of their coach's coaching skill/behaviour (Chen et al., 2019; Hsieh & Huang, 2018; Kunst et al., 2018; Matsuo, 2018; Ribeiro et al., 2020; Weer et al., 2016); however, none were found that measure the coachee's perception of the quality of the coaching relationship itself. This tends to suggest the effect of the PQECR is not well understood. Unsurprisingly, there has been a call to investigate the mediators of employee coaching generally (Pousa, Liu, & Aman, 2020), and in particular, the PQECR (Echeverri, 2020; McCarthy & Milner, 2020; Ribeiro et al., 2020). Therefore, a study such as this is likely to

be welcome by the scholarship because it investigates the mediating effect of the PQECR on employee engagement and intention to turnover, and by doing so, it helps to fill a gap in the literature.

2.4.2 Defining the employee coaching relationship

For the purpose of this study, the definition of the employee coaching relationship developed by Gregory and Levy (2010, p. 111) has been adopted:

> *A working partnership between an employee and his/her supervisor that is focused on addressing the performance and developmental needs of that employee.*

The reason why this definition is appropriate is because is captures the 'partnership' dimension of a successful coaching relationship. It is not a case of the supervisor instructing the employee what they need to do to improve, and the employee obeying (one way). Rather, the supervisor and employee work together to identify performance/development gaps and collectively agree on a plan to close them (two-way).

2.4.3 Coaching frequency and PQECR

It is hypothesised that there is a positive association between coaching frequency and the PQECR. The basis for this is threefold. Firstly, as per LMX theory, when making an offer to coach, the coach initiates a social exchange (Ali et al., 2018; Hsieh & Huang, 2018; Raza & Ahmed, 2020; Raza et al., 2017). Secondly, as the frequency to which coaching is offered increases, so do the opportunities to build and maintain a high-quality interpersonal relationship. Indeed, coaching is perceived as a valuable job resource (Ali et al., 2018; Ribeiro et al., 2020; Tanskanen et al., 2019) and a signal that the coachee has achieved insider status (Zhao & Liu, 2020), so the more frequently coaching takes place, the more opportunities there are for value to be exchanged (Hsieh & Huang, 2018). Finally, as the quality of the interpersonal relationship between the coach and the coachee strengthens, it is expected that the PQECR will do the same. The rationale for this is that a coachee who enjoys a high-quality exchange relationship with their coach is more likely to judge the coaching relationship more favourably than a coachee who has a low-quality exchange relationship with their supervisor (Pousa et al., 2018; Ribeiro et al., 2020). Therefore, it is hypothesised that when employee coaching frequency increases, it will have a positive effect on the PQECR.

- **H.1**: There is a positive association between coaching frequency and the perceived quality of the employee coaching relationship

2.5 Employee engagement

The reason why employee engagement has been chosen as an outcome variable for this study is because the target organisation, NZ Police, has a legislative monopoly over the provision of policing services across New Zealand, and given the organisation has no competitors, but does experience high demand for service, employee engagement is likely to be a more useful indicator of organisational performance than market share or profit and loss. In addition, NZ Police have an interest in employee engagement, evidenced by the measurement of this variable in an organisational survey conducted in 2017, so the measurement of employee engagement aligns with the interests of the organisation.

2.5.1 The relevance of employee engagement

The degree to which employees are 'engaged' in the context of their work experience has gained considerable attention in recent times (Shuck et al., 2017) and this has resulted in calls for further research on the effect of leadership behaviour (i.e. employee coaching) on employee engagement (Bakker, Albrecht, & Leiter, 2011). One of the main drivers behind the call is a belief that engaged employees perform better than ambivalent or disengaged employees (Ali et al., 2018; Hagen & Peterson, 2014; Ismail et al., 2016; McCarthy & Milner, 2020; Shuck et al., 2017; Tanskanen et al., 2019; Wagstaff et al., 2018). This association also explains why many organisations, including NZ Police, strive for an engaged workforce. Indeed, from an engagement survey, conducted in 2017, NZ Police reported that 28% of its employees were engaged, 56.4% were ambivalent, and 15.6% were disengaged (Gallup, 2017). Such results highlight room for improvement.

Along with improved performance, engaged employees appear to be happier. Drawing on the previous research of Kahn (1990); and Rich, Lepine and Crawford (2010), Shuck et al. (2017) suggest that engaged employees are cognitively attentive, emotionally vested, and physically energetic in the work environment. McCarthy and Milner (2020) support this view, saying that engaged employees are happier at work; they are more positive, get promoted, and earn more money. This suggests that when employee engagement is strong, the situation is win-win for the organisations, employees, and the wider community.

2.5.2 Defining employee engagement

The concept of engagement is broad, and it is fair to say that there is significant variety within the scholarship as to what it means for an employee to be engaged. Of note, there are several dimensions of engagement highlighted in the literature: job engagement, work engagement, and organisational engagement (Shuck et al., 2017). This may explain why there are multiple psychometric scales that measure engagement in different ways. Examples include the Job Engagement Scale (JES) produced by Rich et al. (2010), that focuses on the employee's engagement with their specific job; the Utrecht Work Engagement Scale (UWES) developed by Schaufeli, Bakker and Salanova (2006), that focuses on an employee's engagement with their work activity; and the Organizational Engagement Scale (OES), produced by Saks (2006), that focuses on the employee's engagement with their organisation.

For the purpose of this study, the definition of the employee engagement developed by Shuck et al. (2017, p. 955) has been adopted:

> *"An active, work-related positive psychological state operationalized by the intensity and direction of cognitive, emotional, and behavioural energy."*

This definition is preferred because it is grounded in the personal engagement framework developed by Kahn (1990), and is not confined to any particular engagement dimension. Moreover, Shuck et al. (2017) describe a highly engaged employee in the following way: attentive and concentrated; directing mental energy into their work; connected with their work experience; believing in their organisation's vision; ready and willing to go the extra mile; and believing that being a member of the organisation gives their life meaning. On the face of it, this strikes as a thoroughly engaged employee and is precisely what this study strives to measure.

2.5.3 Coaching frequency and employee engagement

It is hypothesised that there is a positive correlation between coaching frequency and employee engagement. The basis for this is that as coaching frequency increases, it becomes more effective (Gregory & Levy, 2010; Matsuo, 2018; Weer et al., 2016) and makes the coachee feel well supported (Ali et al., 2018; Tanskanen et al., 2019), and the more supported a coachee feels, the more engaged they are likely to be (McCarthy & Milner, 2020; Tanskanen et al., 2019; Weer et al., 2016). Therefore, it is expected that as coaching frequency increases, there will be a positive effect on employee engagement.

- **H.2**: There is a positive association between coaching frequency and employee engagement.

2.5.4 PQECR and employee engagement

It is hypothesised that a positive correlation exists between the PQECR and employee engagement. The rationale for this is that the PQECR is a subjective measure of the quality of the coaching relationship (Gregory & Levy, 2010, p. 118), and when the coachee feels well-supported by their coach/organisation, they are more likely to be engaged in their work (McCarthy & Milner, 2020; Tanskanen et al., 2019; Weer et al., 2016). Therefore, it is expected that as the PQECR increases, there will be a positive effect on employee engagement.

H.3: There is a positive association between the perceived quality of the employee coaching relationship and employee engagement.

2.6 Intention to turnover

The reason why intention to turnover was selected as an outcome variable for this study is that NZ Police is a not-for-profit government sector organisation, and turnover rate (among others) is a useful indicator of organisational wellbeing (Forsyth, Ashby, Gardner, & Tappin, 2021). Whilst not the same as intention to turnover, since 2017, the actual turnover rate for constabulary NZ Police employees has reduced by 1.4% from 5.3% in 2017 to 2.9% in 2021; for non-constabulary employees, the turnover rate has reduced by 2.4% from 10.4% in 2017, to 8.0% in 2021 (New Zealand Police, 2021). This is comparatively low considering the average termination rate for organisations across New Zealand in 2018 was reported as 20.5% (Lawson Williams, 2019, p. 4). Given that NZ Police regularly report the organisation's turnover rate in its annual report, it is likely that gaining insight into employee intention to turnover will be of considerable interest to the organisation.

2.6.1 The relevance of intention to turnover

Employee turnover can be a serious problem for organisations. One of the more obvious maladies is the financial cost associated with recruiting, inducting, and training new employees, followed by the loss of productivity when existing employees leave (Dwivedi, 2015; Stone, 2017). Referring to an example from Holman, Batt and Holtgrewe (2007), Dwivedi (2015), describes a situation in a call centre based in India in 2007 where 60 out of 100 new employees quit the organisation within their first year of employment. Given that it takes NZ Police five months to train every new Constabulary employee (not including pre-

training costs), an outcome such as this would be disastrous. Another malady is that coachees with a strong intention to turnover tend to psychologically withdraw from the organisation, reducing the amount of effort they are prepared to invest into their work (Ali et al., 2018).

In addition to the financial cost, a high employee turnover rate can have a negative effect on an organisation's reputation as a good place to work, meaning it may attract less talent (Stone, 2017). It is understandable why employee turnover is considered a dysfunctional behaviour, particularly when the departing employee is a high performer (Griffin et al., 2017). Considering the above, organisations across the globe are interested in the antecedents of employee turnover and how to prevent it (Dwivedi, 2015); however, it appears there are no stand-out reasons why employees choose to quit their jobs (Dwivedi, 2015). Rather, at the present time, the scholarship has identified 'intention to turnover' as the most accurate predictor of actual employee turnover (Bluedorn, 1982; Mobley, Griffeth, Hand, & Meglino, 1979).

2.6.2 Defining intention to turnover

Drawing on the earlier work of Price (1977), Dwivedi (2015) defines employee turnover (distinct from intention to turnover) as a ratio of the number of employees that have left an organisation compared with the average number of employees that remained in the organisation over a specified period of time. Citing Boshoff, Van Wyk, Hoole and Owen (2002); Currivan (1999), Dwivedi (2015) defines intention to turnover (distinct from employee turnover) as the strength of one's attitude and cognitive drive to quit an organisation. This includes thinking of quitting, intending to look for another job outside the organisation, and intending to quit. For the purpose of this study, this definition of intention to turnover has been adopted.

2.6.3 Coaching frequency and intention to turnover

It is hypothesised that there is a negative correlation between coaching frequency and intention to turnover. The basis for this is that as coaching frequency increases, it becomes more effective (Gregory & Levy, 2010; Matsuo, 2018; Weer et al., 2016) and makes the coachee feel well supported (Ali et al., 2018; Tanskanen et al., 2019), and the more supported a coachee feels, the less intent they will have to quit their job (McCarthy & Milner, 2020; Ribeiro et al., 2020; Tanskanen et al., 2019; Weer et al., 2016). Therefore, it is expected that as coaching frequency increases, there will be a negative effect on intention to turnover.

- **H.4**: There is a negative association between coaching frequency and intention to turnover.

2.6.4 PQECR and intention to turnover

It is hypothesised that a negative correlation exists between the PQECR and intention to turnover. The rationale for this is that the PQECR is a subjective measure of the quality of the coaching relationship (Gregory & Levy, 2010, p. 118), and when the coachee feels that their coaching relationship is valuable, they are more likely to have a low intention to turnover (McCarthy & Milner, 2020; Ribeiro et al., 2020; Tanskanen et al., 2019; Weer et al., 2016). Therefore, it is expected that as the PQECR increases, the effect on intention to turnover will be negative.

- **H.5**: There is a negative association between the perceived quality of the employee coaching relationship and intention to turnover.

2.7 Tenure

The rationale for choosing tenure as a moderating variable for this study is that little is known about its effect on the linear relationship between coaching frequency and PQECR. In support of this, Pousa et al. (2018) have suggested that coachee tenure could have a significant moderating effect on the linear relationship between employee coaching and other organisational outcomes because coachees with less experience may be affected by employee coaching to a greater extent than those with more tenure. Moreover, other scholars have noted that conditions such as coachee tenure are likely to moderate the effect of employee coaching on organisational outcomes, and there have been calls to investigate the effect (McCarthy & Milner, 2020; Weer et al., 2016).

2.7.1 The relevance of tenure

Logically, coachees who are new hires and still 'learning the ropes' are likely to be more dependent on their supervisor-coach than peers who have more tenure (Pousa et al., 2018), and this may impact on how the coachee perceives the quality of the coaching relationship with their coach. By the same logic, coachees with more tenure are less likely to depend on their coach for advice and support because their depth of knowledge, skill, and ability is likely to develop with experience. The question here is how does a coachee's tenure affect the linear relationship between coaching frequency and PQECR? Does the slope of the regression line increase, decline, or remain static as the coachee's tenure extends?

2.7.2 Defining tenure

For the purpose of this study tenure is defined as the length of time, measured in years, a person has been employed by an organisation (*Dictionary.com*, 2021).

2.7.3 Coaching frequency, PQECR, and tenure

From the literature, no evidence was found to suggest how coachee tenure might moderate the linear relationship between coaching frequency and the PQECR, which is a gap. That said, from a position of logic, it is suspected that as coachee tenure increases, their knowledge, skill, and ability will develop, meaning they will rely less on their supervisor-coach for support. Therefore, it is hypothesised that as a coachee's tenure increases, it will have a negative moderating effect on the linear relationship between coaching frequency and the PQECR.

- **H.6**: The association between coaching frequency and the perceived quality of the employee coaching relationship is negatively moderated by tenure.

2.8 The hypothetical model of employee coaching

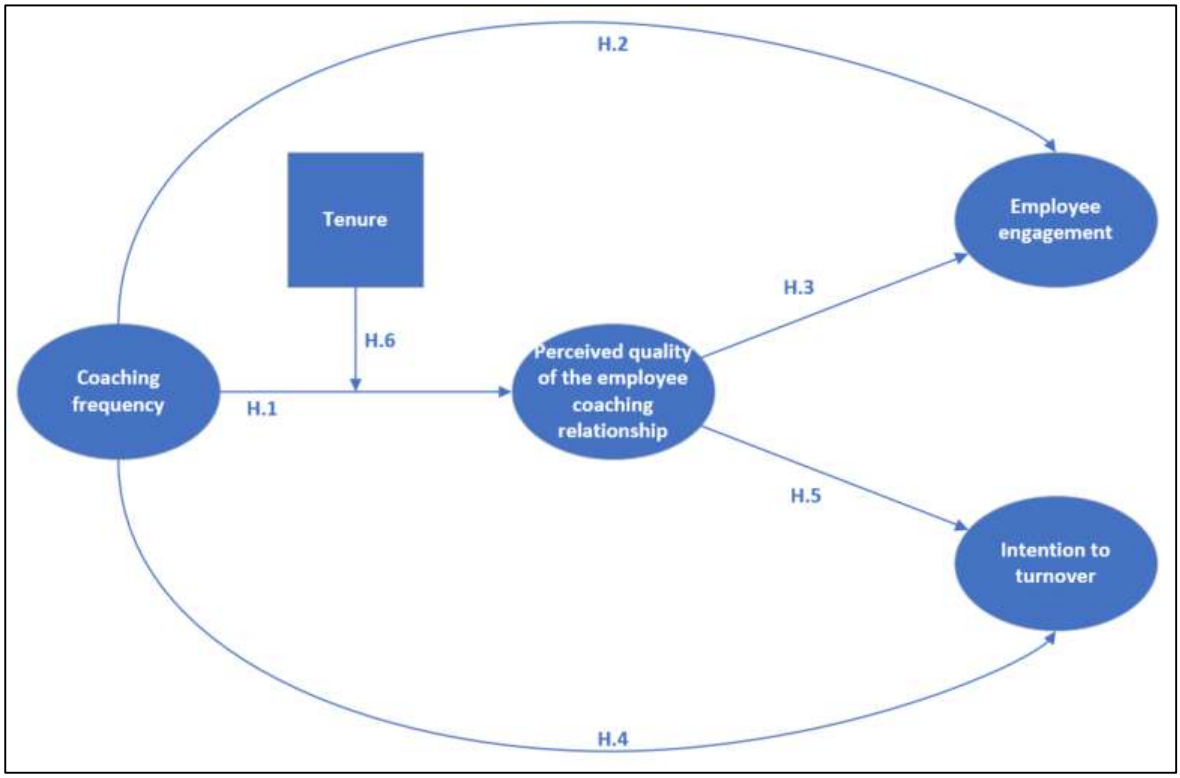

Figure 2: The hypothetical model of employee coaching

2.8.1 Hypotheses

- H.1: There is a positive association between coaching frequency and the perceived quality of the employee coaching relationship
- H.2: There is a positive association between coaching frequency and employee engagement
- H.3: There is a positive association between the perceived quality of the employee coaching relationship and employee engagement
- H.4: There is a negative association between coaching frequency and intention to turnover
- H.5: There is a negative association between the perceived quality of the employee coaching relationship and intention to turnover
- H.6: The association between coaching frequency and the perceived quality of the employee coaching relationship is negatively moderated by tenure

CHAPTER 3: STUDY DESIGN AND METHODOLOGY

Research paradigms can be conceived as distinct frameworks of interrelated assumptions, theories, and values upon which knowledge is built (Kuhn, 1970), and this study has been conducted within the post-positivist paradigm because post-positivism aligns with the researcher's worldview. One of the assumptions shared by post-positivists is the existence of a single objective reality (Davies & Fisher, 2018). Moreover, post-positivists presume that humanity is unable to observe reality directly because our cognitive and physical limitations constrict our ability to achieve this (Panhwar, Ansari, & Shah, 2017). Indeed, when discussing linear regression modelling, Professor Andrew Hayes makes clear that whilst statistical models are useful tools, they remain approximations of reality (Hayes, 2017). From this perspective, the absolute truth about anything (i.e. employee coaching) can never be known (Heywood, 2013; Panhwar et al., 2017); however, this does not prevent post-positivists researchers from trying to get as close to the truth as possible and explains the preoccupation with the replication of results (Bell et al., 2019).

As the name suggests, post-positivists tend to build knowledge about the social world by observing and measuring phenomena from the outside, indirectly, using the scientific method (Bell et al., 2019). This often means engaging in detached, objective measurement of social phenomena (i.e. employee coaching) using remote observation, questionnaires, and psychometric scales (among other techniques) to collect data from large samples of participants, and using statistical techniques (i.e. linear regression modelling) to identify patterns of association (Bell et al., 2019). Moreover, like their positivist cousins, post-positivists tend to be deductive, meaning they start with a theory, develop a hypothetical model, and confirm or reject the hypotheses depending on how well the hypothetical model fits the sample data and whether the observed effects are statistically significant (Davies & Fisher, 2018).

A significant difference between post-positivists and their positivist cousins is that post-positivists tend to be more sceptical of the idea that the scientific method alone can be relied upon to reveal the absolute truth. Indeed, post-positivists tend to be more open to interpretivist perspectives to inform their research and garner support for their results (Davies & Fisher, 2018; Panhwar et al., 2017). Unlike post-positivists, interpretivists reject the single reality ontology, positing that every person has their own unique reality, and each is valid (Davies & Fisher, 2018). This multi-reality worldview enables interpretivists to

explore phenomenon from multiple angles, developing a multi-dimensional picture of reality. Therefore, by considering interpretivist and positivist data collectively, post-positivists can produce findings that have greater breadth and depth than the interpretivist or positivist approach in isolation (Hesse-Biber, 2016). This can help to reduce potential bias of a positivist-only or interpretivist-only approach (Morgan, 2007). For this reason, the literature review informing this study included several articles that explore employee coaching from an interpretivist perspective (Barry et al., 2021; Echeverri, 2020; McCarthy & Milner, 2020), and as the preceding sections testify, each has shaped the development of this study.

3.1 Design

A cross-sectional design was chosen for this study for three main reasons. Firstly, it aligns with the researcher's post-positivist worldview (Davies & Fisher, 2018; Denzin, 2012; Stokes, 2011). Secondly, it serves the aim of this study (Bell et al., 2019; Ghauri, Grønhaug, & Strange, 2020), and finally, it was the pragmatic choice given this study is a non-funded university research project (Bell et al., 2019). As conveyed in previous sections, the research choices for this project are post-positivist in nature: the theoretical approach is deductive; the research questions are focused; hypotheses are posited to test the hypothetical model; a quantitative strategy is adopted; and there is intent is to generalise to the wider NZ Police workforce (the population). A cross-sectional design supports these choices because it enables the researcher to measure multiple variables across a large sample of participants (Bell et al., 2019). Most importantly, a cross-sectional design enables the researcher to confirm or reject the hypotheses and answer the research questions. This is because it incorporates the statistical analysis of data, including the identification of correlations between variables, effect size, direction, and statistical significance (Bell et al., 2019; Ghauri et al., 2020). In short, the cross-sectional design is a good fit for this study.

Despite the popularity of the cross-sectional design, it has some weak points (like all designs) (Bell et al., 2019). One of the main disadvantages is that researchers sometimes struggle to make a case that an antecedent variable (X) 'causes' a consequential variable (Y). This is because when data is collected at a single point in time, one cannot say absolutely that X causes Y or Y causes X. Indeed, whilst an 'association' between X and Y can be observed, this is not evidence of causation (Hayes, 2017). Another disadvantage of the cross-sectional design is that the researcher cannot claim that X causes Y through experimentation

because there is no random assignment of participants to control and experimental groups, and no manipulation of the antecedent variable (Bell et al., 2019).

Notwithstanding these weak points, it is possible to make causal claims when using a cross sectional design. Indeed, Hayes (2017) objects to those who say it cannot be done, arguing that causation has little to do with study design and data analysis. Rather, causation ultimately rests on the strength of the argument proposed by the researcher (Hayes, 2017). It is a product of the researcher's mind, evidenced by a combination of strong theory (logic), study design (cross-sectional or otherwise), and data analysis combined (Hayes, 2017). Moreover, whilst features of experimental design such as random assignment to control/experimental groups and the manipulation of X can provide a strong argument in favour of causality, experimental designs also have limits. For example, in a mediation model, whilst the researcher can manipulate the antecedent variable (X), demonstrating that it causes an effect on the mediator variable (M), it is not possible to do the same for the mediator variable (M) on the consequential variable (Y). This means the researcher is unable to claim that M causes Y and not the other way around (Hayes, 2017).

In the case of the current study, a critic may be able to change the causal flow of the conceptual model of employee coaching by proposing something different. For example, a critic they may find statistical evidence that PQECR influences coaching frequency, and not the other way around. If so, it would confirm that data analysis (on its own at least) is insufficient to make a case for causation (Hayes, 2017). Rather, the argument for causation in this study (that coaching frequency precedes PQECR) is based on the tripod of causation (Hayes, 2017). That is, the combination of a strong theoretical argument (LMX theory), study design (cross sectional), and data analysis (conditional process analysis).

A second weakness of the cross-sectional design is the use of self-completion-surveys as the sole method of data collection (Ali et al., 2018; Kunst et al., 2018; Weer et al., 2016). The problem does not lie with the method itself because all data collection methods are subject to some degree of bias (Bell et al., 2019). Rather, the problem is the absence of variety in the way data is collected to balance against common method variance bias (CMV bias) (Bell et al., 2019). Ideally, data collection for this study would have been supplemented by structured observation; however, due to limited time resources, this was not practicable. Interestingly, within the scope of the literature review, on-line self-completion-surveys (similar to the one used in this study) were used as the sole method of data collection in the majority of the studies. Whilst the majority of authors acknowledged the increased risk of

CMV bias, just one author used a supplementary method. This tends to suggest that the use of online self-completion-surveys in isolation is accepted by the scholarship provided the researcher acknowledges the CMV bias risk.

3.2 Sampling

The population for this study is the NZ Police workforce: a national organisation comprising circa 14,500 employees, spread across 12 geographic districts (New Zealand Police, 2021).

