A Validation Study of the Leadership Behaviour Assessment

A Dissertation Submitted in Partial Fulfilment of the Requirements for the Degree of

Master of Science

In Applied Psychology

At the University of Canterbury

By Michael Heslop

2021

Research Supervisors

Associate Professor Christopher Burt, University of Canterbury

Professor Katharina Näswell, University of Canterbury

Table of Contents

List of Tables

List of Figures

Acknowledgements

I want to acknowledge the time and effort of each of the participants, without whom this research would not be possible. Thank you to Simon and the NZDF for their willingness to enable research within the organisation. An enormous thank you to Jo of the Organisational Research Committee whose guidance, support and lightning fast responses got me through some of the most difficult periods of the dissertation. Thank you to Chris whose supervision, guidance and constant optimism kept the project on track despite the many pitfalls of data collection during a pandemic.

But most of all to Katie. This piece of work is built entirely on your love and support, without which I would be lost. You are my island in the tumultuous sea of life. This work is dedicated to you.

Abstract

The purpose of this study was to develop the Leadership Behaviour Assessment (LBA) and examine its criterion related validity. To achieve this, the researcher gathered data from 27 managers within the New Zealand Defence Force (NZDF) and 46 of their subordinates. The managers completed the LBA, while their subordinates rated their leader's overall leadership effectiveness using criterion scales chosen to measure 10 leadership behaviours. The data from the LBA and the criterion measure scales were correlated and three of the covert measures of the LBA had correlations approaching significance. These measures were *total time* spent completing the LBA, *average time* spent making a decision and number of *sections attempted*. These corelations suggest the LBA has potential as an innovative measure for leader selection that mitigates social desirability bias by using non-transparent measures. Further research needs to be conducted to explore other covert measures within the LBA, further examine its criterion related validity, and develop a profiling method for the tool.

A Validation Study of the Leadership Behaviour Assessment

Overview

The introduction section of this dissertation includes a literature review which explains the justification for the study and outlines its intended goal. It begins by explaining the importance of selecting the right people for leadership positions within an organisation and the organisational outcomes associated with effective or ineffective leadership. The current methods used by organisations to select personnel for leadership positions are discussed and their various disadvantages were examined in detail. It was made clear that one of the key disadvantages of current leadership selection methods is their vulnerability to social desirability bias. This was a key factor in the justification for this study because the assessment methods that mitigate this key bias are often very resource intensive and thus, impractical for some organisations and roles. It is explained that this dissertation is part of a wider research programme to design a selection tool that attempts to minimise the impact of social desirability bias, while being less resource intensive than other leadership assessment options. The Leadership Behaviour Assessment is the outcome of this process and the theoretical background behind its design is explained with reference to the literature that influenced its design. Finally, the introduction explains the importance of criterion related validity for any selection tool and how this study answered the research question: "Does the LBA have criterion related validity?".

The importance of leadership selection

One of the most crucial factors for the success of any organisation is the selection of personnel for leadership and management positions (Carnes, Houghton, & Ellison, 2015; Lashway, 1998). Successful leadership behaviour by managers within an organisation can be linked to positive organisational outcomes including reduced turnover intention, increased employee engagement and increased employee commitment (Mendes & Stander, 2011; Lee, 2005). In addition, specific aspects

of leadership behaviour in managers can predict the likelihood of a team reaching specific organisational goals (Alimo-Metcalfe, Alban-Metcalfe, Bradley, Mariathasan, & Samele, 2008). Leadership behaviours are directly linked to team performance which in turn impacts the overall performance of an organisation. The selection of the right leaders is crucial for an organisation because unlike other selection processes that only impact individual performance, leadership selection impacts the leader's performance as well as the performance of each of their followers (Carnes, et. al. 2015). In this way an error made in the selection of a leader can be multiplied by the number of followers under their influence. It is unsurprising then that most large organisations allocate a significant portion of resources to leadership selection and development (Saari, Johnson, McLaughlin, & Zimmerle, 1988). To maximise efficiency, it is crucial to ensure organisations invest these resources into selection techniques with high validity. Concerningly though, according to Howard, (2001) the most popular methods used to select CEOs are often the ones most lacking in rigour.