The NZ Police workforce comprises 36% females, 12.3% Māori, 6.8% Pacific Islander, and 6.0% Asian; however, the majority of employees are males of European decent (New Zealand Police, 2021). According to (Stats NZ, n.d.), the general population of New Zealand comprises just over 50% female; 16.5% Māori, 8.1% Pacific Islander, and 15.1% Asian. This suggests that in terms of gender and ethnicity, the NZ Police workforce is somewhat diverse but not to the same degree as the general population, and it means that findings would be particular to the NZ Police population (Bell et al., 2019). The targeting of distinct populations (such as NZ Police) is common in the employee coaching literature. For example, Ismail et al. (2016) targeted midlevel military officers in Malaysia's armed forces; Dahling et al. (2016) targeted pharmaceutical sales employees based in the USA; and Matsuo (2018) target engineering employees in Japan: a sample that was 96% male.

Given that this study aims to explore the effect of coaching frequency on three outcome variables, it makes sense to recruit an organisation in which employee coaching frequency can be measured with a high degree of precision (Bell et al., 2019). Whilst it is acknowledged that many organisations encourage employee coaching (Lawrence, 2017), it is less clear how many operationalise employee coaching as a formal, regular, one-on-one leadership practice. Indeed, selecting an organisation that purportedly encourages employee coaching yet lacks a formal system to identify how and when the practice takes place would likely result in measurement error and make a nonsense of the results. This is why NZ Police was chosen for this study: the organisation has a formal performance management system (PHPF) that prescribes formal one-on-one coaching conversations between all supervisors and their direct reports at least once per month (New Zealand Police, 2017).

3.2.1 Data collection and challenges

To collect data to answer the research questions, a research proposal was presented to the NZ Police on the 19th of January 2022. Among other recommendations, the rationale for a short (five minute) organisation-wide survey was made manifestly clear to NZ Police from

the outset. Following a robust assessment process, including several requests for additional information and the signing of a comprehensive research agreement by the researcher and three academic staff from Massey University (see Appendix F), the project was formally approved by NZ Police on the 27th of April 2022 (Appendix A), three months after submission.

The rationale for requesting an organisation-wide survey for this project was threefold. Firstly, NZ Police make clear that PHPF applies to every employee (New Zealand Police, 2017), so by inviting the entire workforce to participate, the study could capture a reasonable number of participants at all hierarchical levels. Secondly, it was expected that a significant proportion of employees would chose not to participate, but by including the entire workforce in the sampling frame, the probability of obtaining a sample size large enough to reasonably represent the population from which it was drawn would increase (Bell et al., 2019). Thirdly, the larger the sample size, the more likely the results will be statistically significant (Field, 2013; Ghauri et al., 2020; Lakens, 2017). This means having greater confidence that any effect that is identified from the sample data is precise and genuine in the population rather than a random coincidence (a Type I error) (Field, 2020, October 1), something that those adopting a quantitative strategy strive for (Appelbaum et al., 2018). Finally, as the size of a sample increases, so does the power of the inferential tests applied to the data (Field, 2013), and this decreases the probability of failing to identify an effect that genuinely exists in the population (a Type II error) (Bell et al., 2019; Field, 2013). The main point here is that when it comes to identifying effects, precision tends to plateaux once a sample size exceeds 1,000 participants (Bell et al., 2019), but as a sample recedes from N=1,000, it withers in terms of statistical power.

Whilst supporting the project generally, on the 21st of June 2022, two months following formal approval, NZ Police declined the researcher's request to distribute the project survey organisation-wide. The basis for this decision was two-fold. Firstly, NZ Police did not rank the project as a strategic priority. Secondly, the organisation had other surveys lined-up for distribution and considered the risk of survey fatigue to be too great. In fairness to NZ Police, the organisation receives a high number of survey requests (internal and external), so the risk of survey fatigue is legitimate. For this reason, the organisation maintains a schedule of surveys, ranking them in accordance with its priorities. That said, NZ Police also agreed to partner with the researcher to complete the project and as per clause 2.3 of the NZ Police

Research Agreement (Annex F), the organisation agreed to make available resources and help as reasonably necessary to complete the project.

As an alternative data collection strategy, NZ Police recommended that the researcher work with the NZ Police Research Liaison Officer and distribute the project survey to a smaller more targeted group of employees (as opposed to the entire organisation). Following this advice, the researcher approached a District Commander in Tāmaki Makaurau Auckland, seeking permission to distribute the survey to circa 3,000 police employees based in the upper North Island (a smaller targeted group, as recommended). During the initial face-to-face meeting, the District Commander expressed support for the research project; however, four weeks later, and after several chase-up emails from the researcher, the same District Commander declined the request, citing the same reasons: the risk of survey fatigue, and other project surveys taking priority.

With options diminishing, alternative data-collection strategies were explored. After further consultation with the NZ Police Research Liaison Officer, it was agreed that an attempt would be made to distribute the survey by leveraging internal networks. This was a labour-intensive process, requiring the researcher and the Police Research Liaison Officer to engage in numerous conversations with workgroup managers and to re-present the project several times over. That said, it proved to be the only successful method for collecting data. As noted by Harvey (2010), organisational gatekeepers play a crucial role in protecting the interest of their company and employees, but they can also free-up research opportunities, a phenomena that is acknowledged in the literature (Harvey, 2010). In the case of this project, several senior organisational gatekeepers put barriers in front of the researcher (i.e. declining survey distribution); however, the Police liaison Officer was extremely supportive and instrumental in navigating a way through (i.e. leveraging internal networks).

Despite being the only successful method for data collection, one of the risks associated with collecting data via internal networks is that the sample became less random and more convenient in nature, meaning an increased risk of sampling bias (Bell et al., 2019; Ghauri et al., 2020). Indeed, one of the internal networks utilised to collect data for this study was a Womans' Advisory Network. Notwithstanding the outstanding work this network does advocating for women in police, membership is biased toward those who identify as female. Another internal network targeted was a group of twelve employees whose full-time role is to promote work-place learning and development across the organisation. Whilst this group of individuals deserve a voice, it is not unreasonable to suggest that their perspective and

experience of work-place coaching is unlikely to be the same as the average police officer operating on the front-line.

Between 1 September and 11 November 2022, nearly ten months after applying to do research with NZ Police, 150 NZ Police employees had completed the survey for this project. It is impossible to calculate a response rate because the survey was distributed in a snowball fashion through internal networks. That said, given the organisation employs circa 14,500 people, the sample size roughly equates to just over 1% of the NZ Police population.

The challenges experienced in collecting data for this project illustrate the difference between theorising about data collection in the classroom and the practical realities in the field. Within the scope of the literature review, the quantitative sample sizes ranged between N=117 (Zuñiga-Collazos et al., 2020) → N=605,387 (Ye et al., 2016). Of the articles comprising the literature review, at least eight had sample sizes exceeding 500 participants, and three had sample sizes exceeding 1,000 participants. Therefore, whilst not unprecedented, the sample size for this project, N=150, is comparatively modest.

3.3 Ethical considerations

Preventing harm to participants, organisations, and the wider community is a priority for this research project and is encouraged in the literature (Bell et al., 2019; Ghauri et al., 2020; Massey University, 2017). The researcher is aware that NZ Police may act on the findings and recommendations contained in this thesis, so to give the organisation confidence to do so, the study has been conducted in alignment with Massey University's code of ethics for conducting research involving human participants (Massey University, 2017), and the guidelines for Māori research ethics, Te Ara Tika (Hudson, 2010). Also, as recommended by Aguinis et al. (2018), every reasonable effort has been made to be open and transparent about how the research has been done, and to highlight the risks of acting on the findings. Indeed, prior to any data being collected, the study was subjected to a full ethical review by the Massey University Human Ethics Southern A Committee and was granted approval on the 18th of February 2022 (Appendix C).

During the design stage of this study, several ethical issues came to the surface. The first was that the researcher is also an employee of the NZ Police, holding the level of position of Senior Sergeant. As such, consideration was given to the possibility that participants may feel obliged to participate because of the organisational authority held by the researcher. Following discussion with an academic supervisor, the researcher agreed that the risk was

low given that all of the participants are adults, and it would be inaccurate to describe the NZ Police workforce as a vulnerable community. None-the-less, it was initially decided that the researcher should identify to participants as a post-graduate student only. Interestingly, when the research proposal was put to NZ Police, it was recommended that the researcher identify to participants as a police employee to give the project more credibility! Indeed, advice from NZ Police was to identify as a police employee who was conducting research as a post-graduate student, and this advice was taken. It is possible that disclosing this 'insider-status' was advantageous because it created a shared belonging with the participants (Harvey, 2010). That said, to further reduce the possibility of involuntary participation, the researcher made clear in the Participant Information Sheet and the survey introduction that participation was entirely voluntary.

The second ethical issue to surface centred on respecting participant anonymity (manaakitanga). Given that participants were advised that the results and findings of this study are to be shared widely, it was anticipated that some may be reluctant to frame responses accurately due to privacy concerns and/or fear of being sanctioned if their identity were revealed. To reduce this risk, participants were not asked for their names, employee numbers, or any other particulars that would readily reveal their identity, and this approach is supported in the literature (Kara, 2020). Moreover, in the Participant Information Sheet and within the survey itself, participants were assured that their identity would be kept anonymous.

Staying with participant anonymity a bit longer, the initial intent was to collect demographic information about gender, the geolocation in which each participant works (District/Service Centre), and their exact level-of-position (rank). Whilst this information is not directly related to employee coaching and it may seem unusual to request it, the rational for doing so was two-fold. Firstly, these demographic details enable readers to assess for themselves the degree to which the data sample reflects the target population (Appelbaum et al., 2018; Bell et al., 2019). Secondly, by including these variables in the regression analysis (and holding them constant), any suggestion that the observed effects are a result of their confounding influence can be addressed (Hayes, 2017).

Despite having valid reasons for collecting such information, it soon became apparent that the data could be filtered, revealing the identity of individual participants, particularly those in senior ranks who identify as female or gender diverse. This is because the number of participants meeting these criteria is small. Given the importance of anonymity, it was

decided to remove the question of gender entirely, and the question on level-of-position was modified so that those in positions of Inspector and above (including non-sworn Police Employee equivalents) were aggregated into a single category of Senior Police Manager. By doing this, the number of potential filters was reduced, and the risk of identifying any particular individual was minimised to an acceptable level.

The third ethical issue to surface was gender identity itself and whether it should be excluded from the study. Like geolocation and level-of-position, gender is a useful demographic because it enables the assessment of sample-population alignment and can operate as a control variable. Moreover, in the context of the literature review, gender identity was collected in three quarters of the studies, so the practice appears to be commonplace. However, when the survey for this study was pilot-tested with academic colleagues from Massey University, feedback was mixed. Some agreed that the question of gender was important to ask to compare the experiences of those who identify as different genders; others recommended eliminating gender from the questionnaire entirely, suggesting it would enable individuals to be identified. Ultimately, it was decided to remove the question of gender identity from the study due to the anonymity risk; however, the experience illuminates that fact that research decisions are not always black-and-white, there are risks and benefits to weigh-up, and tensions are abound.

The fourth ethical issue given serious consideration was whether or not to collect data on participant ethnicity. An obvious advantage is the ability to determine whether the experience of employee coaching differs depending on the ethnic group or groups one identified with. Also, it would enable readers to make a comparison as to whether the data sample accurately reflects the NZ Police population (in terms of ethnicity). That said, anonymity risks were identified. For example, a high-ranking employee who also identifies as Tongan may be easily identifiable given that just 6.8% of the NZ Police identify as a Pacific Islander (New Zealand Police, 2021).

During early stages of the project, including a Māori perspective was prioritised. To achieve this, the researcher contacted the Māori Responsiveness Manager (MRM) for the Waitemata District. The MRM (who identifies as Māori) took several days considering the potential impact of the project for Māori generally and for potential participants who might identify as Māori. During a second discussion, it was agreed that the primary focus of the project was employee coaching, not ethnic identity, and because the topic is not uniquely special from a Māori tikanga cultural perspective, it was not deemed necessary to make any cultural

allowances. Like gender identity, ethnic identity is not the focus of this project, so a decision was made not to collect this data because it was the right thing to do (tika).

The final ethical issue to mention relates to survey distribution. For quantitative investigations, sample size matters. Indeed, it has been demonstrated that when sample sizes reach 1,000 participants and over, the precision of results plateaus (Bell et al., 2019); however, as sample sizes trend downward, below 1,000 participants, statistical power increasingly wanes, confidence intervals increase, and the risk of committing Type I and Type II errors grow (Bell et al., 2019). That said, the potential for over-sampling must also be considered. NZ Police is a large organisation funded by the taxpayer, and diverting the attention of employees away from their work costs the organisation in time and productivity (Bell et al., 2019). This created a research dilemma: aim for a high response rate and risk oversampling, or avoid oversampling and produce imprecise results? This decision was further complicated by the fact that the response rate to a voluntary survey can never be known ahead of time.

Ultimately, the desire to achieve the largest possible response rate won-out, and a request for organisation-wide-email was made to the NZ Police's research gatekeeper, the Evidence Based Policing Centre. The rationale for this was that it would be easier for willing participants to be part of the project. Indeed, recipients of the email would simply have to click on a single link to be navigated to the survey. Whilst it was anticipated that many police employees would decline the invitation, it was expected that the number agreeing to participate would exceed the 1,000-participant-mark.

As discussed in greater depth in Section 3.2.1, a little over a month later, the request was declined by NZ Police. As an alternative strategy, the organisation recommended that the researcher survey a smaller more targeted audience. The rationale for this was the competing demands for organisation-wide survey and concern about subjecting the organisation's employees to survey fatigue.

3.4 Materials

To collect data from participants, an online self-completion-survey was created using the Qualtrics^{XM} software application (Appendix E). The survey contained two parts: part one measured the five main variables of interest and part two measured demographic variables. Both coaching frequency and tenure are manifest variables and were measured directly. The remaining three variables of interest are latent, meaning they could not be measured directly.

As recommended by Bell et al. (2019, p. 169), when measuring latent variables it is necessary to do so indirectly by measuring 'indicators' that represent the variable rather than the variable itself. To this end, psychometric scales were incorporated into the survey to measure the three latent variables. The scale used to measure the perceived quality of the employee coaching relationship was the PQECR; employee engagement was measured using the EES; and intention to turnover was measured using the TISCALE. These three psychometric scales were easily combined into the survey questionnaire as each contain items measured on a five-point Likert scale (see below for a thorough discussion regarding the validity and reliability of the psychometric scales).

Within the literature, psychometric scales are a popular choice for measuring latent variables. According to (Lawrence, 2017), the most popular scales for measuring employee coaching are the Coaching Behaviour Measure (Ellinger et al., 2003), and the Measurement Model of Coaching Skills (Park et al., 2008); however, these scales did not dominate the literature review. The Coaching Behaviour Measure (Ellinger et al., 2003) was used on just two occasions, and the Measurement Model of Coaching Skills (Park et al., 2008) was used on four occasions. In addition, authors also used the Behavioural Observation Scale (BOS) on four occasions, the Survey of Management Practices (SMP) on one occasion, scales adopted from the literature on four occasions, and proprietary scales on two occasions. This suggests that the scholarship is open to a variety materials for measuring employee coaching.

Table 1: Variable measures

Variable	Measurement method	Reliability of method (Cronbach's coefficient alpha)	
		Current study	Previous study
1. Coaching frequency	Survey question (single item)	-	-
2. Tenure	Survey question (single item)	-	-
3. Perceived quality of the employee coaching relationship	Psychometric scale: PQECR (12 items)	0.97	0.95[a]
4. Employee engagement	Psychometric scale: EES (12 items)	0.90	0.91[b]
5. Intention to turnover	Psychometric scale: TISCALE (6 items)	0.85	0.84[c]

a. Gregory and Levy (2010)
b. Shuck et al. (2017)
c. Dwivedi (2015)

3.4.1 Measuring coaching frequency

Coaching interactions take a variety of forms. For example, in their study of managerial coaching activity among tram drivers in Holland, Echeverri (2020) describe how company appointed coaches pose as passengers and observe a tram driver's behaviour for an hour or more before identifying themselves to the coachee and initiating a one-on-one coaching conversation. By the same token, Grant (2017, p. 9), proposes a "quality conversation framework" which highlights several alternative coaching interactions:

- Quick, collaborative conversations (less than 5 minutes in duration)
- Corridor-conversations (3-5 minutes in duration)
- Goal-oriented informal coaching (10-15 minutes in duration)
- Formal coaching sessions (30+ minutes in duration)

Given that employee coaching is the topic of this study, quantifying every coaching interaction experienced by the participants would be ideal; however, achieving this is problematic for at least two reasons. Firstly, in some coaching contexts, such as the aforementioned corridor conversations, it is possible that many coachees may not be consciously aware they are being coached (Dixey, 2015; Lawrence, 2017). Secondly, because the data for this study was collected using an online survey questionnaire, it made sense to seek information that participants could recall easily and accurately (Bell et al., 2019).

For the purpose of this study, data relating to coaching frequency was collected as part of a self-completion survey instrument. A single item asked participants the following question: "In the past six months, how many one-on-one PHPF meetings (coaching conversations) has your direct supervisor had with you to discuss/improve your performance on the job (if nine, enter 0)?" The decision to exclude informal coaching interactions from the data was deliberate given the reduced risk of data collection error due to imprecise memory and/or lack of awareness of participants (Bell et al., 2019). Moreover, by measuring formal coaching interactions exclusively, this study aligns with the NZ Police preference for monthly formal one-on-one conversations between supervisors and employees as part of the PHPF (New Zealand Police, 2017).

3.4.2 Measuring tenure

Given the above definition, there is no requirement to develop indicators for tenure because like age, place of birth, and renumeration, it is a variable that can be measured more or less

directly (Bell et al., 2019). Indeed, data confirming an employee's length of service (tenure) can be gathered from organisational records, or by asking the employee directly. For the purpose of this study, data relating to tenure was collected as part of a self-completion survey instrument. A single item asked participants the following question: "How many years have you worked for the New Zealand Police (enter 0 if less than 12 months)?"

3.4.3 Measuring PQECR

According to (Lawrence, 2017), the most frequently used psychometric scales for measuring employee coaching practice are the Coaching Behaviour Measure (Ellinger, Ellinger, & Keller, 2003), and the Measurement Model of Coaching Skills (Park, Yang, & McLean, 2008). At first glance, it makes sense to go with an existing scale that the scholarship has already accepted; however, for this study, a considered decision was made to go with the PQECR, developed and validated by Gregory and Levy (2010). The justification for using the PQECR scale over one of the more popular coaching scales is because it was developed for the specific purpose of measuring the coachee's perceived quality of the employee coaching relationship, and this serves the study aim. In contrast, the two aforementioned popular scales seek to measure either the coach's behaviour (Ellinger et al., 2003) or the coach's skill (Park et al., 2008), not the coachee's perceived quality of the coaching relationship. Therefore the PQECR is a better fit for the current study. Moreover, despite their popularity, both the Coaching Behaviour Measure and the Measurement Model of Coaching have been found to have questionable validity (Hagen & Peterson, 2014). The Military Coaching Behaviour Scale (MCBS) developed by Wagstaff, Arthur and Hardy (2018) was also considered; however, this scale was deemed unsuitable because the study samples were all military personal, overwhelmingly male, and confirmatory factor analysis benchmarks were not achieved (Wagstaff et al., 2018).

The PQECR is a 12-item scale grounded in the extant workplace coaching literature. Scale items were selected, developed, and refined by subject matter experts and statistically validated over two independent studies. Responses to the PQECR's 12 items are scored on a five-point Likert scale (1 = strongly disagree to 5 = strongly agree). A sample of items includes the following: My supervisor and I have mutual respect for each other, and My supervisor helps me to identify and build upon my strengths. When assessing the PQECR for this study, the advice from Hagen and Peterson (2014) was followed in that the full article reporting the development of the PQECR was obtained and read several times, and an appreciation for how scale items were selected and refined to meet the purpose was gained.

Hagen and Peterson (2014) are clear that three overarching standards for psychometric scales are reliability, validity, and efficacy. Content (face) validity of the PQECR began with initial item development by 25 subject matter experts (graduate students of industrial psychology) and the creation of the first iteration of the PQECR. Following this, data was collected from a convenience sample of 158 workers based in the USA, working in a variety of industries at different levels of responsibility. Analysis of the data revealed that the model fit of the first iteration of the PQECR was acceptable; however, some of the item loadings were considered too weak and were removed, and fit indices of the adjusted model indicated an improved fit. Data was then collected from a second convenience sample of 558 workers based in the USA. Model fit indices for the second iteration of the PQECR were acceptable: CFI was found to be 0.96, above the benchmark of $\geq$0.95 (Bagozzi & Yi, 1988; Hu & Bentler, 1999); RMSEA was found to be 0.10, exceeding the benchmark of $\leq$0.5 (Bagozzi & Yi, 1988; Brown, 2015; Hu & Bentler, 1999); and SRMR was found to be 0.04, exceeding the benchmark of $\leq$0.5(Bagozzi & Yi, 1988; Hu & Bentler, 1999). Factor loadings were found to be high, ranging from 0.88-1.00, and exceeded the benchmark of $\geq$0.5 (Hair, Black, Babin, & Anderson, 2014). Moreover, factor loadings were found to be statistically significant to the level of $p<0.001$. The internal reliability of the final 12-item PQECR was found to be high, with the Cronbach's alpha coefficient being calculated as α=0.95.

3.4.4 Measuring employee engagement

For the purpose of this study, the variable of employee engagement has been measured using the Employee Engagement Scale (EES) developed and validated by Shuck et al. (2017). The Utrecht Work Engagement Scale (UWES) developed and validated by Schaufeli et al. (2006) was also considered for this study; however, it was deemed less suitable because it specifically targets the dimension of work activity (Ali et al., 2018; Shuck et al., 2017) whereas the EES measures a broader perspective of employee engagement in the work context, and for this reason, the EES is preferred.

The EES comprises a twelve-item scale grounded in personal engagement theory (Kahn, 1990) and the extant engagement literature. The scale items were selected, developed, and refined by subject matter experts and statistically validated over four independent studies. The EES is scored on a five-point Likert scale (1 = strongly disagree to 5 = strongly agree). A sample of items includes the following: I am really focused when I am working, and I believe in the vision and purpose of my organisation.

Validation of the EES began with an exploratory factor analysis. Using data from a sample of 283 service workers based in the USA, the initial 14 scale items were loaded onto three lower order factors and one higher order factor. The score from a Kaiser-Meyer-Olkin test (KMO) was 0.89 which exceeds the benchmark of $\geq$0.6 (Shuck et al., 2017). Moreover, a Bartlett's test for sphericity was used to confirm that the KMO result was statistically significant to the level of $p<0.001$, meaning the data sample was ready for confirmatory factor analysis. In the developer's third study, using data from a sample of 1,067 financial service workers in the USA, the developers conducted a confirmatory factor analysis. In terms of model fit indices, the χ^2 was reported as 459.89 with 51 degrees of freedom (*df*), $p<0.01$. Concerningly, the ratio of χ^2/df is 9.02, exceeding the benchmark $\leq$5.0 (Hagen & Peterson, 2014; Hayduk, 1987). That said, the CFI, and TLI were both high at 0.99, exceeding the benchmark of 0.90 (Bagozzi & Yi, 1988; Hu & Bentler, 1999).

In terms of convergent validity, the developers compared the EES with the JES (Rich et al., 2010), which also purports to measure engagement. As expected, because the EES and the JES both measure engagement-like psychological states, the correlation coefficients between the scale factors were positive, ranging from r =0.549 to 0.849. The developers say the correlation coefficients are statistically significant but do not specify to what level. In terms of discriminative validity, the developers calculated heterotrait-monotrait ratios (HTMT) of the correlations between the EES and scores from several other theoretically related scales including the ITS (Colarelli, 1984), which measures intention to turnover (quit one's job), and the JSS (Luthans, Avolio, Avey, & Norman, 2007), which measures job satisfaction. As expected, the calculations were <1.0 (a score of 1.0 would have indicated no discrimination), ranging from 0.51→0.78.

To confirm nomological (predictive) validity of the EES, the developers compared it with several theoretically related scales including the aforementioned JSS. As expected, because the constructs that that these two scales measure are theoretically related (i.e. engagement theoretically predicts job satisfaction), the correlation coefficient was found to be positive and moderate, r = 0.77. Moreover, it was found to be statistically significant to the level of $p<0.001$. Finally, the internal reliability of the EES was tested and the coefficient alpha was found to be α=0.91, demonstrating good internal reliability and exceeding the accepted benchmark of $\geq$0.5 (Hair et al., 2014).

3.4.5 Measuring intention to turnover

For the purposes of this study, the variable of intention to turnover was measured using the Turnover Intention Scale (TISCALE) developed and validated by Dwivedi (2015). The TISCALE is a six-item scale, comprising items developed by other scholars and synthesised into a single instrument. In this sense, whilst lacking a specific theoretical foundation, the TSCALE is grounded in the extant employee turnover literature. Responses to the TISCALE's six items are scored on a five-point Likert scale (1 = strongly disagree to 5 = strongly agree). A sample of items includes the following: I occasionally think about leaving this organisation, and I intend to search for a position with another employer.

Using data from a sample of 524 call centre workers based in India, the scale developer conducted an exploratory factor analysis and found that the six items comprising the TISCALE comfortably loaded onto a single factor of turnover intention. The score from a KMO test was 0.87 which exceeds the acceptable benchmark of $\geq$0.6 (Shuck et al., 2017). Moreover, the Bartlett's test for sphericity was used to confirm that the KMO result was statistically significant to the level of $p<0.05$. The single factor loadings of the six items ranged from 0.67 – 0.78, exceeding the benchmark of $\geq$0.5 (Hair et al., 2014), and the average variance extracted by the TISCALE was calculated to be 60.1%, exceeding the benchmark of $\geq$0.5 (Fornell & Larcker, 1981; Hair, Anderson, Tatham, & William, 1995).

In terms of discriminate validity, the TISCLAE was compared with a scale that measures sportsmanship. As expected, because turnover intention and sportsmanship are distantly related at best, no statistically significant correlation was found between them. To confirm nomological (predictive) validity of the TISCALE, the developer compared it with scales measuring organisational culture and organisational commitment. As expected, because these scales measure constructs that are theoretically related (intention to turnover might predict both organisational culture and organisational commitment) correlations between the TISCALE and both were both found to be negative and moderate in strength, r=-0.68, and r=-0.60. Moreover, both correlations were found to be statistically significant to the level of $p<0.05$.

The internal reliability of the TISCALE was calculated and the Cronbach's coefficient alpha was α=0.84, demonstrating good internal reliability and exceeding the benchmark of $\geq$0.5 (Hair et al., 2014).

Despite demonstrating good validation and internal reliability, the TISCALE is unlikely to be a perfect measure of intention to turnover. As pointed out by the developer, convergent validity (the correlation between two scales that purported measure the same construct) was not tested. Moreover, whilst the results of an exploratory factor analysis were acceptable, a confirmatory factor analysis was not conducted. This is an important gap because without this evidence, there is less certainty that the TISCALE can capture the significant facets of intention to turnover. Moreover, results obtained using the TISCALE may be distorted because of the scale's inability to explain all of the variance within the data. That said, no psychometric instrument is flawless (Bell et al., 2019), and on the face of it, the TISCALE is considered fit for purpose.

3.5 Data collection procedure

An on-line self-completion survey was used to collect the data for this study. The Majority of the studies included in the literature review adopt a self-completion-survey as the sole method of data collection, and this method proved ideal in this study for several reasons. Firstly, it aligns with the post-positivism tradition of measuring variables in a detached and objective fashion (Davies & Fisher, 2018). Secondly, all NZ Police employees have access to the internet and an organisational email address, so no employee was disadvantaged due to lack of internet access. Thirdly, the use the Qualtrics^{XM} software application made generating an on-line survey relatively straight forward. Finally, the data was able to be uploaded directly into SPSS (the analytical software program chosen for this project). This saved time and avoided potential data entry errors associated with the manual transfer of data from pen-and-paper responses (Bell et al., 2019).

As noted in Section 3.1, the potential for CMV bias looms large when data collection is constrained to a single method (i.e. a self-completion-survey). The rationale for adopting this very approach came down to time and budgetary constraints, and because the project has been carried out be a single researcher. Adopting supplementary methods such as participant observation and/or structured interviews would have been preferred; however, doing so would have extended the time to complete the research and necessitated the help of research assistants which would increase the cost (Bell et al., 2019). In contrast, the on-line self-completion-survey method is fast and low-cost (Bell et al., 2019).

3.6 Common method variance

It is important to highlight the potential for common method variance bias (CMV bias). In essence, CMV bias means the potential for error where the researcher uses just one method of data-collection (i.e. a survey). CMV bias is well-recognised in the scholarship, and it is acknowledged that CMV bias can impact the results (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). Addressing CMV bias is relevant to this study because a self-completion-survey was the only method used to collect the data. A popular method for testing for CMV bias is the Harman's single factor test (Pousa et al., 2020; Ribeiro et al., 2020; Zhao & Liu, 2020). This test involves an exploratory factor analysis where all study variables are loaded onto a single factor. If the analysis reveals that the single factor accounts for a majority of the total variance explained, it is proposed that CMV bias 'may' be a significant problem (Podsakoff et al., 2003). Despite the popularity of this test, not everyone supports its use as a means of checking for CMV bias. According to Podsakoff et al. (2003, p. 889), the Harman's single factor test is "insensitive" and is not recommended as a means to check or control for CMV bias. Conducting a Harman's single factor test was considered for this study; however, given the warning about its use, it was decided not to do so because the results could provide false-assurance that CMV bias is absent. Rather, in the context of this study, it is important to consider that the results may be affected by CMV bias because data was collected using a single method.

3.7 Data analysis

The purpose of analysing the data collected for this study is threefold. Firstly, by carrying out various assessments and tests on the sample data, the researcher is able to estimate the parameters (effects) of the proposed conceptual model of employee coaching and make inferences about whether these parameters are more or less likely to be present in the wider population (Field, 2013; Hayes, 2017). Secondly, statistical analysis can be used as evidence to support or reject the hypotheses and answer the research questions (Bell et al., 2019). Finally, statistical analysis enables readers of the research to make an informed judgement as to whether they support or reject the researcher's argument (Bell et al., 2019; Ghauri et al., 2020).

3.7.1 Data management

After the survey was closed off, the data was exported from the Qualtrics^{XM} software application directly into SPSS. This is a useful feature of the Qualtrics^{XM} software because

direct export saves time, and avoids data entry error associated with the manual entry of pen and paper responses (Bell et al., 2019). To safeguard the security and integrity of the data, copies of the exported file were stored electronically in four locations (in accordance with the data management plan):

- The Researcher's Personal Vault repository hosted on Microsoft's One Drive cloud storage (password protected)
- The hard drive of the Researcher's personal computer (password protected)
- A USB flash drive belonging to the Researcher (password protected)
- The Researcher's online data repository, hosted by figshare.com (password protected)

Following this, a small amount of data organising was performed on a working data set. Firstly, the scores from the multi-items used to measure the constructs of PQECR, engagement, and intention to turnover were aggregated and then averaged into a single score for each construct. Secondly, for the purpose of assessing model fit, as recommended by Hayes (2017), indicator (dummy) variables were created for two multi-categorical variables: District (geolocation), and level of position.

3.7.2 Missing data

For any number of reasons, respondents may choose not to fully complete a survey questionnaire, including not understanding a particular question or questions, exercising their right not to respond, or because they wish to withdraw from the study (Bell et al., 2019; Field, 2013). Missing data can create difficulties for statistical analysis (Field, 2013). For example, an estimate of the mean number employee coaching sessions experienced by a respondent is calculated by dividing the sum of coaching sessions by a count of participants. In such as case, the statistic would be inaccurate where one or more respondents decided not to answer the question relating to coaching frequency because, despite not providing any data, the participant(s) would be included in the count of respondents.

Several options exist for dealing with missing data (incomplete responses). These include removing the entire response, assigning a value to the missing data and programming SPSS to ignore it, or imputation (replacing missing data with estimated values) (Martin, 2020, August 1). When using the PROCESS macro for SPSS (discussed below) to generate linear regression output, the program avoids such difficulties by eliminating cases that have missing data on any variable (Hayes, 2017). Whilst this eliminates statistical difficulties,

one must be mindful that missing data may not be a random occurrence. Rather, there may be a pattern as to why certain respondents are unable or unwilling to answer one or more questions. In such a case, the removal of incomplete responses might arbitrarily eliminate a legitimate category of people and bias the results (Martin, 2020, August 1).

To eliminate the possibility of missing data for this study, the forced response option was activated in Qualtrics^{XM}. The single exception to this was the final question where a free text response was made optional. This meant respondents were unable to progress from one item to the next without first providing a response, and incomplete surveys were excluded from study. Indeed, participants were given two weeks from the time they last edited their response to complete the survey before their response was deleted. Ideally, responses from all participants (including incomplete responses) would have been included in the final analysis; however, as per the participant Information Sheet (Appendix D), participants were assured that they could withdraw from the study at any time by not completing the entire survey. Therefore, incomplete responses were not included in the final analysis.

3.7.3 Analytical strategy

The overarching analytical strategy for this study was conditional process analysis (CPA), a term created by Professor Andrew Hayes in 2013 (Hayes, 2017). CPA is a form of linear regression modelling that uses the partial-least-squares method. The term CPA is used to describe the merging of mediation and moderation analysis, both conceptually and analytically (Hayes, 2017). For the most part, CPA was chosen because it is easy to use and meets the aims of this study.

Despite sounding quite complex, CPA is relatively straightforward: sample data is plotted on a graph and a straight line (the regression line) is 'fitted' to it (Field, 2020, Sep 23). As pointed out by Field (2013), most social scientists use linear regression modelling to describe their data and the literature review undertaken for this this study corroborates this claim with the majority of the studies reviewed adopting linear regression modelling as the overarching data analysis strategy. Unsurprisingly, the data points collected for this study did not perfectly line-up with (fit) the regression line (there was some residual error). That said, by adopting the partial-least-squares method, the regression line that was fitted to the sample data was the one that contained the least amount of residual error (Field, 2013).

3.7.4 Statistical significance

In terms of accepting/rejecting the hypotheses and answering the research questions, one must always consider the level of statistical significance that is necessary because there is always a risk (be it larger or smaller) that any effect observed in a sample may not be real in the wider population (Bell et al., 2019). But what level of risk is acceptable? By tolerating a high level of risk, the researcher may commit a Type I hypothesis error. This means positing that an effect is real in the population when it is not (Ghauri et al., 2020). Alternatively, if the researcher is too risk adverse, they may commit a Type II hypothesis error (Ghauri et al., 2020). This means positing that an effect is not real in the population when it is.

Interestingly, in the context of the articles that comprised the managerial/employee coaching aspect of the literature review, just one author reported an *a priori* alpha value for statistical significance prior to data being collected and analysed (Ali et al., 2018). In this case it was $\alpha=0.05$. This means that the result of an inferential test such as a NHST or confidence interval construction tell us that the probability of an observed effect being the result of a random coincidence is no greater than one in twenty. (Ali & Aziz, 2018). An alpha value of $\alpha=0.05$ is typical for the social sciences (Bell et al., 2019; Field, 2013; Hayes, 2017).

Whilst the absence of an *a priori* alpha value does not infer that the statistical significance of observed effects is unimportant, it does infer that a pass-fail point is viewed by some as somewhat arbitrary (Field, 2020, October 1). That said, others disagree. According to Appelbaum et al. (2018), the setting of an *a priori* alpha value is best practice for those wanting to conduct high-quality quantitative research. Therefore, in the context of this study, effects observed from sample data are not considered to be statistically significant unless an inferential test confirms that the chance of the effect being a random coincident is no greater than one in twenty (avoiding a Type I error). Where this benchmark is satisfied, the null hypothesis is rejected in favour of the alternative.

3.7.5 Analytical software

Within the scope of the literature review, the choice of analytical software was broad. For example Kunst et al. (2018) use Latent Gold; Hsieh and Huang (2018); and Tanskanen et al. (2019) use Mplus; and Zuñiga-Collazos et al. (2020) use EQS. For the purpose of this study, IBM's SPSS was chosen because the researcher is familiar with SPSS and the software package features prominently in the literature (Chen et al., 2019; Ribeiro et al., 2020; Zhao & Liu, 2020). In addition, SPSS is one of the main statistical applications used by Professor

Hayes when he instructs students how to conduct CPA (Hayes, 2017). Indeed, Professor Andrew Hayes developed the macro program, PROCESS, to work within SPSS to explore and describe the mechanism(s) by which one variable transfers its effect to another and to test hypotheses about the effect(s) (Hayes, 2017), so it makes sense to use this application.

3.7.6 The PROCESS macro

After preparing the data for analysis, the PROCESS macro for SPSS (version 4.0) developed by Professor Andrew Hayes (Hayes, 2017) was used to conduct a first-stage CPA for the proposed conceptual model of employee coaching. Given that there are two consequential variables for this study (employee engagement and intention to turnover), a separate PROCESS routine was executed for each. As stated by Professor Hayes, the PROCESS macro comes into its own at this point because, in terms the statistical calculations, it does all of the 'heavy-lifting' for the researcher and produces an output report that is easy to interpret (Hayes, 2017).

3.7.7 The statistical model of employee coaching

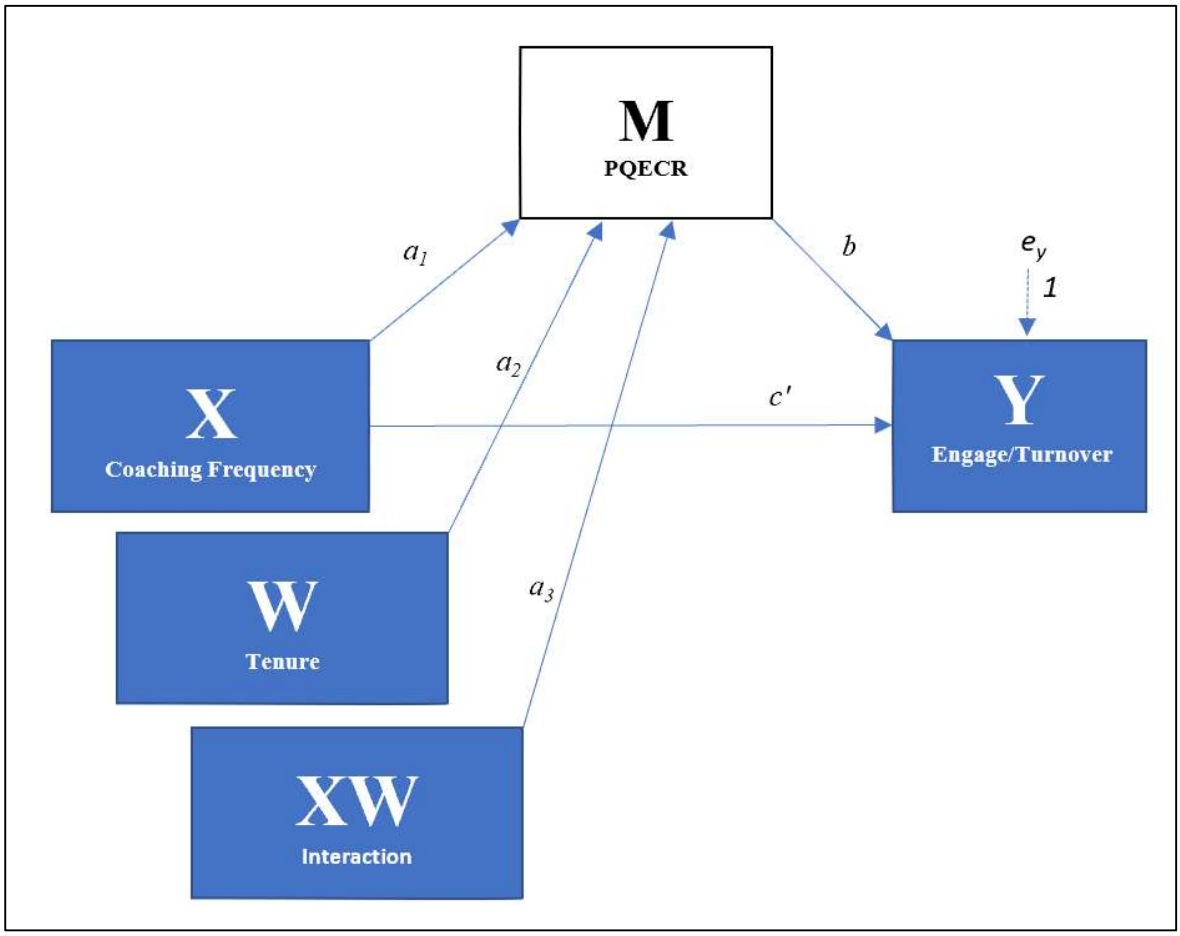

Figure 3: The statistical model of employee coaching

3.7.8 Regression equations

With reference to Figure 3, two linear regression equations were constructed to estimate the regression coefficients (effects) represented by the symbols a_1, a_2, a_3, b, and c'. In both equations, the Y-axis-intercept is represented by the i symbol; employee engagement and intention to turnover are represented by the Y symbol; coaching frequency is represented by the X symbol; PQECR is represented by the M symbol; and tenure is represented by the W symbol.

- The conditional effect of coaching frequency on PQECR, conditioned by tenure:

  $$M=i_M+(a_1+a_3W)X+a_2W+e_M$$

- The partial effects of coaching frequency and PQECR on engage/intention to turnover (unconditioned):

  $$Y=i_Y+c'X+bM+e_Y$$

Despite appearing dissimilar, both equations are mathematical representations of a linear regression. The difference being that in the first equation, the effect of the antecedent variable (X) on the consequential variable (M) is permitted to vary as a function of the moderator variable (W). In other words, the effect of X on M is a 'conditional effect', meaning it may vary as the value of the moderator (W) varies (Hayes, 2017). In the second equation, the effect of the antecedent variables X and M on the consequential variable Y are both mathematically constrained (held constant). In other words, the effect of X on Y and the effect of M on Y are both 'partial effects' and they remain constant regardless of any change to the value of the moderator (Hayes, 2017).

3.7.9 Interpreting the regression coefficients

The regression coefficient represented by the a_1 symbol estimates the conditional effect of the focal predictor, coaching frequency (X), on the mediator variable , PQECR (M), when the value of the moderator variable, Tenure (W), equals zero. In the frame of this study, this can be interpreted as the difference on PQECR for two people who differ by one unit on coaching frequency, when tenure equals zero. By the same token, the regression coefficient represented by the a_2 symbol estimates the conditional effect of the moderator (W) on the mediator variable (M) when the value focal predictor (X) is zero. This means the difference on PQECR for two people who differ by one unit on tenure, when coaching frequency equals zero. It is worth reiterating the point that the a_1 and a_2 regression coefficients estimate

"conditional" effects meaning they are subject to change in size and/or sign and should not be misinterpreted as "main" or "average" effects of X on Y, or W on Y respectively (Hayes, 2017, p. 311).

The regression coefficient represented by the a_3 symbol estimates the effect of the two-way interaction between the focal predictor (X) and the moderator variable (W) on the mediator variable (M). The XW interaction is the product of coaching frequency and tenure. The a_3 regression coefficient is an unconditional effect, meaning its size and sign are not conditioned by the value of any other variable in the model. Depending on which variable has been identified as the focal predictor and which has been identified as the moderator, the *a3* regression coefficient can be interpreted differently (Hayes, 2017); however, in alignment with Figure 3, the focal predictor is coaching frequency, so the a_3 regression coefficient can be interpreted as the systematic (constant) change to the effect of coaching frequency on PQECR when tenure increases by one unit (Hayes, 2017).

The regression coefficient represented by the *b* symbol estimates the partial effect of the moderator variable, PQECR (M), on the consequential variable (Y), holding the focal predictor, coaching frequency (X), constant. That is, the difference on engagement/intention to turnover for two people who are equal on coaching frequency and differ by one unit on PQECR. Likewise, the regression coefficient represented by the *c'* symbol estimates the partial (direct) effect of the focal predictor (X) on the consequential variable (Y), holding the mediator variable (M) constant. That is the difference on engagement/intention to turnover between two people who are equal on PQECR but differ by one unit on coaching frequency. It is worth noting that neither the *b* nor *c'* regression coefficients are moderated by another variable in the model, meaning their effects are constant.

According to Professor Hayes, depending on the boundaries of one's measurements of the focal predictor and the moderator variables, the regression coefficient represented by a_1 and a_2 may not be meaningfully interpretable (Hayes, 2017). For example, if one measures the moderator variable using a Likert scale ranging in increments from 1-5, it is a nonsense to interpret the conditional effect of X on Y when the moderator equals zero because the range of measurement does not permit a score of zero.

This advice was considered; however, in the context of this study, it was deemed unnecessary to make any adjustment because the variables of coaching frequency and tenure permit a score of zero within the range of measurement. Moreover, the sample data contains

scores of zero for both of these variables. Therefore the regression coefficients represented by a_1 and a_2 in Figure 3 are meaningfully interpretable. If this were not the case, the problem could be resolved by mean-centring the moderator variable. By doing so, the a_1 and a_2 regression coefficients are always meaningful because the conditional effects they estimate are always within the range of the data (Hayes, 2017). That said, the interpretation would change slightly. Instead of estimating the conditional effect of the focal predictor on the consequential variable when the moderator is equal to zero, the a_1 and a_2 regression coefficients estimate the conditional effect on the consequential variable when the moderator is equal to its mean value (Hayes, 2017).

In the context of this study, the regression coefficient for any particular conditional effect of X on M conditioned by W can be estimated using the equation $f=a_1+a_3W$, where f represents the conditional effect of X on M and W is a placeholder for any particular value of the moderator one is interested in. In his teachings, Professor Andrew Hayes uses the symbol θ (theta) to represent a conditional effect and subscript to identify the statistical pathway (Hayes, 2017). For example, $\theta_{X\rightarrow M}$ represents the conditioned direct effect of X on M; and $\theta_{X\rightarrow M\rightarrow Y}$ represents the conditional indirect effect of X on Y, mediated by M (Hayes, 2017). The same convention has been adopted in this study.

3.7.10 Testing for moderation

It is important to appreciate the difference between a conditional effect and the systematic (constant) effect of moderation (interaction). A conditional effect is the effect of the focal predictor on the consequential variable at a particular point-value of the moderator, and as the point-value of the moderator changes, so does the conditional effect (Hayes, 2017). Therefore, conditional effects are variable. Moderation, on the other hand, quantifies the systematic variation to the effect of the focal predictor on the consequential variable as the moderator variable increases by one unit (Hayes, 2017), so unless the primary moderator is itself moderated by a secondary moderator (which is not the case in this study), the moderating effect is constant (Hayes, 2017).

In the current study, it is hypothesised that the effect of coaching frequency on PQECR is negatively moderated by tenure. That is, the positive effect of coaching frequency on PQECR will systematically reduce as employee tenure extends. But how can this be determined? One mistake to avoid is to start by probing an interaction for two or more

conditional effects and claim this as evidence of moderation. This is because probing an interaction is not the same as testing whether it is statistically significant (Hayes, 2017).