Current measures of leadership for personnel selection.

According to Howard (2007), the techniques currently used to select leaders can be broadly grouped into three categories: those that make inferences about behaviour, those that provide descriptions of behaviour and those that provide demonstrations of behaviour. The techniques in each category tend to have similar disadvantages which can detract from their utility as a selection tool. The types of assessments in each category with examples can be found in Table 1.

Table 1

Leadership Selection Techniques

Type of test	Explanation/example
Inferences about behaviour	
Cognitive tests	Assessments that measure *g, (*eg: Wonderlic test)
Situational judgment assessments	Tests of decision making/judgement in work settings
Personality inventories	Assess traits and characteristic (eg: 16PF)
Integrity tests	Use facets of conscientiousness and emotional stability
Leadership potential inventories	Measure leadership characteristics/potential (eg: MLQ)
Motivational fit	Assesses job fit, organisation fit and location fit
Projective techniques	Presented with ambiguous stimuli and fill in blank
Descriptions of behaviour	
Career achievement records	Past achievements related to job competencies
Reference checks	Evaluation of job history through key contacts
Interviews	Both competency-based, structured interviews and unstructured interviews.
Biographical data	Measures of past events/behaviours to predict performance
Demonstrations of behaviour	
Administrative simulations	Individual problem solving through simulated tasks (eg: In trays, analysis exercises)
Interactive simulations	Interaction with others in a job specific scenario (eg: role plays, presentations, Group discussion)

Note. Adapted from The Practice of Leadership: Developing the Next Generation of Leaders, by Conger, J.A. and Riggio, R.E. (Eds), Jossey-Bass, San Francisco, CA., p 18 Copyright (2007). Reprinted with permission.

Assessments that make inferences about behaviour cannot always be directly translated into performance outcomes (Howard, 2007). This is particularly true for personality inventories which rely on translating personality profiles into behaviour and using that to predict outcomes. Cognitive assessments that measure general mental ability (GMA or *g*) have some of the highest predicative validity for complex jobs (Schmidt & Hunter, 1996). However, research on GMA as a predictor of leadership performance is mixed, as a leader with high GMA can only directly influence group performance if the leader exhibits a directive style of leadership (Fiedler, 1996). In fact, research has found low or negative correlations between leader intelligence and team performance when the leader was non-directive or not supported by the group (Fiedler and Garcia, 1987, p.136). Other studies show

that leader intelligence does not improve group performance when there is a high level of interpersonal stress and uncertainty in the role (Fiedler, 1996).

Another common selection method that makes inferences about behaviour, is leadership potential inventories which often use self-report scales (Carless, Wearing, & Mann, 2000; Howard, 2007). A major disadvantage of psychometrics that use self-report measures is a vulnerability to the common bias of social desirability (Lilienfeld, Alliger, & Mitchell, 1995). Social desirability bias is the tendency by respondents to select answers that may be perceived as more positive in order to be compared more favourably or to avoid criticism (Paulhus, 1986). According to Paulhus (1991, Paulhus & Reid 1991) social desirability consists of two separate factors, impression management and self-deception. Impression management bias is defined as the deliberate tendency to select positive responses to present a positive self-image to others (Paulhus, 1986; Lajunen & Summala, 2003). Self-deception bias is defined as providing false responses due to an unrealistically positive assessment of one's abilities which is inconsistent with their actual behaviour (Paulhus, 1986, 1991). Applicant selection processes are known to elicit high levels of impression management (Lönnqvist, Paunonen, Nissinen, Ortju & Verkasalo, 2011) so, most research is concerned with reducing the impact of impression management (Hunsley, Vito, Pinsent, James, & Lefebvre, 1996); however, in personnel selection self-deception can have an equally confounding impact on the goal of selecting effective leaders.