Hayes recommends that one first conducts an inferential test for the interaction effect before going on to probe it for regions of statistical significance (Hayes, 2017). The reason for this is that the data used to estimate the effect of the interaction is sample specific and will almost certainly contain sampling error (Bell et al., 2019). Indeed, data from a different set of cases is unlikely to produce identical regression coefficients (Field, 2013). Therefore, if the result of an inferential test infers that the interaction is statistically significant, one can go on to probe the interaction; however, if the interaction is not statistically significant, it should be removed from the regression model entirely (Hayes, 2017).

An acceptable method for testing a hypothesis about moderation is to include the moderator variable and its interaction with focal predictor in a model that estimates the consequential variable (Hayes, 2017). Provided the inferential test for the interaction confirms that the regression coefficient (a_3 in the case of this study) is statistically significant, one can reasonably claim that the effect of the focal predictor on the consequential variable varies as a function of the moderator (Hayes, 2017). Following this advice, the variable of tenure and the product of tenure and coaching frequency (the XW interaction) were included as predictor variables in a regression estimating the effect on PQECR. The results of the inferential test for the XW interaction are reported in Chapter 4.

In addition to conducting an inferential test for the effect of the interaction on the consequential variable, one may also test a hypothesis of moderation by comparing the fit of two models, the first of which omits the XW interaction and the second of which includes it (Hayes, 2017). Provided the change to R^2 is found to be statistically significant, one can reasonably claim that the effect of the focal predictor on the consequential variable varies as a function of the moderator (Hayes, 2017). Following this advice, the results of an *F*-test (inferential test) for the change to R^2 between a model that omits the interaction and a second that includes it are reported in Chapter 4.

3.7.11 Probing the interaction

Once sufficient evidence of moderation has been established, one can go on to probe the interaction to gain a sense of 'when' (if at all), across the distribution of the moderator, estimations of the conditioned effect of the focal predictor on the consequential variable are statistically significant from zero (Hayes, 2017). Depending on the precision one requires,

it is possible that the conditioned effect will be statistically significant for the entire distribution of the moderator, parts of it, or none at all (Hayes, 2017).

There are at least two options for probing an interaction. The first is the pick-a-point approach (also known as a spotlight analysis), a process that involves probing the conditional effect of X on Y at one or more point-values of W (Bauer & Curran, 2005; Hayes, 2017; Rogosa, 1980). The second is the Johnson-Neyman technique (also known a floodlight analysis) (Spiller, Fitzsimons, Lynch Jr, & McClelland, 2013), a process that involves probing the conditional effect of X on Y at numerous point-values of W.

In the context of this study, the pick-a-point approach was used to estimate the conditional effect of coaching frequency on PQECR at the 16^{th}, 50^{th}, and 84^{th} percentiles of the tenure distribution. It is acknowledged that these point-values are arbitrary; however, they roughly represent relatively low, relatively moderate, and relatively high points along the distribution of the moderator (Hayes, 2017), so they provide the researcher with a sense of how the conditional effect trends across the distribution of tenure. Conveniently, when using PROCESS, the macro estimates these point-values, producing the conditional effects together with bootstrap confidence intervals (Hayes, 2017).

The Johnson-Neyman technique (Johnson & Fay, 1950) is a useful option for visualising the conditional effect of coaching frequency on PQECR across the distribution of tenure. The PROCESS macro is able to produce 22 conditional effects together with OLS confidence intervals and can be used to plot the conditional effect on a graph for a visual representation. One of the advantages of the Johnson-Neyman technique over the pick-a-point approach is that it identifies the precise areas on the distribution of the moderator where the conditional effect is statistically significant to the required level (Hayes, 2017). Moreover, it can be used to estimate the percentages of the distribution of the moderator where the conditional effect is statistically significant versus those that are not (Hayes, 2017).

Considering the conditional effects described above, how do we know whether each conditional effect is statistically different from the others? Auspiciously, for moderated mediation where there is a single moderator, the inferential test for moderation doubles as a test for the difference between different each of the conditional effects (Hayes, 2017). Simply put, when the XW interaction is found to be statistically significant in the regression model, we know that any two conditional effects are statistically significant from each other (Hayes, 2017).

3.7.12 The conditional indirect effect

In a moderated mediation analysis, the conditional indirect effect takes primacy because it explains the 'when of the how' (Hayes, 2017). In the context of the current study, it explains how coaching frequency indirectly influences employee engagement and intention to turnover through PQECR, and when, on the distribution of tenure, the effect is statistically significant.

The conditional indirect effect ($\theta_{X \rightarrow M \rightarrow Y}$), is a product of two regression coefficients. The first is the conditional effect of the focal predictor on the mediator variable (X on M) and the second is the partial effect of the mediator variable on the consequential variable (M on Y) (Hayes, 2017).

Within the constraints of the statistical model of employee coaching, the conditional indirect effect is the product of the conditional effect of coaching frequency on PQECR by tenure (a_1+a_3W) and the effect of PQECR on engagement/intention to turnover (b). Any particular conditional indirect effect can be calculated using the equation $(a_1+a_3W)b$, where W is a place holder for any particular value of the moderator (tenure) one is interested in. Indeed, this equation was used to estimate the indirect effect of coaching frequency on both engagement and intention to turnover, at the 16th, 50th, and 84th percentiles of the tenure distribution. The results of the estimation and their statistical significance are reported in Chapter 4.

3.7.13 The index of moderated mediation

It is important to appreciate the difference between a conditional indirect effect and the "index of moderated mediation" (Hayes, 2017, p. 425). The index of moderated mediation (Index) is a regression coefficient that estimates the systematic change to the conditional indirect effect ($\theta_{X \rightarrow M \rightarrow Y}$) as the moderator increases by one unit. It can be estimated using the equation a_3b where the a_3 regression coefficient quantifies the systematic change to the conditional effect of the focal predictor (X) on the mediator variable (M) and the moderator variable (W) increases by one unit; and the b regression coefficient quantifies the unconditional effect of the mediator variable (M) on the consequential variable (Y). Within the framework of the statistical model of employee coaching, the Index is a constant effect and should not be confused with a conditional indirect effect (an indirect effect conditioned by a particular point-value of the moderator).

The Index is an important statistic when fitting a moderated mediation model because one cannot argue that an indirect effect is moderated unless the index of moderated mediation has been estimated to be statistically significant to the *a priori* alpha level (Hayes, 2017). Conveniently, the PROCESS macro performs the necessary calculations and generates a bootstrap confidence interval to assist in the assessment of statistical significance of the Index. Results of the inferential test for the Index are reported in Chapter 4.

Unlike the Sobel test (a specialised t-test) (Sobel, 1982), bootstrap sampling (bootstrapping) is an alternative method for assessing the statistical significance of the indirect effect (or any effect for that matter). Bootstrapping takes the original sample data and resamples it many times (using replacement cases) to produce a hypothetical sampling distribution which is statistically powerful (Hayes, 2017). Indeed, as recommended by Professor Hayes, 5,000 bootstrap samples were used in the context of this study (Hayes, 2017). Bootstrap sampling is used to construct bootstrap confidence intervals which are used to assess statistical significance. So long as the upper and lower bootstrap confidence intervals do not cross zero, the indirect effect observed in the data sample can be considered to be statistically significant to a predetermined alpha level (Hayes, 2017). It is important to note that even when bootstrap confidence intervals do not include zero, one cannot say that the probability of the null is $p \leq 0.05$ because bootstrapping is not a NHST and therefore does not produce a test-statistic and associated p-value (Hayes, 2017).

A reasonable question to ask is why chose the bootstrapping method over the Sobel test? Professor Hayes recommends avoiding the Sobel test for two main reasons. Firstly, the Sobel test assumes that the data in the sample distribution is distributed normally, and this is not the case for the indirect effects because they contain the product of two regression coefficients (a_1b) (Hayes, 2017). Secondly, the Sobel test has low statistical power, producing confidence intervals that are less precise than other methods (Hayes, 2017). Professor Hayes recommends using bootstrap confidence intervals when estimating the statistical significance of the indirect effect because this method does not rely on the sample data being normally distributed and bootstrapping has higher power than the Sobel test (Hayes, 2017).

Another reasonable question to ask is whether the conditional indirect effect at every point-value of the moderator is statistically different from the other points of the moderator? Conveniently, the bootstrapping technique used to estimate confidence intervals for Index doubles as an assessment of this very thing (Aiken, West, & Reno, 1991; Hayes, 2017). Put

simply, when the 95% bootstrap confidence interval for the Index omits zero, one can be assured that the conditional indirect effect at any particular point-value of the moderator is statistically significant from the conditional indirect effect at every other point-value of W (Hayes, 2017).

3.7.14 The direct effect

When the focus is on moderated mediation, direct effects do not particularly matter; however, they may be of interest because they represent the part of the hypothesised model that the indirect effect is unable to explain (Hayes, 2017). As discussed in more detail in Chapter 4, no statistical model is perfect; all are incorrectly specified to some degree (Field, 2013); however, some models can be more incorrect than others (Hayes, 2017). Indeed, the poorer the model fits the data it purportedly represents, the more likely something else is going on within the data that the model is failing to explain (Bell et al., 2019). In the case of this study, the direct effect is a regression coefficient that quantifies the effect of coaching frequency on employee engagement/intention to turnover whilst holding the effect of PQECR on employee engagement/intention to turnover constant. In other words, it quantifies by how much two cases that differ by one unit on coaching frequency are estimated to differ on employee engagement/intention to turnover, when PQECR is held constant. Like the conditional indirect effect, an inferential test was conducted to estimate the probability of the direct effect being different from zero, and the results of the inferential test are reported in Chapter 4.

3.7.15 The total effect

The total effect is of little interest in conditional process analysis because the sum of the indirect and direct effects and can be misleading, particularly when the sign (positive or negative) of the direct and indirect effect are not the same (Hayes, 2017). In such a case, the total effect might come close to zero and be statistically insignificant, and at the same time, the indirect and direct effects might be large and be statistically significant, albeit with opposite signs. Therefore, just because the total effect is close to zero and not statistically significant, this is not evidence of no association between the antecedent variable and the consequential variable (Hayes, 2017). For these reasons, the total effect is not reported.

3.7.16 Confounding variables and epiphenomenal associations

A confounding variable is one that has a causal effect on both the antecedent and consequent variable, giving a false impression that the antecedent variable and the consequential

variable are linearly associated when they are not (Field, 2013). An epiphenomenal association occurs when the association between an antecedent and consequential variable appears to be mediated by a third variable; however, in such a case, the true mediator is absent from the model. This mirage occurs because the excluded mediator is linearly associated with mediator that is included in the model (Hayes, 2017).

Professor Andrew Hayes is clear that output from a regression analysis (on its own at least) is incapable of determining whether an association or mediation between an antecedent and consequential variable is genuine (Hayes, 2017). This means that there is always a risk that a confounding variable and/or an epiphenomenal association is concealing the truth. Indeed, this risk is greater when the study design is cross-sectional because, unlike experimental designs, there is no random assignment, no use of control v experimental groups, and no controlled manipulation of any independent variable. This is why causal arguments must be based on more than data analysis alone (Hayes, 2017).

One way to reduce (but not eliminate) the risk of such spurious associations in a cross-sectional design is to anticipate potential confounding variables, and include them in the regression model as covariates (Hayes, 2017). In doing so, the effect of each covariate is controlled for by holding it constant, allowing the researcher to determine if the effect of interest persists (Field, 2013). Of course, it is impracticable to include every potential confounder as the list is infinite; however, one may anticipate a modest number to ward off potential critics who might suggest that such a third (or subsequent) variable (absent from the linear model) is responsible for effects that are observed (Hayes, 2017).

In the context of this study, consideration was given to the possibility that participant age, sworn/non-sworn status, full/part-time status, level of position, and district (geolocation) might be wholly or partially responsible for the effect of the main variables of interest on the three consequential variables. To test this, the aforementioned covariates were included in the initial regression analysis (indicator coding was used for the multi-categorical variables). By including these covariates in the regression, the effect of each could be controlled (Hayes, 2017). As anticipated, the magnitude of change to the effects of interest was miniscule and there was no change in sign (refer Table 1). In this situation, where potential confounders have been accounted for and found to have made little difference to the effects of interest, it is perfectly reasonable to remove the covariates from the regression model and that is what was done.

Table 2: The effect of covariates on the main variables of interest

		Outcome Variables					
		PEQER		Employee Engagement		Intention to Turnover	
1.	Coaching frequency	0.07[a]	0.09[b]	0.01[a]	0.03[b]	-0.04[a]	-0.05[b]
2.	Tenure	-0.03[a]	-0.02[b]	-	-	-	-
3.	PQECR	-	-	0.19[a]	0.16[b]	-0.50[a]	-0.45[b]

a. Regression coefficient with covariates included in the model
b. Regression coefficient with covariates removed from the model

CHAPTER 4: RESULTS

The purpose of this chapter is to report the results of the data analysis. Importantly, the hypotheses proposed in Section 2.8.1 are either accepted or rejected, and the research questions are answered.

4.1 Participant profile

Table 3: Participant profile

Characteristics	Sub profile	Sample frequency	Sample (100%)	NZ Police Population (100%)*
District (geolocation)	Northland	0	0.00	3.17
	Waitemata	5	3.33	6.36
	Auckland	4	2.67	6.73
	Counties Manukau	5	3.33	9.81
	Waikato	2	1.33	5.42
	Bay of Plenty	4	2.67	5.74
	Central	1	0.67	5.76
	Eastern	0	0.00	3.84
	Wellington	9	6.00	6.49
	Tasman	0	0.00	2.75
	Canterbury	5	3.33	7.13
	Southland	56	37.33	4.47
	RNZPC	13	8.67	2.63
	PNHQ + Service Centres	46	30.67	29.70
Level of position	Recruit	1	0.67	0.98**
	Constable + equivalent	93	62.00	73.62**
	Sergeant + equivalent	19	12.67	15.84**
	Senior Sergeant + equivalent	21	14.00	5.75**
	Senior Police Manager	16	10.67	3.81**
Age (years)	0-24	11	7.33	4.81
	25-34	36	24.00	27.77
	35-44	36	24.00	21.60
	45-54	39	26.00	26.85
	55-64	26	17.33	16.46
	65+	2	1.33	2.52
Tenure (years)	0-9	61	40.66	NA
	10-19	46	30.67	
	20-29	28	18.67	
	30-39	13	8.67	
	40-49	2	1.33	

Employment capacity	Constabulary	85	56.67	71.68
	Non-Constabulary	65	43.33	28.32
Hours contracted	Full-time	141	94.00	93.23
	Part-time	9	6.00	6.77

*. NZ Police Annual Report 2020/21 (New Zealand Police, 2021)
**. Constabulary (sworn) employees only (New Zealand Police, 2021)

4.2 Descriptive statistics and bivariate correlations

Descriptive statistics and bivariate correlations for the main variables included in this study are presented in the Table 3. The reporting of these statistics is common practice within the scholarship. For example, in their study exploring the relationship between employee coaching and performance, Ribeiro et al. (2020) produce a table that combines the Mean and SD of each of the study variables along with a bivariate correlation matrix. Readers may find these statistics useful because they help one to visualise the shape and spread of the data around the linear model, providing a useful indicator of the model's precision and predictive ability (Field, 2013). Descriptive statistics include the arithmetic mean (Mean) and standard deviation (SD) for each of the study variables; bivariate correlations report the correlation between each variable with each of the other variables in the linear model.

Table 4: Descriptive statistics and bivariate correlations

Variable	M	SD	1	2	3	4	5
4. Coaching frequency	2.36	3.64	-				
5. PQECR	3.78	1.11	0.28**	-			
6. Employee engagement	4.30	0.61	0.23**	0.34**	-		
7. Intention to turnover	2.72	1.04	-0.31**	-0.53**	-0.45**	-	
8. Tenure	13.93	10.86	0.12	-0.17*	0.14	-0.02	-

**. Correlation is statistically significant at the level of 0.01 (two-tailed)
*. Correlation is statistically significant at the level of 0.05 (two-tailed)

4.3 Data normality

Like linear regression modelling generally, CPA is based on the assumption that the sample data is normally distributed (parametric), approximating a 'bell-shaped' normal distribution curve (Hayes, 2017), so discussing data normality is considered good practice for those

undertaking quantitative research (Appelbaum et al., 2018). To appreciate why, one must understand that among other parameters, CPA estimates regression coefficients (effects), and their statistical significance (Hayes, 2017), and the accuracy of each relies on the sample data approximating a normal distribution curve (Field, 2013). For example, the estimation of regression coefficients tends to be more precise when the errors of estimation (residuals) around the regression line reflect a normal distribution curve (Field, 2013). Also, in terms of statistical significance, OLS confidence intervals tend to be more precise when the sample data itself approximates a normal distribution curve. Furthermore, NHST tends to be more precise when the hypothetical sampling distribution from which the test statistic is estimated approximates the familiar bell-shaped curve (Field, 2013).

Unsurprisingly, data collected from the real-world rarely reflect a perfect bell-shaped normal distribution curve (Bell et al., 2019), and a degree of skewness (asymmetry) and/or kurtosis (flatness/tallness) is almost certain (Bell et al., 2019). This is unhelpful when one considers that CPA operates on the assumption that the sample data is parametric. That said, sample data does not need to be perfectly normally distributed. Rather, it is enough that it reasonably approximates a normal distribution curve (Löfgren, 2013, Aug 5). Controversially, Professor Andrew Hayes argues that the assumption of data normality is one of the least important in linear regression analysis, and that "only the most severe violations of the normality assumption substantially affect the validity of statistical inferences…" (Hayes, 2017, p. 70).

4.3.1 Skewness and kurtosis

SPSS was used to calculate the degree of skewness and kurtosis of the data collected for this study (Table 4). Given that a perfectly normally distributed data set would score a zero for skewness and kurtosis (Löfgren, 2013, Aug 5), the estimated skewness and kurtosis for the variables of interest was found to be substantial.

Table 5: Skewness and kurtosis statistics

	Coaching frequency	PQECR	Employee engagement	Intention to turnover	Tenure
Skewness	3.95	-0.83	-1.01	0.11	0.66
Standard error of skewness	0.20	0.20	0.20	0.20	0.20
Kurtosis	23.20	-0.13	0.39	-0.92	-0.35
Standard error of kurtosis	0.39	0.39	0.39	0.39	0.39

4.3.2 Tests of normality

In addition to calculating skewness and kurtosis, SPSS was used to conduct two commonly used normality tests: the Kolmogorov-Smirnov test, and the Shapiro-Wilk test. According to Löfgren (2013, Aug 5); O'Loughlin (2020, November 14), the null hypothesis for the Shapiro-Wilk test is that the data is normally distributed, so when the probability of the test statistic is greater than the *a priori* alpha value (α=0.05 in the case of this study), the sample data can be considered approximately normal.

As per Table 6, the Shapiro-Wilk test statistic for each variable of interest was found to be statistically significant, meaning that the sample data cannot be considered normally distributed. Moreover, a visual inspection of the histograms, normal Q-Q plots, and box plots produced by SPSS confirmed the same.

Table 6: Tests for data normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Coaching frequency	0.30	150	<0.001	0.61	150	<0.001
PQECR	0.14	150	<0.001	0.90	150	<0.001
Employee engagement	0.13	150	<0.001	0.90	150	<0.001
Intention to turnover	0.06	150	0.200*	0.97	150	0.002
Tenure	0.12	150	<0.001	0.93	150	<0.001

*. This is a lower bound of the true significance
a. Lilliefors significance correction

A possible explanation for the non-parametric nature of the sample data is the study's relatively small sample size and the presence of outliers. For example, coaching frequency data was collected from 150 participants. The mean was estimated to be 2.36 and the 95% OLS confidence interval ranged from 1.77 to 2.95, yet nine participants scored between 9 and 30. Such extreme outliers have caused the distribution to become skewed and kurtotic. Indeed, had the sample size been significantly larger, the impact of these outliers may have been less severe (Bell et al., 2019).

It is not unusual for extreme outliers to be deleted or the scores modified. This action may be justified on the grounds that the person who supplied the data does not appear to be part of the population that the researcher intended to sample (Field, 2013), or it is clear that the person supplying the information is being disingenuous. Data "trimming" (Field, 2013, p. 196) can be done in several ways. One option is to remove an equal percentage of responses from each end of the sample; another it to remove samples that sit beyond a particular standard deviation from the mean (Field, 2013). If trimming is not preferred, one may choose to 'winsorize' the data. This involves replacing extreme outlier scores with more reasonable ones (i.e. the next highest score that is not an outlier) (Field, 2013). Whilst this may feel mischievous, the accuracy of results will ultimately improve because the loosening effect of a small number of highly irregular outliers is removed (Field, 2013).

The most extreme outlier for this study is the case of a participant scoring 30 for coaching frequency. For this to be accurate, it means that on average, the participant's supervisor facilitated slightly more than one formal one-on-one coaching conversations per week for six months. Given the average participant received a one-on-one coaching conversation with their supervisor once every 11 weeks, the outlier's situation seems highly unlikely; however, it is not impossible. Indeed, a study of large Australian Organisations found that of the 580 Human Resource managers and General managers sampled, 94% coached their direct reports and 73% of those that coached did so at least once per week (McCarthy & Milner, 2020). Therefore, after careful consideration, it was decided to leave this response (and less extreme outliers) unchanged.

Thankfully, in accordance with the central limit theorem, the problem of non-parametric shaped sample data (does not approximate a bell-shaped curve) can be overcome (Field, 2013, February 15). To achieve this, one resamples the abnormally shaped sample data many times over (say 5000 times) using smaller samples of at least 30 cases and above (randomly selected from the sample data). As a result of this process, the sampling distribution

approximates a normal bell-shaped distribution curve, meaning the confidence intervals surrounding the model parameters (i.e. regression coefficients) are more precise (Field, 2013).

Given the non-parametric nature of this study's sample data, 95% bootstrap confidence intervals, derived from 5,000 bootstrap samples, are reported for all of the regression coefficients. The rationale for this is that unlike OLS confidence intervals, bootstrap confidence intervals are not dependent upon the data being normally distributed (Field, 2013; Hayes, 2017), and this provides increased confidence in the findings.

4.4 Model fitness

To achieve the aims of this study, the hypothetical model of employee coaching was fitted to the data collected from NZ Police employees, but how can we be assured this linear model is a good fit for the data it represents? In reality, no hypothetical model fits real-world data perfectly (Bell et al., 2019). That said, a common approach to demonstrating model-data-fit is to report model-fit-indices (Pousa et al., 2020; Ribeiro et al., 2020; Zhao & Liu, 2020; Zuñiga-Collazos et al., 2020).

Model fit indices are statistical indicators that tell us how well a hypothetical model explains the data it is fitted to. Fundamentally, they estimate a ratio of pattern to disorder (signal to noise) within a dataset (Field, 2013). Among an assortment of others, three commonly reported indices are chi-squared/degrees of freedom (χ^2/df); comparative fit index (CFI); and the goodness of fit index (GFI). For example, in their study exploring the effect of employee coaching on sales performance, Pousa et al. (2020, p. 1265) report χ^2/df=2.024; CFI= 0.935; and GFI=0.870. Each of these indices comes with a conventional benchmark indicating an acceptable fit. For example, the ratio of the chi-square to degrees of freedom should not exceed 5.0 (Hagen & Peterson, 2014; Hayduk, 1987); CFI ranges from 0–1, with larger numbers indicating a better fit: >0.90 (adequate fit); >0.95 (good fit) (Bagozzi & Yi, 1988; Hu & Bentler, 1999); and GFI ranges from 0–1 with larger numbers indicating a better fit: ≥0.90 (acceptable fit) (Baumgartner & Homburg, 1996).

The reporting of multiple model fitness indices is common for researchers choosing structural equation modelling (SEM) as an analytical strategy. Indeed, SEM enables the researcher to estimate multiple fitness indices for an entire hypothetical model (referred to as the measurement model in SEM) (Hayes, 2017). This is not the case with CPA. The sole measure of model fitness produced by the PROCESS macro is R^2 score (also known as the

coefficient of determination). Moreover, PROCESS produces an R^2 score for each of the consequential variables contained within the model. In the case of this study, it meant that four R^2 scores were generated and are presented in Table 7.