According to Nederhof (1985), social desirability and other biases explain between 10% and 75% of the variance in participants' responses on self-report measures. These biases can have a confounding impact on the relationships between variables by either concealing variable relationships (King & Brunner, 2000) or producing non-existent relationships which undermine valid interpretation (Podaskoff, MacKenzie & Lee, 2003). For example, Bass and Avolio (2000), found in a study of military leaders that "…relying only on self-ratings of leadership is simply inaccurate to predict performance…" (p. 85). In the high stakes area of leadership assessment measurement mistakes can

have disproportionate impact on team performance, therefore it is crucial that the impact of social desirability bias is minimised.

Leadership selection methods that use descriptions of behaviour such as structured interviews, can also be vulnerable to impression management (Howard, 2007). In a study on military cadets, Lönnqvist, et. al. (2011) found that cadets with higher tendencies for self-enhancement through impression management had a higher probability of being selected into a military officer training programme. Lönnqvist et. al. (2011) hypothesised that the cadets' self enhancement tendencies had a persuasive effect on the interviewing psychologist, (even though they had been trained to minimise bias). In another study, Carnes et. al. (2015) found that the personality variables of the applicant accounted for over 18% of the variance in interview score, further demonstrating the potential impact of self enhancement on leadership selection.

In addition, descriptions of behaviour such as reference checks, behavioural type structured interviews and biographical data, often rely on leader experience which has been found to have a mixed relationship with team performance outcomes (Fiedler, 1996). Fiedler (1996), states that under stressful conditions more experienced leaders tend to perform better but under non-stressful conditions inexperienced leaders often perform better. This is demonstrated by a study of fire department officers by Frost (1981), which found experience and performance was negatively correlated in low-stress administrative work, but positively correlated with performance in stressful fire-combat. This can be explained by the more experienced officers being more likely to take short cuts during routine administration but using their experience and intuition to make decisions in fire combat.

To mitigate the vulnerability of social desirability bias, organisations can use selection methods that require demonstrations of actual behaviour such as in-trays, work sample tests (administrative simulations), presentations, group exercises or role plays (interactive simulations) (Howard, 2007; Lashway, 1998). Simulations of this type are often associated with assessment centres which have a long track record of high predictive validity (Howard, 1997). Assessment centres and simulations in

general require a significant resource investment in terms of personnel, time and money which is a major disadvantage that can be limiting for some organisations. This dissertation is part of a wider research programme on the design of a selection tool (the LBA) that attempts to minimise the impact of social desirability bias while being less resource intensive than other leadership assessment options.

Reducing the impact of social desirability

For social desirability to impact the validity of a measure, the following three conditions must be met. (1) The applicant must be motivated to portray themselves in a positive light; (2) the measure must have a socially desirable response; and (3) this response must be obvious to the participant (Furnham, 1986; Villanova & Bernardin, 1991). While it is impossible to reduce the motivation of the applicant to present themselves positively in the context of personnel selection; measures can be taken to reduce the transparency of the measure and thus reduce its susceptibility to social desirability bias (Furnham, 1986). A method of reducing the transparency of an assessment is to measure behaviour through indirect sources that are not obvious to the applicant. This reduces the impact of social desirability bias as the applicant is less likely to understand the method of assessment and adapt their responses to present themselves in a more positive manner (Reiners & Wood, 2014). We expected to achieve this through the LBA by using a web-based assessment that covertly measures leadership behaviours in a simulated management environment. The web-based nature of the assessment allows the LBA to measure the behaviours of an applicant as they interact with the LBA in ways that are not immediately obvious to the applicant. For example, a data log from the LBA can record how long an applicant spends in each menu before making a selection and whether an applicant changes their answer before submitting it. These data may then be able to be used to provide a measure of leadership effectiveness in a less transparent manner.

Assessments that use a simulated work task have other benefits that improve the authenticity of a measure, making them ideal for personnel selection. Simulations allow applicants behaviour to be measured in authentic environments and this is especially true for web-based simulations that can

immerse the applicant in a virtually simulated environment (Howard 2007, Clarke-Midura, 2010). By creating a replication of a real-life situation, a simulated measure can give an immersive experience that provokes a more authentic performance from the applicant (Clarke-Midura, 2010). Additionally, according to Ajzen, (1991) past behaviours are the best predictors of future behaviour, so creating a realistic environment requires applicants to draw on their previous experience, providing an accurate prediction of future behaviour.