Table 7: Goodness of fit indices (R-squared) for the proposed conceptual model of employee coaching

Outcome Variable	X→M path	X→Y and M→Y paths
Employee engagement	$R^2 = 0.12$, $p<0.01$	$R^2 = 0.13$, $p<0.01$
Intention to turnover	$R^2 = 0.12$, $p<0.01$	$R^2 = 0.31$, $p<0.01$

R^2 is a goodness-of-fit measure that is commonly reported in linear regression modelling (Frost, 2019). It estimates the proportion of variance in a consequential variable (dispersal from the mean) that can be explained by the model fitted to it (Frost, 2019). In essence, the closer the data points are to the regression line, the better the model fits the data it purportedly represents. R^2 scores range from 0-1, with 1 indicating a perfect fit and 0 indicting no fit whatsoever. Researchers strive for an R^2 score of $\geq$0.90 (Baumgartner & Homburg, 1996); however, even a near perfect R^2 score may be misleading. Indeed, low R^2 scores can still be perfectly acceptable (particularly in the social sciences), provided one acknowledges that there is greater residual error between the data points and the regression line, and any predictions made using the model may be less precise (Frost, 2019).

One of the limitations of R^2 as an indicator of model fitness is that it fails to alert the researcher to the presence any 'patterns' of residual bias (Frost, 2019). For example, one can obtain a high R^2 score, yet the data points can be heteroscedastic, meaning the distribution around the regression line has an obvious pattern to it. Alternatively, one can obtain a low R^2 score, yet the data points can be homoscedastic, meaning the distribution around the regression line is random (no obvious pattern). If the data points are scattered randomly, the R^2 score is likely to be stochastic (unbiased). Conversely, if the data points have a pattern to them (predictable), it is more likely that the model contains stochastic error (Frost, 2019). When stochastic error is present, it is likely that the model fitted to the data is failing to pick up all of the variance in the consequential variable. The inference being that the model is missing one or more components and may need adjustment to better fit the data (Frost, 2019).

Ultimately, any residual error ought to be randomly distributed (Frost, 2019). As suggested by Frost (2019); Hayes (2017), SPSS was used to plot the residual errors for each of the

antecedent variables onto the model's fitted (predicted) values. This enabled a visual assessment for stochastic error.

Focusing first on PQECR and employee engagement (refer Figures 4 and 5 below), it is clear that as predicted value (standardised) increases, the residual value (standardised) decreases. This pattern suggests the presence of some stochastic error in both models is likely (Frost, 2019). In contrast, with reference to intention to turnover (Figure 6), there are no obvious patterns, so the presence of stochastic error in this model is considered unlikely (Frost, 2019).

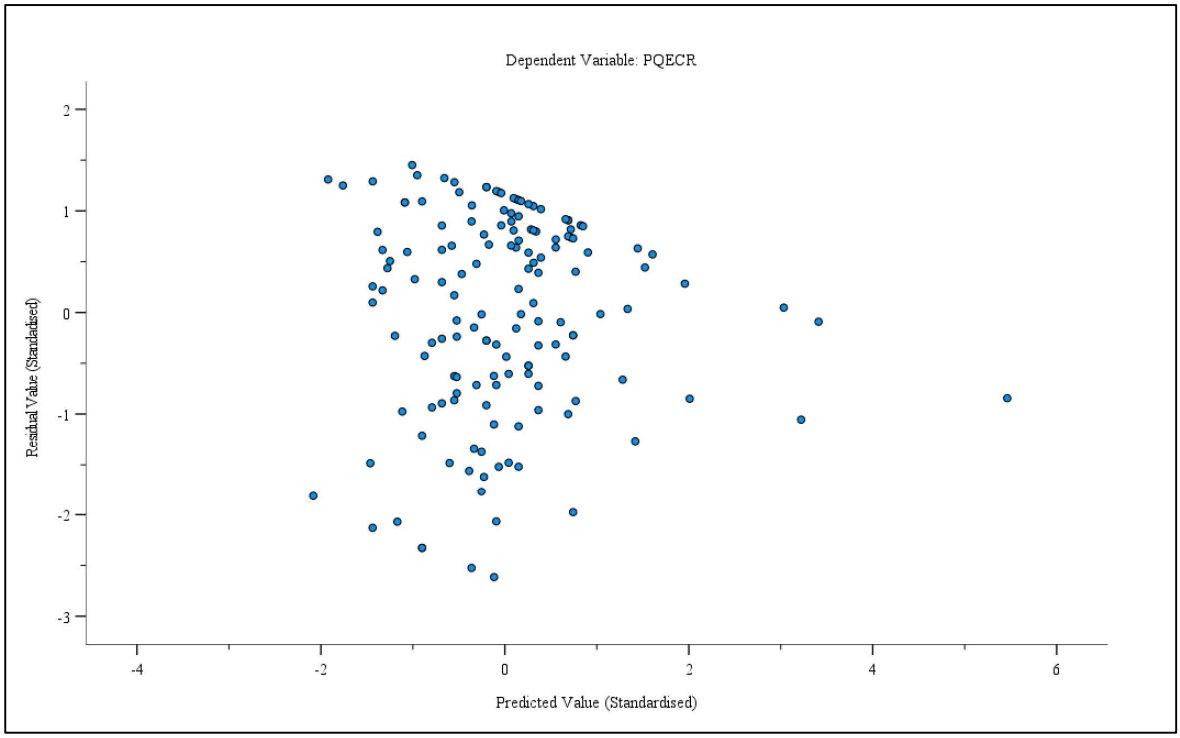

Figure 4: Residual plot (PQECR)

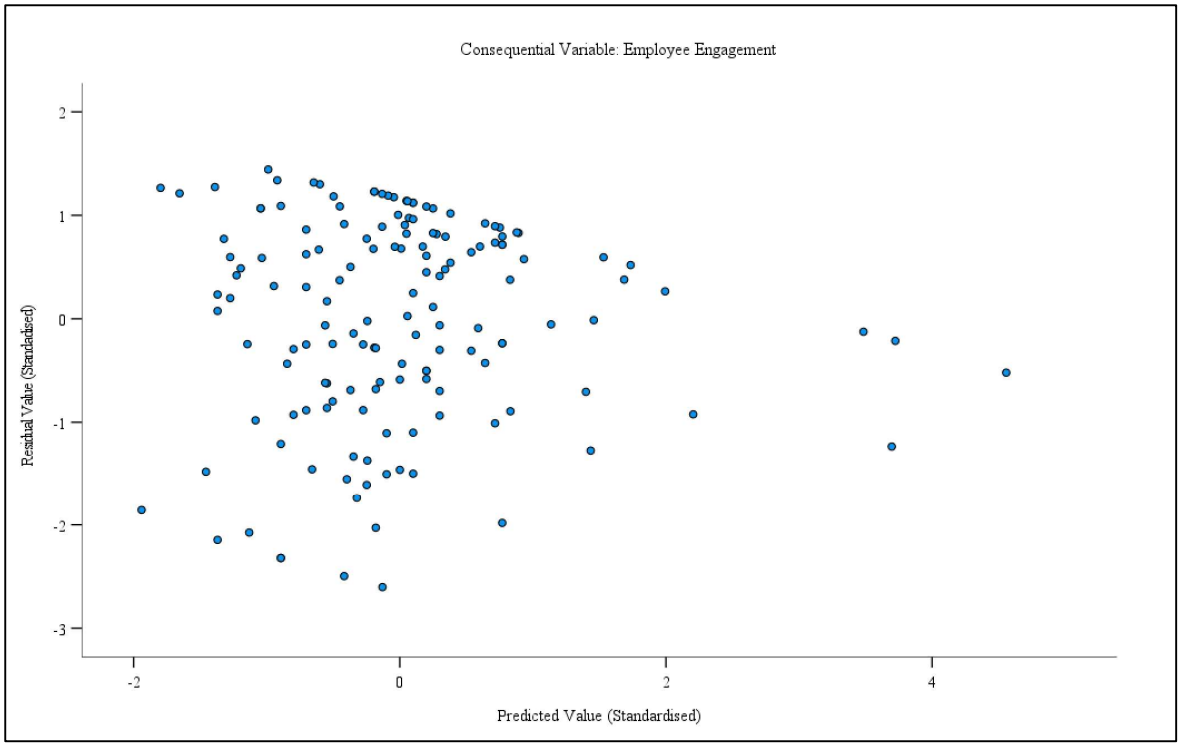

Figure 5: Residual plot (employee engagement)

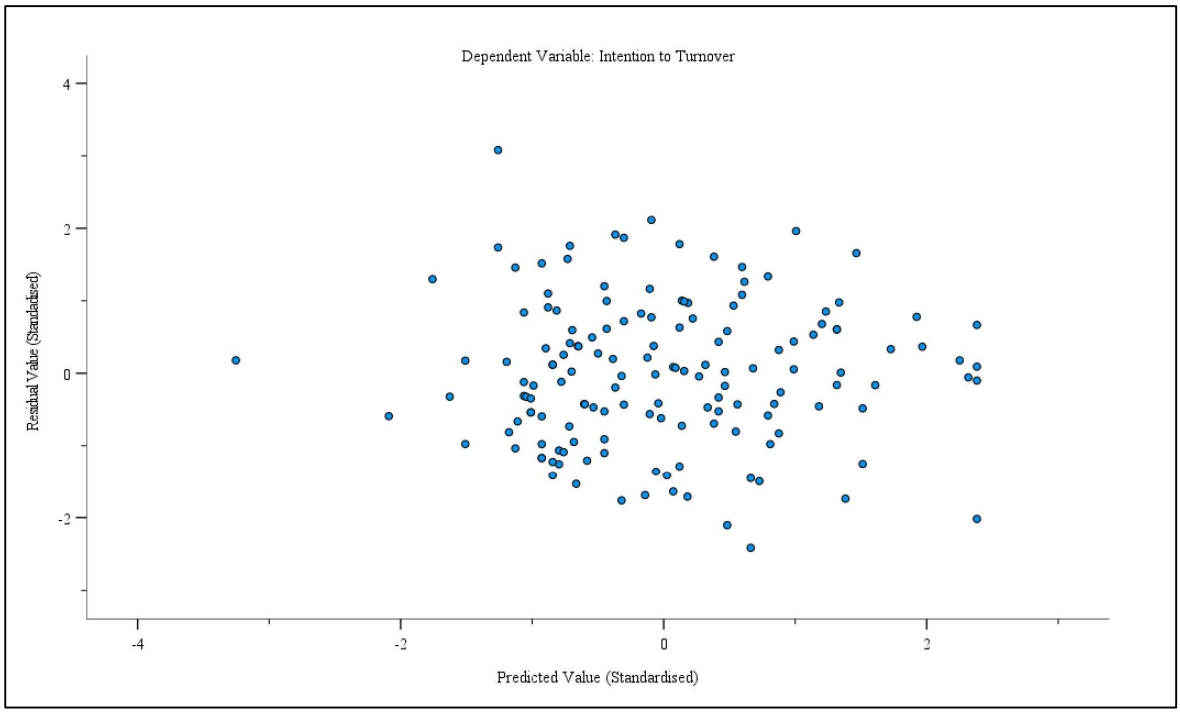

Figure 6: Residual plot (intention to turnover)

Having now estimated how well the linear model fits the dataset, it is equally important to think about the probability that the R^2 scores were obtained by random chance. To test the null hypotheses that R^2 is actually zero and the data has no pattern to it (random), the PROCESS macro carries out an *F*-test using the *F*-distribution and produces a p-value for the null hypothesis. In the case of this study, the *F*-statistic for each of the estimation models

was sufficiently large to produce p-values of <0.05, (Table 7). Given this, we can be assured that it is highly improbable that R^2 scores were obtained by chance.

4.5 Validity and reliability

Because this study used three psychometric scales to measure three independent variables, it makes sense to test their validity and reliability. Indeed, validity and reliability are cornerstones of quantitative research (Appelbaum et al., 2018; Bell et al., 2019; Ghauri et al., 2020). To this end, Chapter 2 provides accounts of the development, validity, and reliability of each of the psychometric scales, but how robust were they in the context of the sample specific data collected for this study? The following sections provide an account of this.

4.5.1 Scale validity

A popular way for demonstrating scale validity is to report the average variance extracted (AVE) (Ali & Aziz, 2018; Chen et al., 2019; Ismail et al., 2016; Pousa et al., 2020; Ribeiro et al., 2020; Tanskanen et al., 2019; Zhao & Liu, 2020; Zuñiga-Collazos et al., 2020). The rationale for this is that AVE is an indicator of convergent validity. That is, how well a scale's items cluster around the variable it purportedly measures (Bell et al., 2019). Citing Farrell (2010), Matsuo (2018) posits that the AVE estimates the average amount of variance within a data sample that the scale items explain. Indeed, to be acceptable, the AVE score should be ≥0.5 (Fornell & Larcker, 1981; Hair et al., 1995).

Another popular procedure for assessing scale validity is to calculate and report the square-root of the AVE (Ali & Aziz, 2018; Ismail et al., 2016; Pousa et al., 2020; Ribeiro et al., 2020). The reason for this is to demonstrate discriminant validity. That is, the degree to which each scale item discriminates in favour of the variable it is supposed to measure and against the variables it is not supposed to measure. When the square root of the AVE is greater than the bivariate correlations between variables, discriminant validity is achieved (Fornell & Larcker, 1981).

To calculate the AVE (and its square-root) for the psychometric scales used in this study, procedures recommended by Grande (2016, December 4) were followed. Sample data for each scale was entered into SPSS and a factor analysis was conducted using the varimax rotation method. The factor loadings produced by SPSS for each of the scales were transferred into Excel and labelled λ (lambda). At this point, an Excel calculation were performed to produce the AVE and its square root for each of the psychometric scales.

$\frac{\sum \lambda^2}{n}$ was the equation used to calculate the AVE, where n = the number of factor loadings. As per Table 8, the three psychometric scales used in this study exceed the minimum requirements for convergent and discriminative validity: AVE is greater than 0.5, and the square root of AVE is greater than the bivariate corelations reported in Table 4. This means that the PQECR, EES, and TISCALE are sufficiently valid for the purpose of this study.

4.5.2 Scale reliability

Having now covered scale validity, it is time to pivot toward scale reliability, and many authors achieve this by reporting on the internal consistency of the scale items (Ali & Aziz, 2018; Chen et al., 2019; Dahling et al., 2016; Hsieh & Huang, 2018; Ismail et al., 2016; Kunst et al., 2018; Pousa et al., 2020; Pousa et al., 2018; Ribeiro et al., 2020; Tanskanen et al., 2019; Weer et al., 2016; Ye et al., 2016; Zhao & Liu, 2020; Zuñiga-Collazos et al., 2020). Internal consistency means the degree to which the items within a psychometric scale are scored consistently. In other words, if a respondent scores high on one scale item, they should, theoretically, score high on all of the scale items because the scale items are intended to measure the same variable (Bell et al., 2019). To this end, composite reliability (CR) and Cronbach's coefficient alpha (α) are consistently reported within the literature. To be acceptable, the CR score for each variable must be $\geq$ 0.70 (Bagozzi & Yi, 1988; Fornell & Larcker, 1981; Hair et al., 1995; Kline, 2015; Nunnally, Bernstein, & Berge, 1978), and Cronbach's alpha (α) should be $\geq$ 0.5 (Hair et al., 2014).

To calculate the CR for the psychometric scales used in this study, procedures recommended by Grande (2016, December 4) were followed. Sample data was entered into SPSS and a factor analysis was conducted using the varimax rotation method. The factor loadings produced by SPSS were then transferred into Excel and labelled λ (lambda). At this point, an Excel calculation was performed to produce the CR for each of the psychometric scales.

$\frac{\sum \lambda^2}{\sum \lambda^2 + \sum \varepsilon}$ was the equation used to calculate the composite reliability for each psychometric scale, where the symbol ε (epsilon) represents the error variance ($1-\lambda^2$). Results are as per Table 8.

To calculate Cronbach's coefficient alpha (α) for the scales used in this study, procedures recommended by Grande (2014, November 11) were followed. Items pertaining to each scale were entered into SPSS and a reliability analysis was conducted.

As per Table 8, the three psychometric scales used in this study exceed the minimum requirements for CR and Cronbach's alpha: CR is greater than 0.70, and the α score for each scale exceeds 0.50. This means that the PQECR, EES, and TISCALE are sufficiently reliable for the purpose of this study.

Table 8: Validity and reliability statistics for the psychometric scales

Psychometric scale	AVE	Square-root of AVE	CR	Cronbach's alpha
PQECR	0.77	0.88	0.98	0.97
EES	0.68	0.82	0.96	0.90
TISCALE	0.58	0.76	0.71	0.85

4.6 Regression coefficients and confidence intervals for the statistical pathways

Table 9: Coefficients and confidence intervals for the statistical pathways

Model 1 Outcome variable: PEQCR (M)		Model 2 Outcome variable: employee engagement (Y)		Model 3 Outcome variable: intention to turnover (Y)	
Coefficients	Bootstrap CI[a]	Coefficients	Bootstrap CI[a]	Coefficients	Bootstrap CI[a]
a_1=.1129 Coaching frequency (X)	.0179→.2160	*b*=.1618 PQECR (M)	.0674→.2591	*b*=-.4522 PQECR (M)	-.5695→-.3318
a_2=-.0184 Tenure (W)	-.0432→ .0005	*c'*=.0245 Coaching frequency (X)	.084→.0473	*c'*=-.0478 Coaching Frequency (X)	-.0848→-.0144
a_3=-.0009 Interaction (XW)	-.0044→.0072	Index[b]= -.0001	-.0008→.0012	Index[b]= .0004	-.0032→.0021

a. Bootstrap samples=5,000
b. Index of Moderated Mediation

4.7 Hypothesis testing

Hypothesis 1 postulates that a positive association exists between coaching frequency and PQECR, and a statistical analysis supports this. As expected, the estimated effect is positive albeit small, *r*=0.11, and the 95% bootstrap confidence interval does not straddle zero, meaning that the effect is statistically significant. We may interpret this to mean that coaching frequency has a positive direct effect on PQECR, rejecting the null in favour of Hypothesis 1.

Hypothesis 2 postulates that a positive association exists between coaching frequency and employee engagement, and a statistical analysis supports this. As expected, the estimated effect is positive, albeit tiny, r=0.02, and the 95% bootstrap confidence interval does not cross zero meaning the effect is statistically significant. We may interpret this to mean that coaching frequency has a positive direct effect on employee engagement, rejecting the null in favour of Hypothesis 2.

Hypothesis 3 postulates that a positive association exists between PQECR and employee engagement, and a statistical analysis supports this. As expected, the estimated effect is positive albeit small, r=0.16, and the 95% bootstrap confidence does not include zero, meaning the effect is statistically significant. We may interpret this to mean that PQECR has a positive effect on employee engagement, rejecting the null in favour of Hypothesis 3.

Hypothesis 4 postulates that a negative association exists between coaching frequency and intention to turnover, and a statistical analysis supports this. As expected, the estimated effect is negative albeit tiny, r=-0.05, and the 95% bootstrap confidence interval does not include zero. We may interpret this to mean coaching frequency has a negative direct effect on intention to turnover, rejecting the null in favour of Hypothesis 4.

Hypothesis 5 postulates that a negative association exists between the perceived quality of the employee coaching relationship and intention to turnover, and a statistical analysis supports this. As expected, the estimated effect is negative and moderate, r=-0.45, and the 95% bootstrap confidence interval does not include zero. We may interpret this to mean that PQECR has a negative direct effect on intention to turnover, rejecting the null in favour of Hypothesis 5.

Hypothesis 6 postulates that the association between coaching frequency and PQECR is negatively moderated by tenure; however, a statistical analysis does not support this. As expected, the estimated effect of the interaction between coaching frequency and tenure on PQECR is negative but minuscule, <-0.01. Unsurprisingly, the 95% bootstrap confidence interval includes zero, >-0.01→<0.01, signalling the effect is not statistically significant. Moreover, the difference in model fit between two models, the second of which includes the interaction between coaching frequency and tenure is minuscule and not statistically significant, ΔR^2=<0.01, p=0.62. We may interpret this to mean that the association between coaching frequency and PQECR is not moderated by tenure, rejecting Hypothesis 6 in favour of the null.

4.8 Miscellaneous

Whilst the effect of coaching frequency on employee engagement and intention to turnover was found to be tiny, the same cannot be said for the effect of PQECR. In contrast to coaching frequency, the effect of PQECR on employee engagement and intention to turnover was found to be moderate, suggesting that PQECR is a meaningful predictor from an academic and practical perspective. Given that intention to turnover has been identified as an indicator of organisational wellbeing (Forsyth et al., 2021), this result may be of particular interest to NZ Police.

As expected, when coaching frequency is held constant, the partial effect of PQECR on employee engagement was found to be positive albeit small, r=0.16. Furthermore, the 95% bootstrap confidence interval excluded zero, 0.07→0.26, inferring that the effect is also statistically significant. This has been interpreted to mean that for two people who are equal on coaching frequency but differ by one unit on PQECR, the effect on employee engagement is 0.16 units.

Pivoting to intention to turnover, as expected, when coaching frequency is held constant, the partial effect of PQECR on intention to turnover was found to be negative and moderate, r=-0.45. Moreover, the 95% bootstrap confidence intervals exclude zero, -0.57→-0.33, inferring that the effect is statistically significant. This has been interpreted to mean that for two people who are equal on coaching frequency but differ by one unit on PQECR, the effect on intention to turnover is -0.45 units.

CHAPTER 5: DISCUSSION

5.1 Revisiting the problem

NZ Police have a laudable vision of making New Zealand the safest country in the world (New Zealand Police, 2018) and PHPF is a key mechanism for realising the vision. Given PHPF was introduced in 2017 (New Zealand Police, 2017), one would expect to observe a vigorous coaching culture by now; however, a small-scale study conducted in 2020 (Mowat, 2020) suggested otherwise and the results of this study confirm that progress has been slow. This could be a serious problem for NZ Police because its vision is ambitious and is not helped when employee coaching conversations have not been delivered as frequently as prescribed. Indeed, from the sample data, the mean number of one-on-one coaching conversations over the past six months was found to be 2.42, well short of the monthly frequency prescribed by PHPF (New Zealand Police, 2017). Concerningly, 44 respondents (29.33% of the sample) reported that they had not had a single coaching conversations with their direct supervisor in the past six months. This result is consistent with the small-scale study conducted in 2020 (Mowat, 2020) in which 31% of coachees reported the same.

5.2 Addressing the problem

Hypothesising that coaching frequency matters, this quantitative research project fitted the conceptual model of employee coaching to the sample data, revealing tiny but statistically significant direct and indirect relationship between coaching frequency and employee engagement, and coaching frequency and intention to turnover. As anticipated, when coaching frequency increases, so does employee engagement (directly and indirectly). Furthermore, intention to turnover abates (directly and indirectly). The data analysis further confirmed that PQECR is a legitimate mechanism by which coaching frequency exerts influence on the aforementioned outcome variables.

Somewhat surprisingly, it was found that tenure had no moderating effect on the above-mentioned relationships and as recommended, it was removed from the model (Hayes, 2017). The absence of any moderation effect answers the ‘when’ of the how. That is, the effect of coaching frequency on PQECR is constant regardless of how long a coachee has been employed by NZ Police.

5.3 Answering the research questions

This study set out to find answers to four research questions:

1. Is coaching frequency a useful predictor of employee engagement and intention to turnover in the context of the NZ Police?
2. Is the association between coaching frequency and employee engagement mediated by the perceived quality of the employee coaching relationship?
3. Is the association between coaching frequency and intention to turnover mediated by the perceived quality of the employee coaching relationship?
4. Is the direct association between coaching frequency and the PQECR moderated by employee tenure?