As a Web-based simulation, the LBA also provides more opportunities to measure data than traditional self-report measures (Jaffal & Wloka, 2015). By recording data logs of the applicant's actions during the assessment, more measurement items are available, providing a better understanding of the applicant's performance (Jaffal & Wloka, 2015). This wealth of authentic data can give a greater insight into how the behaviour is applied to simulated leadership situations. By contrast, self-report measures have comparatively few data points from which to classify an applicant's leadership ability, narrowing the scope of their method of assessment. The LBA measures more authentic behaviour and as a result may be a more accurate predictor of future behaviour than self-report assessments alone.

Users of self-report measures tend to accept the impact of social desirability bias in a compromise for efficiency, requiring far fewer resources to conduct compared to simulations (Howard, 2007). Recent developments in the area of selection; however, have seen the use of web-based simulation assessments, which significantly reduce the time, money and personnel required to conduct an assessment (Howard, 2007). As a result of these advantages, web-based simulation assessments are the most efficient way to mitigate the disadvantages of self-report assessments in the high stakes area of leadership selection.

Assessing a range of leadership behaviours.

Research suggests that different leadership behaviours can be more or less desirable depending on context or culture of an organisation (Garg, & Jain, 2013; Ogawa & Bossert, 1995). Successful

leadership behaviours can even differ at each management level within the organisation, with higher levels involving greater scope, complexity and ambiguity (Howard, 2007). Instruments that consider a wide range of leadership behaviours can provide a better measure of overall leadership effectiveness.

In a review of leadership research Cohn and Moran (2011) conclude that relationship between executive leadership style and success is mixed at best. What seems to be more important than having a certain leadership style is being able to adapt one's style to meet differing demands and context (Cohn & Moran, 2011). Ogbonna and Harris (2000), explain that leadership effectiveness is dependent on the managers analysis of situational factors, followed by the implementation of an appropriate style to meet the situation. This is easier said than done as research by Bititci, Mendibil, Nudurupati, Turner, and Garengo, (2004) suggests managers find it difficult to adapt their leadership styles without external stimuli forcing them into it. This research further reinforces the importance of measuring various leadership behaviours when designing the LBA.

Blake and Mouton (1964) were one of the first researchers to categorise leadership behaviours into task or relationship-orientation in their Managerial Grid Model. This model was developed over the decades and is very useful for broadly categorising leadership behaviours. Relationship-oriented behaviours focus on the quality of the relationship with followers while task-oriented behaviours focus on the task to be accomplished by the followers (Blake & Mouton, 1982; Fiedler, 1996).

Various research has examined the relationship between task or relationship-oriented behaviours and performance with mixed results. Yammarino, Spangler, and Bass (1993) conducted a study on 276 US Naval Officers and found that relationship oriented-leadership behaviour was a better predictor of military performance than task-oriented leadership behaviour. Blake and Mouton (1964) in a study of 716 managers reported that the managers who displayed a combination of relationship-oriented and task-oriented behaviours advanced more quickly in their careers than managers with other styles. Klimoski and Hayes (1980) in research conducted on 241 personnel in a production department

of a large information processing firm, concluded that performance, job satisfaction and effort were higher under managers that demonstrated a mix of task and relationship-oriented behaviours.

These mixed research results demonstrate the importance of applying a combination of task and relationship-oriented leadership behaviours and the correct balance between these behaviours is often dependant on the context a leader finds themselves in (Garg, & Jain, 2013). Therefore, a tool that provides awareness of a leader's task and relationship-oriented behaviours is valuable for leadership selection. The LBA was designed to simulate these types of leadership behaviours.

Leadership Behaviour Assessment (LBA) Development

The LBA was conceptualised by Associate Professor Chris Burt who developed the initial idea for the assessment. These ideas were operationalised into the LBA by Chris and two master's students, Michael Heslop (the author) and Jessica Lord. To determine the leadership behaviours to be included in the LBA a literature review of leadership models was conducted. Behaviours were selected due to their frequency of occurrence across the different leadership models and their potential for measurement in the LBA. Table 2 shows the leadership models that were examined, and the most common behaviours listed in these models.

Table 2

List of Leadership Behaviours Mentioned in Leadership Models

	Full Range Leadership Model [a]	Authentic Leadership Model [b]	Leadership Performance Model [c]	Humble Leadership [d]	High Impact Leadership Model [e]	Well Rounded Leader Model [f]	Transitional Leadership Model [g]
Shares Knowledge		✓	✓		✓	✓	✓
Delegates	✓		✓	✓			✓
Decisive							✓
Feedback Orientation	✓	✓		✓			
Supports Innovation	✓		✓	✓	✓		
Supports Mentoring	✓		✓		✓		✓
Collaborates			✓	✓	✓	✓	✓
Just	✓	✓					
Develops Trust	✓	✓		✓		✓	
Inclusive			✓			✓	✓

Note. [a] *Bass & Avolio, 1994;* [b] *Avolio & Gardner 2005* [c] *Turor Versal, 2020;* [d] *Anderson, 2018;* [e] *Linkage, 2003;* [f] *Leadership Mastery, 2012;* [g] *Lotus leadership institute, 2011.*

The behaviour types chosen for inclusion in the LBA were knowledge sharing, delegation, decisiveness, feedback orientation, support for innovation, support for mentoring, collaboration, justness, development of trust and inclusiveness. In addition, activities in the LBA were designed to have either a task or relationship focus, which would allow the LBA to simulate the task/relationship focus of the applicant. The construction of the LBA is discussed in further detail in the method section. But in brief, the LBA was designed to simulate an online management system (OMS) which presents the applicant with a number of different choices that require their input as a manager. All activity conducted by the participant in the OMS is recorded in the data log, including the navigation of the menus and the time spent in each menu. The information recorded in this data log provides the basis of the covert measures that the LBA assesses in a non-transparent manner.

The covert measures recorded are as follows. *Total time* is defined as the amount of time spent completing the LBA in minutes (up to a maximum of 20 when the assessment ends). *Actions taken* is defined as the number of times a participant selected a response to one of the activities with the LBA. *Response changes* is defined as the number of times an applicant changed the response option they had selected. *Sections attempted* is the number of sections that an applicant participant selected at least one response option in (out of a maximum of seven sections). *Sections opened* is defined as the number of times a section was opened (may be greater than seven if the same section was opened multiple times). *Opened but no response* is defined as the number of activities within the seven sections that the participant initiated by clicking on the activity but not selecting a response despite being given options. *Time to decide* is defined as the average time in seconds that elapsed between the initiation of an activity by clicking on it and a response being selected by the participant (this was only available for two of the seven sections).

Validation

Before the LBA can be used in personnel selection, it must be proven as a valid tool for predicting future leadership behaviour and to achieve this its criterion related validity must be established. Criterion related validity measures how well a measure predicts an outcome. An assessment is considered to have criterion related validity if it can be used as an accurate predictor of future behaviour or outcomes (Schmidt & Hunter, 1980). According to the American Psychological Association (APA) (1999) for a psychometric tool to be considered fit for purpose, it must accurately predict the construct it claims to; thus, demonstrating criterion related validity. Criterion related validity can be established using two methods, by establishing predictive validity or concurrent validity (Cascio, & Aguinis, 2011). Predictive validity relates to the assessments ability to predict the result of future performance and usually involves measuring the predictor data first and subsequently collecting the criterion data. Concurrent validity is when the predictor data and the criterion data are collected at the same time (Cascio, & Aguinis, 2011). Both instances involve determining the strength of the relationship (correlation) between the predictor measure and the criterion data with the strength of the relationship determining the criterion related validity of the measure. For the LBA to demonstrate criterion related validity it should have significant and adequately strong correlations ($r > 0.3$, $p < 0.05$) with the criterion measure.