As recommended by Hayes (2017), a conditional process analysis was undertaken to help answer these questions and necessitated the estimation of three regression models, one for each of the outcome variables: PQECR, employee engagement, and intention to turnover (Table 9).

The answer to question 1 is no, albeit with some caution, and the answer to questions two and three is yes.

As anticipated, coaching frequency was found to predict employee engagement and intention to turnover both directly and when mediated via PQECR, and the estimated effects were found to be statistically significant (Table 9). Whilst this makes the predictions reliable and supports the case for PQECR as a legitimate mediator, the estimated effects are tiny, ranging between r=-0.05 and 0.11. This suggests that whilst coaching frequency is a meaningful predictor from an academic perspective, in practice it is not particularly useful.

Beginning with employee engagement, as expected, when tenure equals zero, the conditional indirect effect of coaching frequency on employee engagement was found to be positive albeit tiny, r=0.02; however, the 95% bootstrap confidence interval excludes zero, <0.01→0.04, inferring that the effect is statistically significant. We may interpret this to mean that for two people who differ by one unit on coaching frequency, there is a difference of 0.02 units on employee engagement, when tenure equals zero, and the effect is mediated via PQECR.

As expected, when PQECR is held constant, the partial direct effect of coaching frequency on employee engagement is positive albeit tiny, r=0.02; however, the 95% bootstrap confidence intervals exclude zero, 0.01→0.05, signalling the effect is statistically significant. This has been interpreted to mean that for two people who differ by one unit in

coaching frequency, there is a difference of 0.02 units on employee engagement, when PQECR is held constant.

Considering intention to turnover now, as expected, when tenure equals zero, the conditional indirect effect of coaching frequency is negative albeit tiny, *r*=-0.05; however, the 95% bootstrap confidence interval excludes zero -0.10→-0.01, signalling the effect is statistically significant. We may interpret this to mean that for two people who differ by one unit on coaching frequency, there is a difference of -0.05 units on intention to turnover when tenure equals zero, and the effect is mediated via PQECR.

As expected, when PQECR is held constant, the partial direct effect of coaching frequency on intention to turnover is negative albeit tiny, *r*=-0.05; however, the 95% bootstrap confidence intervals exclude zero, -0.08→-0.01, signalling the effect is statistically significant. This has been interpreted to mean that for two people who differ by one unit in coaching frequency, there is a difference of -0.05 units on intention to turnover, when PQECR is held constant.

With the rejection of Hypothesis 6 in favour of the null; the answer to question 4 is a firm no; however, this is constricted to the first stage of the two-stage mediation model, and some readers may be interested in whether the entire two-stage (mediated) indirect effect of coaching frequency on employee engagement and intention to turnover is moderated by tenure.

With reference to Figure 7, as expected, the Index for the conditional indirect effect of coaching frequency on employee engagement is negative albeit miniscule, *r*=-0.0001. Furthermore, the 95% bootstrap confidence interval includes zero, -0.0008→0.0012, signalling the effect in not statistically significant. We may interpret this to mean the indirect effect of coaching frequency on employee engagement is not moderated by tenure.

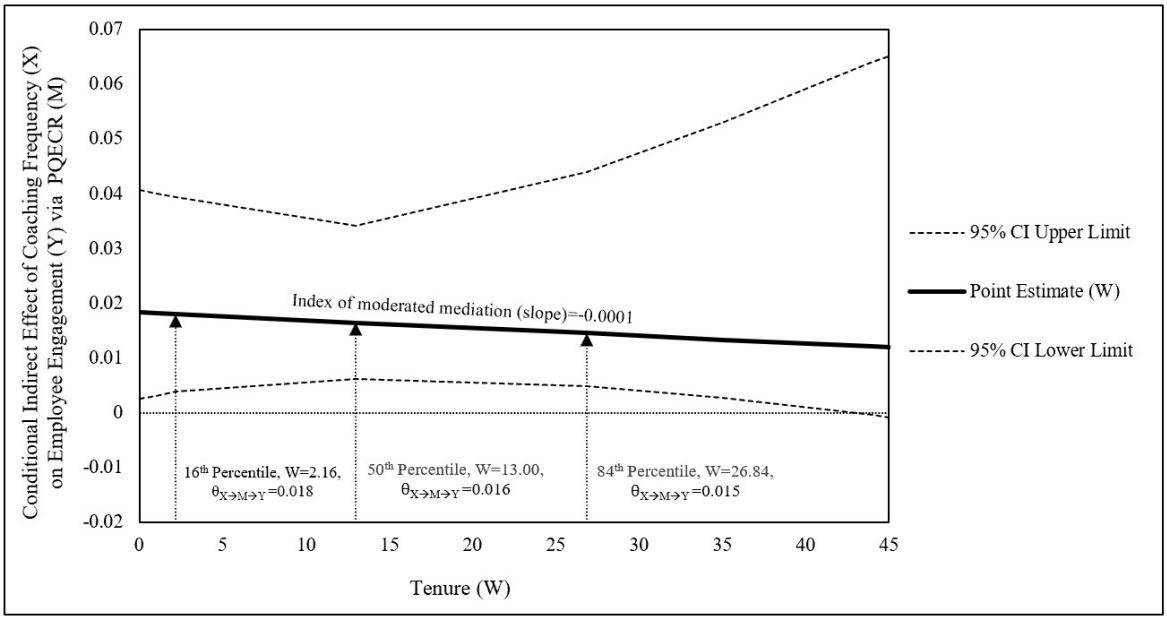

Figure 7: The conditional indirect effect of coaching frequency on employee engagement

Considering intention to turnover (Figure 8), as predicted, the Index for the indirect effect of coaching frequency is positive, albeit miniscule, r=0.0004; Moreover, the 95% bootstrap confidence interval includes zero, -0.0032→0.0021, signalling that the effect is not statistically significant. We may interpret this to mean the indirect effect of coaching frequency on intention to turnover is not moderated by tenure.

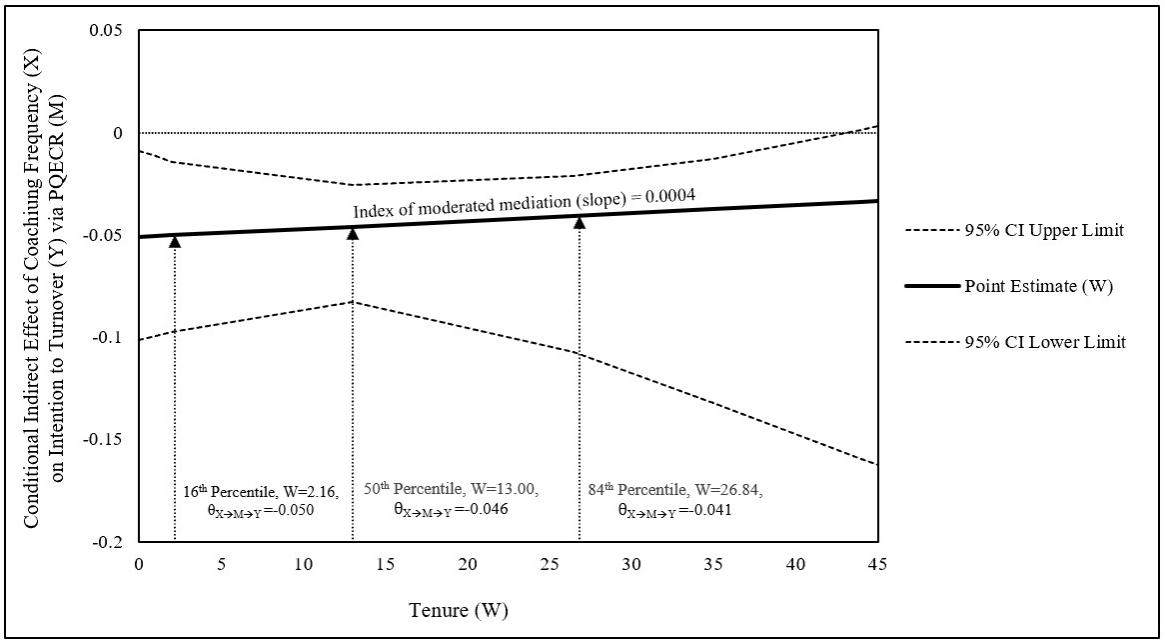

Figure 8: The conditional indirect effect of employee coaching on intention to turnover

Given that the moderating effect of tenure has been found to be miniscule and not statistically significant for either employee engagement or intention to turnover, it is entirely reasonable to remove tenure and its interaction with coaching frequency from the proposed

conceptual model of employee coaching (Hayes, 2017). Indeed, despite first appearances, the regression lines in Figure 7 and Figure 8 are both verging on the horizontal, meaning the variable represented on the X-axis (tenure) has very low predictive ability (Field, 2013). If this were not the case, the gradient of the regression lines representing each Index would be more substantive and statistically significant to the level of the *a priori* alpha value (Hayes, 2017).

5.4 How is this study different?

Arguably, the most interesting finding to emerge from this study is that, from an academic perspective at least, coaching frequency matters. This does not completely align with current thinking, with some suggesting that it is the coach's skill level that matters most, not the frequency to which coaching occurs (Dahling et al., 2016; Pousa et al., 2018). Indeed, in a study conducted by Dahling et al. (2016), the authors found that coaching frequency on its own (high and low) had no statistically significant effect on coachee goal attainment. Moreover, they suggest that low-performance is associated with unskilled coaches who coach frequently. Dahling et al. (2016) make their position clear: it is the quality of coaching that matters, not the quantity.

In contrast, the results of this study found that coaching frequency has a positive, albeit small, direct and indirect effect on employee engagement, and a negative, albeit small, direct and indirect effect on intention to turnover. Importantly, these effects were found to be statistically significant. The contrasting results could be caused by any number of factors, such as the differing consequential variables. In the study conducted by Dahling et al. (2016), the consequential variable was sales goal attainment whereas in the current study, it was employee engagement and intention to turnover. If true, it means that coaching frequency is not a universally influential antecedent variable. Rather the strength of its influence may vary depending on the consequential variable(s).

Within the frame of LMX theory, the results obtained were anticipated because in accordance with LMX theory, employee coaching is likely to be seen as a valuable resource within the context of an exchange relationship between a leader (coach) and each member of their team (coachee) (Ali et al., 2018; Griffin et al., 2017; Hsieh & Huang, 2018). Moreover, the success of employee coaching hinges on the perceived quality of the interpersonal relationship that exists between the coach and their coachee, not the coach's objective skill level (Dansereau et al., 1975). This is not to say that a coach's skill does not

matter because it almost certainly does (Dahling et al., 2016); however, even the most highly skilled coach is unlikely to have a significant impact unless the practice of coaching takes place on a reasonably frequent basis. Indeed, as per LMX theory, when making an offer to coach, the coach initiates a social exchange (Ali et al., 2018; Hsieh & Huang, 2018; Raza & Ahmed, 2020), and as the frequency to which coaching is offered increases, so do the opportunities to build and maintain a high-quality interpersonal relationship (and improve coaching skill). Therefore, the more frequently coaching takes place, the more opportunities there are for value to be exchanged (Hsieh & Huang, 2018).

Another interesting feature of this study is that it is almost certainly the first independent study to use the PQECR and EES psychometric scales developed to measure the perceived quality of the employee coaching relationship and employee engagement. The rationale for choosing the PQECR and EES scales over more popular options was because they were the best fit for the study aim. The PQECR was developed for the specific purpose of measuring a coachee's perceived quality of the employee coaching relationship they share with their direct supervisor, and the EES measures a broad perspective of employee engagement in the work context. Along with the TISCALE, the validity and reliability of the PQECR and EES were tested and exceeded the acceptable benchmarks (Chapter 4). Therefore, this study makes a small contribution toward strengthening the validity and reliability of all three psychometric scales.

Notwithstanding that there are almost certainly additional mediators capable of transmitting the effect of coaching frequency onto any number of consequential variables (Hayes, 2017), this study has presented empirical evidence that the PQECR is one such mediator variable (Chapter 4), and it is the first independent study (outside of scale development) to do so. This contribution may be considered valuable because it answers calls to explore different mediators of employee coaching generally (Pousa et al., 2020), and in particular, the PQECR (Echeverri, 2020; McCarthy & Milner, 2020; Ribeiro et al., 2020). By doing so, this study fills a gap in the literature.

Along with exploring unique mediators, this study also explored the moderating effect of employee tenure. From the literature, no empirical evidence was found to suggest that tenure would moderate the effect of coaching frequency on the consequential variables of PQECR, employee engagement, and intention to turnover, which is another gap in the literature. That said, it was hypothesised that as a coachee's tenure increases, their knowledge, skill, and ability is expected to develop, meaning they are less likely to rely less on their coach.

Therefore, it came as a surprise that tenure had no moderating effect on any of the aforementioned relationships. This finding is likely to be appreciated because others have posited that coachee tenure could have a significant moderating effect on the practice of employee coaching (McCarthy & Milner, 2020; Pousa et al., 2018; Weer et al., 2016) and have called for further investigation.

5.5 Implications for theory

As highlighted in Section 2.2, the theoretical frames through which employee coaching has been explained in the literature are abundant; however one of the more popular theories among the scholarship is LMX, which is the theoretical lens used in the current study. The results obtained are consistent with LMX in that they support the argument that employee coaching frequently forms part of the social exchange that takes place between a supervisor and each of their direct reports, the success of which depends on the quality of the interpersonal relationship (Dansereau et al., 1975). Indeed, this study found that in general terms, as coaching frequency increases, the PQECR strengthens, and as PQECR strengthens, there is a positive effect on employee engagement and a negative effect on intention to turnover. Therefore, this study makes a small contribution to theory by validating LMX as an appropriate theoretical lens for understanding the employee coaching phenomenon.

5.6 Implications for practice

Within the literature, numerous studies advocate for supervisor-employee coaching as a means of enhancing organisational performance (Hicks & McCracken, 2013; Hunt & Weintraub, 2016; Longenecker, 2010), improve personal gratification (McCarthy & Milner, 2020), improve job satisfaction (Hagen & Peterson, 2014), and improve perceived job value (Lee, Liu, Rousseau, Hui, & Chen, 2011; Lewis & Smithson, 2000; Rousseau & Arthur, 1999). Given that this study has produced empirical evidence that coaching frequency matters, and the effect of coaching frequency on employee engagement and intention to turnover is statistically significant, the NZ Police Executive may consider making changes to organisational settings to strengthen the organisation's existing coaching culture.

It must be reiterated that despite being statistically significant, the effects of coaching frequency on the three outcome variables of PQECR, employee engagement, and intention to turnover were found to be small, and from a practice perspective, may hardly matter. That said, when coaching frequency was held constant, the partial effect of PQECR on intention

to turnover was found to be moderate in size, r=-0.45, and statistically significant, which is meaningful from both an academic and practice perspective.

The practice of employee coaching is not without challenges. In their study of 580 coaches based in large Australian organisations, McCarthy and Milner (2020) identified a list of six barriers to employee coaching:

- Lack of opportunity: some coaches struggle to balance the demands of their dual role as coach and manager
- Hesitancy to coach: some coaches are fearful of a backlash from their coachees, particularly if constructive feedback is perceived to be critical
- Reluctant coachees: some coachees prefer to be given direct instructions rather than to be coached
- Lack of coaching ability (skill): some coaches lack the necessary coaching competencies
- Lack of motivation: some coaches resist devolving power and responsibility to their coachees
- Lack of organisational support: the structures and systems of some organisations restrain coaching practice

The above list is consistent with a small-scale study conducted by the researcher in 2020 (N=22), in which six Police Sergeants (first-level-supervisors) perceived PHPF coaching conversations as a low priority. This is because they were rewarded (or at least not reprimanded) for prioritising operational delivery (driving prevention activity, triaging calls for service, and field supervision) over performance management. Other barriers identified by the Sergeants included inappropriate mindset due to shift-work fatigue, Constable unavailability, and being too busy attending calls for service (time-poor).

Interestingly, the results of the current study indicate that the mean value of PQECR is lowest for Senior Police Managers at 3.63. The next is Constables (including Police Employee equivalents) at 3.72, followed by Sergeants (including Police Employee equivalents) at 3.83, and the highest group was Senior Sergeants (including Police Employee equivalents) at 4.13. This suggests that the barriers to employee coaching are most prevalent for those occupying the senior ranks of NZ Police.

5.7 Limitations

This study is not without limitations. In particular, readers should be aware that all of the study participants were employees of NZ Police, a public sector organisation with a unique culture and idiosyncrasies, so caution should be exercised before generalising findings beyond this context.

A second limitation concerns causation. Conditional process analysis (as used in the current study) certainly uses the language of causation (Hayes, 2017). For example, the proposed conceptual model of employee coaching (Figure 1) has directional arrows pointing from the antecedent variable to the consequential variable, and this infers that the antecedent variable 'causes' the consequential. However, it must be made manifestly clear that this study adopted a cross-sectional design, meaning all of the data was collected at a single point in time, so it is not possible to confirm the causal orderings of the variables. For example, it could be that employee engagement causes coaching frequency to increase rather than the other way around. That said, whilst those adopting an experimental, or longitudinal design can make a stronger case for causal order, these designs are not problem-free either, particularly for mediation analysis (Hayes, 2017). Ultimately, causation is an argument developed by the researcher and is based on a 'tripod' of three factors: solid theory (logic), data analysis, and study design (Hayes, 2017).

A third limitation of this project is that the findings are, in part, based on the statistical analysis of a modest data sample, N=150. As discussed in greater depth in Chapter 3, despite a concerted effort, there were significant challenges gaining access to the NZ Police population. Indeed the sample obtained is convenient in nature and equates to just over 1% of the NZ Police workforce. That said, the majority of the results reported in Chapter 4 were found to be statistically significant to the *a priori* alpha value of $\alpha=0.05$ which provides a measure of reassurance to readers. Moreover, the results of this study are consistent with the findings of the small-scale study carried out by the researcher in 2020 (Mowat, 2020).

Finally, it is worthwhile pointing out the participants in this study were not randomly selected from the NZ Police population. Rather, as permitted by the Evidence Based Policing Centre, the researcher distributed the survey to smaller targeted groups via internal networks. The potential risk with such an approach is that the results may be non-representative of the average NZ Police employee (Bell et al., 2019). For example, it could be that a disproportionate number of participants are people who are interested in employee coaching. Alternatively, they may be a cohort who enjoy filling in surveys. The point here is that the safeguard (check-and-balance) against potential biases that random selection can reduce was

not applied to this study. That said, the majority of social science researchers do not adopt random sampling, choosing alternative techniques to control for potential bias, i.e. including potential confounders in a regression analysis (Hayes, 2017).

5.8 Future research opportunities

This study was carried out in the context of the NZ Police, and it would be interesting to see if similar effects are observed in one or more different organisations. One of the ways that the external reliability of the results from this study could be strengthened is by replicating the study in a different context, such as within a private organisation that encourages employee coaching. It is suspected that the effects will be similar; however, one never knows for certain until similar results are gathered from a variety of contexts (Bell et al., 2019).

Because this study adopted a cross sectional design, it is essentially a snapshot in time. This means it is possible that the observed effects are atypical, like an unusually hot day in winter. As a check against this, an opportunity exists to repeat the study at one or more future points in time. This would make the study design longitudinal in nature, and such a test-retest approach may build confidence that the observed effects are reliable if they are found to be stable over time (Bell et al., 2019). Again, this is something that is currently unknown.

As per Chapter 2, some have argued that coaching skill is more important than coaching frequency (Dahling et al., 2016; Pousa et al., 2018). Whilst acknowledging that coaching skill almost certainly matters, the aim of this study was to determine whether the same can said about coaching frequency, and the results suggest it is (albeit the effect is small). That said, one of the ways this study could be extended would be to repeat it using coaching skill as a moderator of the effect of coaching frequency on PQECR rather than tenure. Should the index of moderated mediation produce a meaningfully sloped regression line, showing that the indirect effect of coaching frequency on employee engagement and/or intention to turnover becomes statistically significant only when coaching skill strengthens, this would support an argument that coaching skill matters over coaching frequency. However if the index of moderation produced a flat regression line, yet the indirect effect of coaching frequency on employee engagement and intention to turnover remained meaningful and statistically significant from zero, this would support an argument that coaching frequency is effective regardless of the supervisor's level of coaching skill.

CHAPTER 6: CONCLUSIONS

NZ Police has set itself the aspirational vision of making New Zealand the safest country in the world (New Zealand Police, 2017) and is to be admired for accepting such a challenging objective. However, something more is needed to take the lead from Iceland, who, according to the Institute of Economic and Peace, has been the most peaceful country in the world for the past 14 years (Institute for economics & peace, 2021). To help the organisation achieve its vision, the NZ Police introduced the PHPF, a performance management system designed to draw-out the potential of its 14,500 employees (New Zealand Police, 2017). PHPF prescribes that NZ Police supervisors facilitate monthly one-on-one coaching conversations to equip and enable each of their directs reports to perform to their best (New Zealand Police, 2017), but a small-scale study conducted by the author in 2020 (Mowat, 2020) suggested that these coaching conversations are happening less frequently than prescribed and in some cases, not at all. This presents a potential road-block for the NZ Police because the organisation relies on its workforce to deliver outcomes, and success is less likely when employees are not coached to perform to their best.

But does coaching frequency really matter? Some say no, arguing that it is the quality of the coaching that matters, not the frequency (Dahling et al., 2016; Pousa et al., 2018). According to the results of this study, coaching frequency, from an academic perspective at least, does matter. Data collected from 150 NZ Police employees was analysed using conditional process analysis as prescribed by Hayes (2017). The hypothetical model of employee coaching was fitted to the data and the results suggest that when mediated by PQECR, the effect of coaching frequency on employee engagement and intention to turnover is statistically significant, albeit small. Moreover, it was found that employee tenure has no moderating influence on the strength of these effects. The point here is that coaching frequency matters regardless of the coachee's level of service, and it remains advantageous to offer regular employee coaching to all employees.

The results from this study align with and can be explained by LMX theory which posits that employees and their supervisors engage in a form of social exchange, albeit in the context of the workplace. When a supervisor offers an employee an inducement (i.e. employee coaching) the employee will perceive this as valuable and is likely reciprocate by offering their loyalty to the supervisor and/or increased commitment toward the team/organisation's goals (Dansereau et al., 1975; Raza et al., 2017). Therefore, in according

with LMX theory, one would expect that as coaching frequency increases, employee engagement would strengthen and intention to turnover would abate, and this is exactly what the study found.

Given the results of this study, it is recommended that the NZ Police Executive consider what (if any) structural changes could be made to ensure that the frequency to which employee coaching is delivered better aligns with the monthly cycle prescribed by PHPF (New Zealand Police, 2017). Indeed, the average participant of this study experienced a one-on-one coaching conversation with their supervisor once every three months. This is well short of what is prescribed and identifies an area for improvement. Given the above, it is certainly worth considering providing ongoing training to all NZ Police employees in supervisory positions on how to use employee coaching to strengthen their coachees' PQECR, and this is supported in the literature (McCarthy & Milner, 2020).

Done well, employee coaching takes significant effort on behalf of supervisors and their direct reports, but the advantages to organisations and their employees is well recognised in the literature: enhancing organisational performance (Hicks & McCracken, 2013; Hunt & Weintraub, 2016; Longenecker, 2010), improving personal gratification (McCarthy & Milner, 2020), and improving job satisfaction (Hagen & Peterson, 2014). Moreover, according to others, it has been suggested that "coaching is the heart of managerial effectiveness." (Matsuo, 2018, p. 128), and "organizational success can be deeply influenced by how well managers coach subordinates." (Ribeiro et al., 2020, p. 2). In support of this, the results of this study suggest that when coaching frequency increases, it has a positive, (albeit small) effect on PQECR, and when PQECR strengthens, employee engagement increases and intention to turnover wanes. Having a stable and engaged workforce certainly advantage NZ Police as the organisation strives to realise its vision.