To establish criterion related validity for the LBA a concurrent validity study was conducted with subordinate rating data used as the criterion measure. Subordinates were asked to independently rate the leadership behaviours of their manager on ten scales selected as criterion measures for the ten behaviour types used to design the LBA. While it could be tempting to use team performance as a criterion measure, performance often has too many uncontrollable confounding variables (such as team member competence, workplace culture and employee motivation levels) to be a truly accurate measure of leadership ability (Howard, 2001; Oyinlade, 2006). Another option for a criterion measure is supervisor rating data; however, there are many factors that can confound the accuracy of supervisor

performance ratings. For example, supervisors may not have the time to provide comprehensive performance appraisals; supervisors may not be in a position to interact with a subordinate regularly so may be unfamiliar with the leadership behaviours they exhibit; or they may be uncomfortable with conflict and give neutral ratings to avoid the confrontation associated with negative ratings (Hughes, Ginnett & Curphy, 1999).

Subordinate ratings of leadership behaviour are the best option for a criterion measure for the LBA despite there being some disadvantages to using this measure. For example, subordinates may be reluctant to give a leader a negative rating from fear of potential backlash from their supervisor (Oyinlade, 2006). Also, if a leader is able to choose the subordinates that provide the ratings, they may be inclined to choose the ones that they have the best relationship with, resulting in an overly positive rating of their leadership behaviour. These disadvantages can be mitigated by asking participants to include all their direct report subordinates to participate in the subordinate rating questionnaire and to emphasise the anonymity of the survey to the subordinates before they conduct it. These mitigation measures were taken to improve the validity of the subordinate rating data used in this study.

Atwater and Yammarinol (1993) argue that leadership is about the influence on subordinates and therefore the impressions of subordinates are the most important measure of a leader's ability. Additionally, subordinates are best placed to directly observe a superior's leadership behaviour at its most authentic, compared to supervisors who are likely to only see a specific set of behaviours, carefully curated to present a positive image to those in charge of their performance appraisal (Atwater & Yammarinol, 1993). Subordinate rating data has been shown to be a much more accurate predictor of the performance of a leadership team and overall team performance than self-report data or manager rating (Bass & Avolio, 2000). For all of these reasons subordinate rating data of leadership competencies were used as the criterion measures for the validation study.

Current study

The aim of the current study was to establish criterion related validity for the LBA, so it can be used as a selection tool for leadership behaviours in an organisational context. Thus, the following research question was investigated: Does the LBA have criterion-related validity? To establish criterion related validity, each of the covert measures were correlated with the individual scale scores and with an overall measure of leadership effectiveness developed by summing all ten of the subordinate rating scales. It was predicted that each of the covert measures would have significant and adequately strong correlations with the leadership effectiveness rating scores.

Method

Design

This study used a concurrent criterion-related validation design within a single organisation, which was the New Zealand Defence Force (NZDF). To receive permission to conduct the research within the NZDF an extensive application process was followed that involved demonstrating the potential utility of the LBA to the organisation, finding an individual in the senior levels of the organisation to act as the sponsor of the research and having the research approved by the NZDF Organisational Research Committee. This process will be explained in more detail in a later section.

There are two types of participants in this study; *Applicant participants*, who completed the LBA and *subordinate participants* who completed the Manager Behaviour Questionnaire (MBQ) (the MBQ is listed in full in Appendix A). Data from the LBA (the predictor variable) was correlated with the data from the MBQ (the criterion variable) and these analyses were used to answer the research question: does the LBA have criterion-related validity? The study was reviewed and approved by the University of Canterbury Human Ethics Committee (Reference number HEC 2019/10/BL).