REFERENCES

Aguinis, H., Ramani, R., & Alabduljader, N. (2018). What you see is what you get? Enhancing methodological transparency in management research. *Accadamy of Management Annals, 12*(1), 1-28.

Aiken, L., West, S., & Reno, R. (1991). *Multiple regression: testing and interpreting interactions*. Sage Publications.

Ali, M., & Aziz, S. (2018). Relationship between managerial coaching and employee role behaviors: moderating role of south asian culture. *South Asian Studies, 33*(1), 221-238.

Ali, M., Lodhi, S., Orangzab, Raza, B., & Ali, W. (2018). Examining the impact of managerial coaching on employee job performance: mediating role of work engagement, leader-member-exchange quality, job satisfaction, and turnover Intentions. *Pakistan Journal of Commerce & Social Sciences, 12*(1), 253-282.

Appelbaum, M., Cooper, H., Kline, R., Nezu, A., Mayo-Wilson, E., & Rao, S. (2018). Journal article reporting standards for quantitative research in psychology: the APA Publications and Communications Board Task Force report. *American Psychologist, 73*(1), 3-25.

Bagozzi, R., & Yi, Y. (1988). On the evaluation of structure equation models. *Journal of the Academy of Marketing Science, 16*(1), 74-94.

Bakker, A., Albrecht, S., & Leiter, M. (2011). Key questions regarding work engagement. *European Journal of Work and Organizational Psychology, 20*(1), 4-28.

Barry, K., Gloeckner, G., & Kaiser, L. (2021). Managerial coaching competencies used by managers for performance improvement. *Performance Improvement Quarterly, 34*(2), 195-217.

Bauer, D., & Curran, P. (2005). Probing interactions in fixed and multilevel regression: inferential and graphical techniques. *Multivariate behavioral research, 40*(3), 373-400.

Baumgartner, H., & Homburg, C. (1996). Applications of structural equation modeling in marketing and consumer research: a review. *International journal of Research in Marketing, 13*(2), 139-161.

Bell, E., Bryman, A., & Harley, B. (2019). *Business research methods* (5th ed.). Oxford University Press.

Blau, P. (1964). *Exchange and power in social life*. Wiley.

Bluedorn, A. (1982). A unified model of turnover from organizations. *Human Relations, 35*(2), 135-153.

Boshoff, A., Van Wyk, R., Hoole, C., & Owen, J. (2002). The prediction of intention to quit by means of biographic variables, work commitment, role strain and psychological climate. *Management Dynamics: Journal of the Southern African Institute for Management Scientists, 11*(4), 14-28.

Brown, T. (2015). *Confirmatory factor analysis for applied research*. Guilford Publications.

Chatered institute of personnel and development. (2015). *Learning and development 2015 survey report*. https://www.cipd.co.uk/Images/learning-development_2015_tcm18-11298.pdf

Chen, G., Sharma, P., Edinger, S., Shapiro, D., & Farh, J. (2011). Motivating and demotivating forces in teams: cross-level influences of empowering leadership and relationship conflict. *Journal of Applied Psychology, 96*(3), 541.

Chen, I., Khan, S., & Lin, C. (2019). Team resource management perception under managerial coaching skills and organizational climate: cross-level analysis in Taiwan's hospitals. *Health Care Manager, 38*(3), 228-238.

Colarelli, S. (1984). Methods of communication and mediating processes in realistic job previews. *Journal of Applied Psychology, 69*(4), 633-642.

Currivan, D. (1999). The causal order of job satisfaction and organizational commitment in models of employee turnover. *Human Resource Management Review, 9*(4), 495-524.

Dahling, J., Taylor, S., Chau, S., & Dwight, S. (2016). Does coaching matter? A multilevel model linking managerial coaching skill and frequency to sales goal attainment. *Personnel Psychology, 69*(4), 863-894.

Dansereau, F., Graen, G., & Haga, W. (1975). A vertical dyad linkage approach to leadership within formal organizations: a longitudinal investigation of the role making process. *Organizational Behavior and Human Performance, 13*(1), 46-78.

Davies, C., & Fisher, M. (2018). Understanding research paradigms. *Journal of the Australasian Rehabilitation Nurses' Association (JARNA), 21*(3), 21-25.

Deming, W. (2000). *Out of the crisis*. First MIT Press.

DeNisi, A., & Murphy, K. (2017). Performance appraisal and performance management: 100 years of progress? *Journal of Applied Psychology, 102*(3), 421-433.

Denzin, N. K. (2012). Triangulation 2.0*. *Journal of Mixed Methods Research, 6*(2), 80-88.

Dictionary.com. (2021). https://www.dictionary.com/browse/tenure

Dixey, A. (2015). Managerial coaching: a formal process or a daily conversation? *International Journal of Evidence Based Coaching and Mentoring*(9), 77-89.

Du Plessis, T., & Van Niekerk, A. (2017). Factors influencing managers' attitudes towards performance appraisal. *SA Journal of Human Resource Management, 15*(0), 1-10.

Dwivedi, S. (2015). Turnover intentions: scale construction & validation. *The Indian Journal of Industrial Relations, 50*(3), 452-468.

Echeverri, P. (2020). Value-forming micro-practices of managerial coaching. *Coaching, 13*(2), 191-208.

Ellinger, A., Ellinger, A., & Keller, S. (2003). Supervisory coaching behavior, employee satisfaction, and warehouse employee performance: a dyadic perspective in the distribution industry. *Human Resource Development Quarterly, 14*(4), 435-458.

Farrell, A. (2010). Insufficient discriminant validity: a comment on Bove, Pervan, Beatty, and Shiu (2009). *Journal of Business Research, 63*(3), 324-327.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). Sage Publications Ltd.

Field, A. (2013, February 15). *The central limit theorem*. YouTube. https://www.youtube.com/watch?v=ermii2fQWOo

Field, A. (2020, October 1). *Null hypothesis significance testing* [video]. YouTube. https://www.youtube.com/watch?v=IXYDMMBisr8

Field, A. (2020, Sep 23). *Lecture 01: the linear model* [video]. YouTube. https://www.youtube.com/watch?v=7cSArk7tU4w&t=1942s

Fornell, C., & Larcker, D. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39-50.

Forsyth, D., Ashby, L., Gardner, D., & Tappin, D. (2021). *The New Zealand Workplace Barometer: psychosocial safety climate and worker health - findings from the 2020 NZ Workplace Barometer*. https://www.massey.ac.nz/massey/fms/school_of_management/HWG/The%202020%20New%20Zealand%20Workplace%20Barometer%20Report%20FINAL.pdf?

Fournies, F. (1978). *Coaching for improved work performance*. Van Nostrand Reinhold Company.

Frost, J. (2019). *Regression analysis: an intuative guide for using and interpreting linear models*. Statistics by Jim Pulishing.

Gallup. (2017). *NZ Police workplace survey 2017*. https://www.police.govt.nz/sites/default/files/publications/wps-2017-nz-police-organisational-report.pdf

Ghauri, P., Grønhaug, K., & Strange, R. (2020). *Research methods in business studies* (5th ed.). Cambridge University Press.

Graham, S., Wedman, J., & Garvin–Kester, B. (1993). Manager coaching skills: development and application. *Performance Improvement Quarterly, 6*(1), 2-13.

Grande, T. (2014, November 11). *Calculating and interpreting Cronbach's Alpha using SPSS*. YouTube. https://www.youtube.com/watch?v=Kz8OdR6lV44

Grande, T. (2016, December 4). *Average variance extracted and composite reliability after factor analysis using SPSS and Excel*. YouTube. https://www.youtube.com/watch?v=8tyjdfpiAJQ

Grant, A. (2017). The third ‘generation’of workplace coaching: creating a culture of quality conversations. *Coaching: An International Journal of Theory, Research and Practice, 10*(1), 37-53.

Gregory, J., & Levy, P. (2010). Employee coaching relationships: enhancing construct clarity and measurement. *Coaching: An International Journal of Theory, Research and Practice, 3*(2), 109-123.

Griffin, R., Phillips, J., & Gully, S. (2017). *Organizational behavior: managing people and organizations* (12th ed.). Cengage Learning.

Gyllensten, K., & Palmer, S. (2007). The coaching relationship: an interpretative phenomenological analysis. *International Coaching Psychology Review, 2*(2), 168-177.

Hagen, M., & Peterson, S. (2014). Coaching Scales: a review of the literature and comparative analysis. *Advances in Developing Human Resources, 16*(2), 222-241.

Hair, J., Anderson, R., Tatham, R., & William, C. (1995). *Multivariate data analysis with readings*. Prentice Hall.

Hair, J., Black, W., Babin, B., & Anderson, R. (2014). *Multivariate data analysis* (7th ed.). Prentice Hall.

Hamlin, R., Ellinger, A., & Beattie, R. (2006). Coaching at the heart of managerial effectiveness: a cross-cultural study of managerial behaviours. *Human Resource Development International, 9*(3), 305-331.

Harvey, W. (2010). Methodological approaches for interviewing elites. *Geography Compass, 4*(3), 193-205.

Hayduk, L. (1987). *Structural equation modeling with LISREL: essentials and advances*. Jhu Press.

Hayes, A. (2017). *Introduction to mediation, moderation, and conditional process analysis: a regression-based approach*. Guilford Publications.

Heslin, P., Vandewalle, D., & Latham, G. (2006). Keen to help? Managers' implicit person theories and their subsequqnt employment coaching. *Personnel Psychology, 59*(4), 871-902.

Hesse-Biber, S. (2016). Qualitative or mixed methods research inquiry approaches: some loose guidelines for publishing in sex roles. *Sex Roles, 74*, 6-9.

Heywood, A. (2013). *Politics* (4th ed.). Palgrave Macmillan.

Hicks, R., & McCracken, J. (2013). *Coaching as a leadership style*. Taylor & Francis.

Holman, D., Batt, R., & Holtgrewe, U. (2007). *The global call center report: international perspectives on management and employment*. https://ecommons.cornell.edu/bitstream/handle/1813/74325/GCC_Intl_Rept_US_Version.pdf

Hsieh, H., & Huang, J. (2018). Exploring factors influencing employees' impression management feedback-seeking behavior: the role of managerial coaching skills and affective trust. *Human Resource Development Quarterly, 29*(2), 163-180.

Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal, 6*(1), 1-55.

Hudson, M. (2010). *Te ara tika: guidelines for Māori research ethics: a framework for researchers and ethics committee members*. Health Research Council of New Zealand on behalf of thePūtaiora Writing Group.

Hunt, J., & Weintraub, J. (2016). *The coaching manager: developing top talent in business*. Sage Publications.

Institute for economics & peace. (2021). *Global peace index: 2021*. https://www.economicsandpeace.org/wp-content/uploads/2021/06/GPI-2021-web.pdf

Ismail, A., Ahmad, N., & Zainol, A. (2016). Managerial coaching in enhancing employees' motivation [Article]. *Acta Universitatis Danubius: Oeconomica, 12*(3), 98-112.

Johnson, P., & Fay, L. (1950). The Johnson-Neyman technique, its theory and application. *Psychometrika, 15*(4), 349-367.

Kahn, W. (1990). Psychological conditions of personal engagement and disengagement at work. *Academy of management journal, 33*(4), 692-724.

Kara, H. (2020). *Research ethics: theory and practice.* https://www.ncrm.ac.uk/resources/online/all/?main&id=20591

Kerr, S. (1995). An Academy classic: on the folly of rewarding A, while hoping for B. *Academy of Management Perspectives, 9*(1), 7-14.

Kilburg, R. (2001). Facilitating intervention adherence in executive coaching: a model and methods. *Consulting Psychology Journal: Practice and Research, 53*(4), 251-257.

Kline, R. (2015). *Principles and practice of structural equation modeling*. Guilford Publications.

Kuhn, T. (1970). *The structure of scientific revolutions* (2nd ed.). University of Chicago Press.

Kunst, E., van Woerkom, M., Poell, R., & van Kollenburg, G. (2018). Stability and change in teachers' goal orientation profiles over time: managerial coaching behavior as a predictor of profile change. *Journal of Vocational Behavior, 104*, 115-127.

Lakens, D. (2017). *Power analysis and sample size decisions: SPSP experts: Open Science* [video]. You Tube. https://www.youtube.com/watch?v=KIRwsYTR62A

Lawrence, P. (2017). Managerial coaching - a literature review. *International Journal of Evidence Based Coaching & Mentoring, 15*(2), 43-69.

Lawson Williams. (2019). *New Zealand staff turnover: summary survey report 2019*. https://www.lawsonwilliams.co.nz/national-staff-turnover-survey-summary-report

Lee, C., Liu, J., Rousseau, D., Hui, C., & Chen, Z. (2011). Inducements, contributions, and fulfillment in new employee psychological contracts. *Human Resource Management, 50*(2), 201-226.

Lewis, J., & Smithson, J. (2000). Is job insecurity changing the psychological contract? *Personnel Review, 29*(6), 680-702.

Liu, X., & Batt, R. (2010). How supervisors influence performance: a multilevel study of coaching and group management in technology-mediated services. *Personnel Psychology, 63*(2), 265-298.

Löfgren, K. (2013, Aug 5). *Normality test using SPSS: how to check whether data is normally distributed* [video]. YouTube.

https://www.youtube.com/watch?v=IiedOyglLn0&list=FLw9lyIcCdF60aND8Or35bKQ&index=1&t=30s

London, M., & Smither, J. (2002). Feedback orientation, feedback culture, and the longitudinal performance management process. *Human Resource Management Review, 12*(1), 81-100.

Longenecker, C. (2010). Coaching for better results: key practices of high performance leaders. *Industrial and Commercial Training, 42*(1), 32-40.

Luthans, F., Avolio, B., Avey, J., & Norman, S. (2007). Positive psychological capital: measurement and relationship with performance and satisfaction. *Personnel Psychology, 60*(3), 541-572.

Martin, G. (2020, August 1). *Understanding missing data and missing values. 5 ways to deal with missing data using R programming*. YouTube. https://www.youtube.com/watch?v=sUAMiAIUhcI&t=394s

Massey University. (2017). *Code of ethical conduct for research, teaching and evaluations involving human participants: revised code 2017*. https://www.massey.ac.nz/massey/research/research-ethics/human-ethics/code-ethical-conduct.cfm

Matsuo, M. (2018). How does managerial coaching affect individual learning? The mediating roles of team and individual reflexivity. *Personnel Review, 47*(1), 118-132.

Matthews, J. (2010). Can line managers ever be effective coaches? *Business Leadership Review, 7*(2), 1-10.

McCarthy, G., & Milner, J. (2020). Ability, motivation and opportunity: managerial coaching in practice. *Asia Pacific Journal of Human Resources, 58*(1), 149-170.

Misiukonis, T. (2011). The conclusions middle managers draw from their beliefs about organisational coaching and their coaching practices. *International Journal of Evidence Based Coaching & Mentoring*. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.689.3055&rep=rep1&type=pdf

Mobley, W., Griffeth, R., Hand, H., & Meglino, B. (1979). Review and conceptual analysis of the employee turnover process. *Psychological Bulletin, 86*(3), 493-522.

Morgan, D. (2007). Paradigms lost and pragmatism regained: methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research, 1*(1), 48-76.

Mowat, J. (2020). *The systematic constriction of performance conversations between police supervisors and employees* [Unpublished manuscript]. Massey Business School, Massey University.

New Zealand Police. (2017). *The police high perfromance framework*.

New Zealand Police. (2018). *Our business*.

New Zealand Police. (2021). *Annual report: 2020/21*. https://www.police.govt.nz/sites/default/files/publications/annual-report-2020-2021.pdf

Nunnally, J., Bernstein, I., & Berge, J. (1978). *Psychometric theory*. McGraw-Hill.

O'Loughlin, E. (2020, November 14). *How to... test for normality in SPSS*. YouTube. https://www.youtube.com/watch?v=68sYzkCRv9E

Oxford Brookes University. (n.d.). International centre for coaching and mentoring studies.

Panhwar, A., Ansari, S., & Shah, A. (2017). Post-positivism: an effective paradigm for social and educational research. *International Research Journal of Arts & Humanities (IRJAH), 45*(45), 253-259.

Park, J., Johnson, D., Moon, K., & Lee, J. (2019). The interaction effects of frequency and specificity of feedback on work performance. *Journal of Organizational Behavior Management, 39*(3-4), 164-178.

Park, S., Yang, B., & McLean, G. (2008). An examination of relationships between managerial coaching and employee development. *Online submission*. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.943.935&rep=rep1&type=pdf

Podsakoff, P., MacKenzie, S., Lee, J., & Podsakoff, N. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879-903.

Policing Act. (2008). http://www.legislation.govt.nz/act/public/2008/0072/latest/DLM1102125.html

Pousa, C., Liu, Y., & Aman, A. (2020). The effect of managerial coaching on salesperson's relationship behaviors: new evidence from frontline bank employees in China. *International Journal of Bank Marketing, 38*(6), 1259-1277.

Pousa, C., Richards, D., & Trépanier, C. (2018). Managerial coaching of frontline employees: the moderating role of gender. *Human Resource Development Quarterly, 29*(3), 219-241.

Price, J. (1977). *The study of turnover* (1st ed.). Iowa State University Press.

Raza, B., & Ahmed, A. (2020). Linking managerial coaching and workplace deviance: the mediating role of thriving at work. *Iranian Journal of Management Studies, 13*(3), 467-494.

Raza, B., Ali, M., Ahmed, S., & Moueed, A. (2017). Impact of managerial coaching on employee performance and organizational citizenship behavior: intervening role of thriving at work. *Pakistan Journal of Commerce & Social Sciences, 11*(3), 790-813.

Ribeiro, N., Nguyen, T., Duarte, A., Torres de Oliveira, R., & Faustino, C. (2020). How managerial coaching promotes employees' affective commitment and individual performance. *International Journal of Productivity and Performance Management*. https://eprints.qut.edu.au/205603/1/69322968.pdf

Rich, B., Lepine, J., & Crawford, E. (2010). Job engagement: antecedents and effects on job performance. *Academy of management journal, 53*(3), 617-635.

Rogosa, D. (1980). Comparing nonparallel regression lines. *Psychological Bulletin, 88*(2), 307.

Rousseau, D., & Arthur, M. (1999). The boundaryless human resource function: building agency and community in the new economic era. *Organizational dynamics, 27*(4), 7-18.

Saks, A. (2006). Antecedents and consequences of employee engagement. *Journal of Managerial Psychology, 21*(7), 600-619.

Schaufeli, W., Bakker, A., & Salanova, M. (2006). The measurement of work engagement with a short questionnaire: a cross-national study. *Educational and Psychological Measurement, 66*(4), 701-716.

Shuck, B., Adelson, J., & Reio Jr, T. (2017). The employee engagement scale: initial evidence for construct validity and implications for theory and practice. *Human Resource Management, 56*(6), 953-977.

Sobel, M. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological methodology, 13*, 290-312.

Spiller, S., Fitzsimons, G., Lynch Jr, J., & McClelland, G. (2013). Spotlights, floodlights, and the magic number zero: simple effects tests in moderated regression. *Journal of Marketing Research, 50*(2), 277-288.

Stats NZ. (n.d.). https://www.stats.govt.nz/tools/2018-census-place-summaries/new-zealand#ethnicity-culture-and-identity

Stokes, P. (2011). *Key concepts in business and management research methods*. Palgrave Macmillan.

Stone, R. (2017). *Human resource management* (9th ed.). John Wiley & Sons.

Tanskanen, J., Mäkelä, L., & Viitala, R. (2019). Linking managerial coaching and leader–member exchange on work engagement and performance. *Journal of Happiness Studies, 20*(4), 1217-1240.

University of Sydney. (n.d.). *Coaching psychology unit*. https://www.sydney.edu.au/science/our-research/research-areas/psychology/coaching-psychology-unit.html

Wagstaff, C., Arthur, C., & Hardy, L. (2018). The development and initial validation of a measure of coaching behaviors in a sample of army recruits. *Journal of Applied Sport Psychology, 30*(3), 341-357.

Weer, C., DiRenzo, M., & Shipper, F. (2016). A holistic view of employee coaching. *Journal of Applied Behavioral Science, 52*(2), 187-214.

Whitmore, J. (2002). *Coaching for performance*. Nicholas Brealey Publishing.

Ye, R., Wang, X., Wendt, J., Wu, J., & Euwema, M. (2016). Gender and managerial coaching across cultures: female managers are coaching more. *International Journal of Human Resource Management, 27*(16), 1791-1812.

Zhao, H., & Liu, W. (2020). Managerial coaching and subordinates' workplace well-being: a moderated mediation study. *Human Resource Management Journal, 30*(2), 293-311.

Zuñiga-Collazos, A., Castillo-Palacio, M., Montaña-Narváez, E., & Castillo-Arévalo, G. (2020). Influence of managerial coaching on organisational performance. *Coaching, 13*(1), 30-44.

Appendix A: NZ Police approval

EVIDENCE — BASED
POLICING • CENTRE • • •

NEW ZEALAND POLICE
Ngā Pirihimana o Aotearoa

27 April 2022

Jesse Mowat
Email: jesse.mowat.2@uni.massey.ac.nz

Dear Jesse

Research Application: EV-12-575 Employee Coaching Project

Thank you for your interest in conducting research with New Zealand Police.

The New Zealand Police Research Panel have considered the above application and have provisionally approved this request subject to a number of conditions. These include the following:

a) Please sign the New Zealand Police Research Agreement and acknowledge that you will be undertaking the research under these conditions. If the Principal Researcher is employed by or affiliated to a New Zealand University, the University Research Office is required to sign and administer the Research Agreement along with the Principal Researcher.
b) It is important that you are familiar with and adhere to the New Zealand Police Policy for External Researchers.
c) Before requesting data, information or approaching resource that is not defined in the original research proposal, the Principal Researcher will need to seek written permission from the Director of the Evidence Based Policing Centre. This may require a variation or amendment to the research agreement.
d) Your sponsor will determine the best way to distribute your survey throughout New Zealand Police, in consultation with the Evidence Based Policing Centre.

Vanessa Merrett, Principal Advisor Organisational Learning has agreed to be the sponsor and point of contact at NZ Police.

I look forward to seeing the results from this work as I believe it to be important research that will offer new insights. I wish you well with progress and offer my support should you need anything further.

Yours sincerely,

Simon Williams
Director Evidence Based Policing, New Zealand Police
Lambton House, 160 Lambton Quay, Wellington

cc: Vanessa Merrett
Vanessa.merrett@police.govt.nz

Dr Darryl Forsyth
d.forsyth@massey.ac.nz

Associate Professor David Brougham
d.brougham@massey.govt.nz

Appendix B: Employee coaching journal articles

1. Pousa, C., Liu, Y., & Aman, A. (2020). The effect of managerial coaching on salesperson's relationship behaviors: new evidence from frontline bank employees in China. *International Journal of Bank Marketing, 38*(6), 1259-1277.
2. Echeverri, P. (2020). Value-forming micro-practices of managerial coaching. *Coaching, 13*(2), 191-208.
3. Barry, K., Gloeckner, G., & Kaiser, L. (2021). Managerial coaching competencies used by managers for performance improvement. *Performance Improvement Quarterly, 34*(2), 195-217.
4. Raza, B., & Ahmed, A. (2020). Linking managerial coaching and workplace deviance: the mediating role of thriving at work. *Iranian Journal of Management Studies, 13*(3), 467-494.
5. Zhao, H., & Liu, W. (2020). Managerial coaching and subordinates' workplace well-being: a moderated mediation study. *Human Resource Management Journal, 30*(2), 293-311.
6. Zuñiga-Collazos, A., Castillo-Palacio, M., Montaña-Narváez, E., & Castillo-Arévalo, G. (2020). Influence of managerial coaching on organisational performance. *Coaching, 13*(1), 30-44.
7. McCarthy, G., & Milner, J. (2020). Ability, motivation and opportunity: managerial coaching in practice. *Asia Pacific Journal of Human Resources, 58*(1), 149-170.
8. Ribeiro, N., Nguyen, T., Duarte, A., Torres de Oliveira, R., & Faustino, C. (2020). How managerial coaching promotes employees' affective commitment and individual performance. *International Journal of Productivity and Performance Management*.
9. Chen, I., Khan, S., & Lin, C. (2019). Team resource management perception under managerial coaching skills and organizational climate: cross-level analysis in Taiwan's hospitals. *Health Care Manager, 38*(3), 228-238.
10. Tanskanen, J., Mäkelä, L., & Viitala, R. (2019). Linking managerial coaching and leader–member exchange on work engagement and performance. *Journal of Happiness Studies, 20*(4), 1217-1240.
11. Pousa, C., Richards, D., & Trépanier, C. (2018). Managerial coaching of frontline employees: the moderating role of gender. *Human Resource Development Quarterly, 29*(3), 219-241.

12. Hsieh, H., & Huang, J. (2018). Exploring factors influencing employees' impression management feedback-seeking behavior: the role of managerial coaching skills and affective trust. *Human Resource Development Quarterly, 29*(2), 163-180.
13. Kunst, E., van Woerkom, M., Poell, R., & van Kollenburg, G. (2018). Stability and change in teachers' goal orientation profiles over time: managerial coaching behavior as a predictor of profile change. *Journal of Vocational Behavior, 104*, 115-127.
14. Ali, M., Lodhi, S., Orangzab, Raza, B., & Ali, W. (2018). Examining the impact of managerial coaching on employee job performance: mediating role of work engagement, leader-member-exchange quality, job satisfaction, and turnover Intentions. *Pakistan Journal of Commerce & Social Sciences, 12*(1), 253-282.
15. Matsuo, M. (2018). How does managerial coaching affect individual learning? The mediating roles of team and individual reflexivity. *Personnel Review, 47*(1), 118-132.
16. Lawrence, P. (2017). Managerial coaching - a literature review. *International Journal of Evidence Based Coaching & Mentoring, 15*(2), 43-69.
17. Dahling, J., Taylor, S., Chau, S., & Dwight, S. (2016). Does coaching matter? A multilevel model linking managerial coaching skill and frequency to sales goal attainment. *Personnel Psychology, 69*(4), 863-894.
18. Ye, R., Wang, X., Wendt, J., Wu, J., & Euwema, M. (2016). Gender and managerial coaching across cultures: female managers are coaching more. *International Journal of Human Resource Management, 27*(16), 1791-1812.
19. Weer, C., DiRenzo, M., & Shipper, F. (2016). A holistic view of employee coaching. *Journal of Applied Behavioral Science, 52*(2), 187-214.
20. Azman, I., Nursaadatun Nisak, A., & Azmi, Z. (2016). Managerial coaching in enhancing employees' motivation. *Acta Universitatis Danubius: Oeconomica, 12*(3), 98-112.

Appendix C: Ethical approval

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

18/02/2022

Dear: Jesse Mowat

Re: Ethics Application - SOA 22/01 - Employee coaching project

Thank you for the above application that was considered by the Massey University Human Ethics Committee:

Human Ethics Southern A Committee at their meeting held on **Tuesday, 8 February 2022**

On behalf of the Committee I am pleased to advise you that the ethics of your application are approved.

Approval is for three years. If this project has not been completed within three years from the date of this letter, reapproval must be requested.

If the nature, content, location, procedures or personnel of your approved application change, please advise the Secretary of the Committee.

Yours sincerely

Professor Craig Johnson
Chair, Human Ethics Chairs' Committee and Director (Research Ethics)

Research Ethics Office, Research and Enterprise
Massey University, Private Bag 11 222, Palmerston North, 4442, New Zealand T 06 951 6841; 06 95106840
E humanethics@massey.ac.nz; animalethics@massey.ac.nz; gtc@massey.ac.nz

Appendix D: Participant information sheet

Participant Information: Research on Employee Coaching

Kia ora

My name is Jesse Mowat, and I am both a member of the New Zealand Police and a postgraduate student at Massey University. In my capacity as a postgraduate student, I am undertaking a research project that explores the coaching relationship between New Zealand Police employees and their direct supervisors. Specifically, I am interested in the association between coaching frequency, employee engagement, and intention to leave the NZ Police.

It is my pleasure to invite you to be a participant in this research project. Specifically, I am interested in your experience as a 'coachee' (a person who may receive job related coaching from their direct supervisor). Should you accept, your contribution will help to shape the results and will produce knowledge about the nature and quality of employee coaching within the NZ Police. Once complete, the study results will be shared with police employees and the police executive to consider whether the organisation's current policies and/or practice settings are delivering the best outcomes for everyone. In addition, I intend to publish in one or more external academic journals because the study itself and the results/findings are likely to be of interest to others engaged in the academic examination of workplace coaching. The data you provide will be subjected to Massey University's open access policy. This means it could be reused many times and contribute to additional projects.

Please be assured that your identity will remain anonymous. To make sure of this, you will not be asked for your name or any other identifying particulars. Also, the survey itself and the data that is collected from you will be hosted on a secure platform that is independent of the police information system. In addition, individual responses will not be disclosed to Police Managers and the results/findings will be aggregated to ensure that no individual can be identified.

There is an anonymous link (see below) that will navigate you to an online survey. The survey asks several questions about your experience as a coachee in the NZ Police and takes about 5 minutes to complete. Your participation is entirely voluntary, so you can withdraw anytime up until you have completed the questionnaire (non-completed questionnaires will not be included in the study). That is all, nothing more to do!

This project has been reviewed and approved by the Massey University Human Ethics Committee: Southern A, Application 22/01. If you have any concerns about the conduct of this research, please contact Dr Negar Partow, Chair, Massey University Human Ethics Committee: Southern A, telephone 04 801 5799, extension 63363, email humanethicsoutha@massey.ac.nz.

Thanks for your time; it is much appreciated.

Noho ora mai

Jesse Mowat

Click the link below to complete the survey (or paste the link into your internet browser)

https://masseybusiness.asia.qualtrics.com/jfe/form/SV_1z5e0SqOehxhM1M

Appendix E: Survey questionnaire

Perceived quality of the employee coaching relationship (PQECR)

1. My supervisor and I have mutual respect for one another
2. I believe my supervisor truly cares about me
3. I believe my supervisor feels a sense of commitment to me
4. My supervisor is a good listener
5. My supervisor is easy to talk to
6. My supervisor is effective at communicating with me
7. I feel at ease talking with my supervisor about my job performance
8. I am content to discuss my concerns or troubles with my supervisor
9. I feel safe being open and honest with my supervisor
10. My supervisor helps me to identify and build upon my strengths
11. My supervisor enables me to develop as an employee of our organisation
12. My supervisor engages in activities that help me unlock my potential

Employee engagement scale (EES)

1. I am really focused when I am working
2. I concentrate on my job when I am at work
3. I give my job responsibility a lot of attention
4. At work, I am focused on my job
5. Working for the New Zealand Police has a great deal of personal meaning for me
6. I feel a strong sense of belonging to my job
7. I believe in the purpose of the New Zealand Police
8. I care about the future of the New Zealand Police
9. I really push myself to work beyond what is expected of me
10. I am willing to put in extra effort without being asked
11. I often go beyond what is expected of me to help my team be successful
12. I work harder than expected to help the New Zealand Police be successful

Intention to turnover scale (TISCALE)

1. I intend to ask people about new job opportunities outside of the New Zealand Police
2. My current job is not addressing my important personal needs
3. Very often, opportunities to achieve my most important goals at work are jeopardised
4. I intend to search for a position with another employer
5. I occasionally think about leaving the New Zealand Police
6. Very often, I think about becoming an entrepreneur (operating my own business)

Coaching frequency

1. In the past six months, how many one-on-one PHPF meetings (coaching conversations) has your direct supervisor had with you to discuss/improve your performance on the job (if none, enter 0)?

Tenure

1. How many years have you worked for the New Zealand Police (enter 0 if less than 12 months)?

Descriptive statistics

1. Select one option the best represents your current age
2. Select one option that best represents the District or Service Centre you report to
3. In what capacity are you employed by the New Zealand Police?
4. Which of the following best represents the hours you are contracted to work for NZ Police?
5. Select one option that best represents your level of position in the New Zealand Police

Other

1. Are there any additional comments you would like to make about the practice of employee coaching in the New Zealand Police?

Appendix F: NZ Police research agreement

New Zealand Police Research Agreement

THIS AGREEMENT is made on 27 April 2022

BETWEEN Her Majesty the Queen in Right of Her Government in New Zealand acting by and through the Commissioner of Police ("Police")

AND the Principal Researcher, Dr Darryl Forsyth, of Massey University

AND the Researcher, Associate Professor David Brougham, of Massey University

AND the Researcher, Jesse Mowat, of Massey University

AND Massey University

Please note: If the principal researcher is employed by or affiliated to a NZ University the University will be the party to this agreement. This agreement should be entered into through the University's Research Office.

BACKGROUND

The NZ Police (hereinafter Police) want to make high-quality data available for research:

1. in as much detail as is necessary and possible
2. as widely as practicable
3. as soon as possible
4. as conveniently as is reasonable having regard to the impact on the activities of Police

while ensuring all legislative and ethical obligations governing access to, and safekeeping of, individualised and personal information are followed.

a. The Principal Researcher has submitted to Police an application to undertake research, including a Research Proposal as set out as Schedule One ("The Application"). [attached].

b. The Researcher has submitted The Application after having read and understood the Police Policy for External Researchers Access to Resources, Data or Privileged Information

c. Police has accepted and approved The Application.

d. This Agreement documents the terms and conditions upon which Police allows the Researcher to conduct research accessing the resources of Police. The scope of the research is detailed in the approved Research Proposal appended as Schedule One.

e. If the Researcher wishes, at any stage, for additional individuals to undertake research (or to substitute individuals) they must first obtain Police consent in writing and understand that those additional individuals may first need to clear appropriate and reasonable security and additional checks before undertaking research.

f. The Researcher agrees to conduct research in accordance with The Application. [attached]

g. The Researcher has approval from an accredited institutional ethics committee, or the proposal has been reviewed by a recognised human ethics body.

h. Other than information being gathered for the research, the Researcher agrees to keep confidential all information about Police and its operations about which the Researcher becomes aware and where this information is not in the public domain. This condition survives expiry or termination of The Project and this Agreement.

i. Researchers are welcome to provide comments on their experiences with conducting research with Police to research@police.govt.nz that will be included within the review of the Police Policy for External Researchers Access to Resources, Data or Privileged Information that will occur every 24 months.

THEREFORE IT IS AGREED AS FOLLOWS:

1 INTERPRETATION

1.1 "Police" means the New Zealand Police.

1.2 "Principal Researcher" means an individual who takes responsibility for the project or collaboration of researchers, the lead researcher, or the supervisor of any student research.

1.3 "Researcher" means any person working on The Project.

1.4 "The Researcher" means the collaboration that is the Principal Researcher and all researchers directed to work on the project by the Principal Researcher.

1.5 "Police Liaison Officer" means Vanessa Merrett, Principal Advisor, Organisational Learning

1.6 "Police Subject matter Expert" means Vanessa Merrett, Principal Advisor, Organisational Learning

1.7 "The Project" means EV-12-575: Employee coaching project

1.8 "Police information" includes any data held by Police or produced through the use of any Police activity, or produced using any resource(s) that belongs to Police.

1.9 "Privileged Police Information" includes:

(a) information stored within any Police database, file or documentation, not otherwise publically available; and

(b) the views and information supplied by current and former Police employees interviewed or surveyed for The Project;

1.10 "Personal Information" has the same meaning as in the Privacy Act 1993.

1.11 "Ethical Standards" means The Royal Society of New Zealand Code of Professional Standards and Ethics.

1.12 "Release" means submit, present, publish, disseminate or otherwise disclose The Project Findings.

1.13 "Project Findings" includes Privileged Police Information; any derived data, analyses and results from analyses, the recordings of interviews and any draft or completed Project Reports.

1.14 Subject to clause 12, the "Term of this Agreement" is from the date of this agreement until:

(a) The Project Findings are released, or

(b) The Project is terminated, (whichever occurs first).

1.15 Singular and Plural: The singular includes the plural and vice versa.

2 GENERAL TERMS AND CONDITIONS

2.1 The Researcher proposes to undertake a research project, "The Project".

2.2 The Project requires the Researcher to access relevant Police Information.

2.3 Police agree to make available to the Researcher such resources, Police Information and help as is feasible and reasonably necessary for The Project.

2.4 It is the Principal Researcher's responsibility to have had the project assessed by an approved Ethics Committee if such is necessary.

2.5 The Researcher agrees to maintain the security and confidentiality of all Privileged Police Information obtained for The Project.

2.6 Only those individuals identified in The Application and named in this Agreement as Researchers may undertake The Project. Prior to commencing the research, The Researcher will meet all of Police's security or any other requirements (which may relate to Health and Safety, for example) relating to access to any institution and to participants.

3 JURISDICTION

3.1 This Agreement is governed by and construed according to the law of New Zealand.

3.2 Where the Principal Researcher and any other Researcher is resident out of New Zealand, or carries out any work related to The Project out of New Zealand, or in any case where any foreign court might otherwise have been considered to have

any jurisdiction in regard to this Agreement, the Principal Researcher agrees to submit to the exclusive jurisdiction of the New Zealand Courts.

4 STATUS AND OBLIGATIONS OF THE PRINCIPAL RESEARCHER

4.1 The Principal Researcher must ensure that all Researchers conduct The Project in accordance with this Agreement.

4.2 The Principal Researcher has declared any relevant interests, all associations or any employment which might meaningfully impact on this agreement. Where the Principal Researcher is employed by a University the University will be party to this agreement.

5 POSSESS PRIVILEGED POLICE INFORMATION LAWFULLY

5.1 The Principal Researcher and any other Researcher may only access Privileged Police Information if Police approve their access to it pursuant to the Police Policy for External Researchers' Access to Resources, Data or Privileged Information and this Research Agreement.

5.2 The Principal Researcher and any other Researchers must not access, attempt to access, or use Police property or Privileged Police Information unless it is for the purpose of The Project.

5.3 The Principal Researcher and any other Researcher must comply with section 50 of the Policing Act 2008. Under this section, it is an offence to have Police property without lawful authority, or reasonable excuse.

6 MAINTAIN INFORMATION SECURITY AND CONFIDENTIALITY

6.1 The Principal Researcher and any other Researcher shall protect Privileged Police Information against unauthorised access, use, modification, disclosure, misuse or loss by storing Privileged Police Information on computers or devices with at least:

a. individual user challenge and authentication (user name and password)

b. individual user access logging firewalls;

c. intrusion detection system and server authentication;

d. operating system and application security measures (e.g. up to date antivirus screening, anti-"spyware" measures);

e. protecting Privileged Police Information stored electronically with passwords;

f. Protection of information or datasets electronically communicated over a public network (such as the internet).

And in the case of physical storage, printouts or media holding the data must be stored in lock cabinets.

6.2 The Principal Researcher and any other Researcher shall not:

a. disclose any Privileged Police Information to any third party.

b. use for any purpose (including educational or further research purposes) any Privileged Police information obtained during the Term of this Agreement other than for the purpose of The Research Project.

c. attempt to data-match or identify individuals in data which has been confidentialised.

d. make any copies of Privileged Police Information except where reasonably required to permit the research.

6.3 In relation to outputs from data analysis, the Researcher shall:

a. apply any confidentiality rules stipulated in the metadata, by Police and/or the data owner, before outputs are released.

b. actively consider whether outputs could be an inappropriate disclosure risk even with the confidentiality rules applied and take further steps to protect the output if necessary.

c. not provide outputs to anyone who is not an authorised researcher with the Project unless the outputs have been sufficiently confidentialised.

6.4 As soon as possible, and no later than 3 weeks after the end of the Term of this Agreement, the Principal Researcher shall provide the Police Liaison Officer with a schedule outlining all Privileged Police Information that Researchers hold. Police will provide directions as to whether and how to return the information or destroy it securely and the Principal Researcher and any other Researcher will comply with these directions, either returning the information or destroying it securely accordingly.

6.5 When destroying Privileged Police Information the Principal Researcher and any other Researcher:

a. must not place hard copy material in ordinary office wastepaper or rubbish bins;

b. must shred hard copies (to below 5mm), or dispose of them as secure waste (e.g. through a contractor such as the DDS "blue bin" system);

c. must delete all electronic copies from systems, computers, and devices; and

d. must confirm in writing that all Privileged Police Information has been returned or disposed of securely, as directed by Police.

6.6 The Principal Researcher shall agree to release data collected under the provisions of this Agreement to Police upon request, unless there is an ethical consideration that prevents the Principal Researcher from doing so.

6.7 If the Principal Researcher discloses, accidentally or otherwise, any Privileged Police Information in the possession of the Principal Researcher where not permitted by this Agreement, or becomes aware or suspects that any person has

obtained unauthorised access to or has used or attempted to use any Privileged Police Information for purposes not permitted by this Agreement, the Principal Researcher will:

a. immediately notify the Director of the Evidence Based Policing Centre, New Zealand Police;

b. use best endeavours to retrieve the Privileged Police Information and identify the person and secure the Privileged Police Information against any further unauthorised access or misuse.

7 CONDUCT RESEARCH

7.1 Where Police has approved The Project, the Principal Researcher shall ensure that all Researchers:

a. carry out The Project in accordance with The Research Project Proposal, which is attached as a Schedule to this Agreement;

b. follow any statistical obligations and quality rules stipulated in the metadata, by Police and/or the data owner.

c. work in consultation with the assigned Police Subject Matter Expert.

8 ADHERE TO A REPORTING SCHEDULE

8.1 The Principal Researcher will e-mail an update on the progress of The Research Project to research@police.govt.nz at intervals specified in the proposal or every 3 months.

8.2 At any time, Police may request the Principal Researcher to provide an update of progress, covering what has been achieved, what results have been produced, planned publication, and a list of the current research team. Refer to clause 9.4 for requirements in relation to Police review of research outputs prior to dissemination.

8.3 The Principal Researcher will provide Police with advance notification that they are about to seek Police peer review, via an e-mail to research@police.govt.nz at least 10 days before seeking peer review.

8.4 The Principal Researcher and any other Researcher will use the full Project reference (see clause 1.7) in all correspondence with Police.

9 RESULTS OF THE RESEARCH

9.1 The principal researcher must agree to a reasonable timeframe for completing the research and publishing results.

9.2 All results must include an appropriate reference to the source of the data collection and appropriate acknowledgement to the efforts of Police staff.

9.3 All results must include a disclaimer indicating that the researchers take full responsibility for the outputs. For example, "the results presented in this paper are the work of the authors."

9.4 The Principal Researcher must send copies of results, publications and presentations to Police, with sufficient opportunity for material to be reviewed, prior to dissemination, so that Police is informed before any public interest is generated. Progress updates covering results and outputs can be requested by Police at any time (refer to clause 8.2).

9.5 Police may publish links to published work on its external website.

10 COPYRIGHT

10.1 Intellectual property and copyright of any Privileged Police Information provided by Police to the Researchers shall belong to Police.

10.2 The Principal Researcher will provide to Police electronic copies of all final journal articles, thesis, presentations or any other reports prepared for The Project. All such documents may be used and copied by Police for any internal purpose, including, without limitation, publication on internal websites with appropriate acknowledgements as to source.

11 TERMINATION

11.1 The Police Director of the Evidence Based Policing Centre may review The Project operating under this Agreement and present the review to a member of the Police Executive with the recommendation to terminate this Agreement.

11.2 If an Executive member of Police agrees with the recommendation then Police will advise the Principal Researcher of the review and seek their response.

11.3 If no resolution is achieved within 10 working days, the Director of the Evidence Based Policing Centre may exercise the recommendation and terminate this Agreement by giving written notice to the Principal Researcher.

11.4 The foreseeable circumstances in which this clause might be exercised are:

a. The Researcher fails to comply with reasonable timeframes.

b. The Researcher fails to exercise a professional standard of practice.

c. The project creates an unanticipated and/or unreasonable demand upon Police resource beyond that represented in or contemplated by the Research Project proposal.

d. The researcher engages in conduct which, in the opinion of Police, is or may be likely to injure the reputation or interests of Police or brings or potentially brings Police into disrepute.

e. The researcher commits a material breach of this Agreement.

f. Breach of confidentiality and security. Breaches that are deliberate or a result of a lack of due care may also affect future access requests.

11.5 The Researcher can terminate this Agreement for any reason and without cause, by giving the Police one month's written notice.

11.6 If this Agreement is terminated, Police is not obliged to provide any support to The Project or any further information to the Principal Researcher or any other Researcher in relation to this project.

12 NON-TERMINATION

12.1 The Principal Researcher agrees that the obligations in this Agreement to maintain the confidentiality of Privileged Police Information shall be continuous and, in particular, shall continue in force beyond:

(a) the termination of this Agreement;

(b) the completion of The Project;

(c) release of the Project Findings; and

(d) the end of the Principal Researcher's involvement in The Project.

13 AMENDMENTS

13.1 No amendment of this Agreement shall be effective unless it is in writing and signed by both parties.

14 DISPUTE RESOLUTION

14.1 Subject to Clause 11, the parties agree to try to resolve any dispute arising from this Agreement by mediation (using the LEADR New Zealand Inc. standard), before starting any arbitration or legal proceedings.

15 COMPLETE AGREEMENT

15.1 This Agreement constitutes the full and complete agreement between Police and the Principal Researcher for the Project and supersedes all previous agreements, representations and contracts in relation to the Project.

16 GENERAL

16.1 The conditions of this Agreement which are capable of having effect after this Agreement expires shall remain in full force and effect after this Agreement expires.

SIGNED by

.....
(signature)

.....
SIMON WILLIAMS

on behalf of **THE COMMISSIONER OF NEW ZEALAND POLICE**

Police National Headquarters
180 Molesworth Street
Wellington 6140

Director, Evidence Based Policing Centre, New Zealand Police

SIGNED by the

Principal Researcher

Darryl Forsyth Digitally signed by Darryl Forsyth DN: cn=Darryl Forsyth, c=NZ, o=Massey University, ou=School of Management, email=d.forsyth@massey.ac.nz Date: 2022.05.06 14:03:22 +12'00'

.....
(signature)

.....
DR DARRYL FORSYTH

2/45 Rangitoto Tce Milford, AK
.....
(address)

Senior Lecturer
.....
(Position)

SIGNED by Researcher [repeat as necessary]

David Brougham Digitally signed by David Brougham DN: cn=David Brougham , c=NZ, o=Massey, ou=Massey, email=d.brougham@massey.ac.nz Date: 2022.05.06 09:32:59 +12'00'

.....
(signature)

.....
ASSOCIATE PROFESSOR DAVID BROUGHAM

23 Eyre Road, PN 4472
.....
(address)

AP
.....
(Position)

SIGNED by Researcher [repeat as necessary]

DocuSigned by:
11B103162AC2431...

.....
(signature)

.....
JESSE MOWAT

11 Spring Valley Place
Albany, Auckland
.....
(address)

Researcher
.....
(Position)

SIGNED by a representative of the Principal Researcher's Organisation, [*where necessary*]

This may be a Manager or CE and depends on the policy of the organisation that allows the principal to enter into this agreement.

..... (signature)

DR VIV SMITH
..... (name in block letters)

..... (address)

Director Research Operations | Research and Enterprise
..... (Position in Organisation)