New Zealand Defence Force Organisational Approval Process

Any research conducted using members of the NZDF must be approved by the NZDF Organisational Research Committee before it can proceed. The NZDF Organisational Research Committee meets regularly to evaluate applications for research within the organisation. The Committee evaluates potential research on a number of factors, including the utility of the research to the NZDF, the ethical standard of the research and the level of risk the research presents to the organisation or its personnel. Before an application is made to the committee a sponsor must be found within the NZDF that will support the application. The sponsor must be of the equivalent rank of Colonel or higher (general manager level) and be willing to vouch for the usefulness of the research to the NZDF. Three different personnel were approached to sponsor the research and after extensive

discussion and consideration the head of organisational development within the NZDF offered to sponsor the research.

The initial research plan was adjusted to ensure the maximum utility of the research to the NZDF. It was agreed that once the LBA had been validated and the scoring system had been developed, a leadership report would be given to the applicant participants based on their interactions with the Leadership Behaviour Assessment. The LBA report could be used by the applicant participants for personal development. The method was adjusted to ensure that applicant participants were made aware that their subordinates could not be ordered, instructed of otherwise pressured into participating in the research. Additionally, instead of asking the applicants to nominate a certain number of subordinates to participate in the research, the committee asked that every one of the applicant participants' direct report subordinates should be approached to participate. This was done to increase the confidentiality of the research, making it less transparent to the applicant participant which subordinates had participated in the study. After an extensive application process, approval for the research was granted by the NZDF Organisational Research Committee (Organisational Research Committee Minute 500/PB/5/3) in October 2020. Once the research was approved a list of contact details for 706 suitable personnel was given to the researcher and this was used to approach personnel to participate in the research as applicant participants.

Participants

A total of 706 applicant participants and 94 subordinate participants were approached via email to participate in the study. After two months of data collection and two contact attempts 27 applicant participants and 46 subordinate participants had provided enough data to participate in the study (giving a response rate of 3.8% for applicant participants and response rate of 49% for subordinate participants). The desired number of applicant participants was calculated by a-priori power analysis using G*power (correlation; effect size = .55; α error probability = .05; power = .95) which gave a required sample size of 33. The final number of applicant participants was lower than this required

amount and is therefore a limitation of the study which will addressed in the discussion section. Applicant participants were recruited from the NZDF using a haphazard sampling where the most available personnel are studied (Weisberg & Bowen, 1977). The recruitment criteria for applicant participants included; being a member of the NZDF, in full time employment and with at least one direct report subordinate. Subordinate Participants were also recruited from the NZDF through haphazard sampling (Weisberg & Bowen, 1977), with each applicant participant being asked to provide the researcher with a list of all their direct report subordinates' email addresses (see Appendix E). The recruitment criteria for any subordinate participant included being a subordinate of one of the managers participating in the study for at least 12 months and having interacted with the manager at least once per week in the course of their work. Table 3 shows the basic demographic breakdown of the participants.

Table 3

Demographic Information of Participants

	Applicant Participants *n* = 27	Subordinate Participants *n* = 46
Males	20 (74%)	34 (74%)
Females	7 (26%)	12 (26%)
Mean age	41 (*SD* 8)	39 (*SD* 10)
Age range	30-57	23-59
	Ethnicity	
Pakeha only	24	34
Pakeha and Maori	1	8
Pakeha and Pacific peoples		1
Asian	1	2
Other	1	1

Note: *SD = standard deviation*

Sampling procedure

Applicant Participants. Initially, a list of 216 NZDF personnel who had completed a specific leadership course within the last four years was provided by the NZDF. Each of these individuals were contacted by email and asked to participate in the research. A low response rate after three weeks led

the researcher to ask the NZDF for another list of personnel. The two rank brackets with the highest rate of response from the previous sample were identified and a list was provided by the NZDF of the email address of every person in the NZDF at those two ranks. This resulted in a list of 492 individuals (after removing duplicates from the previous sample). Overall a total of 708 personnel were sent an email, asking them to participate in the research (see Appendix E) which included an information sheet (Appendix B) and a link to the electronic format of the consent form (see Appendix D). Only 27 applicant participants completed all the research requirements giving a response rate of less than 4%. Applicant participants were told they could complete the LBA during work hours (as stipulated by the NZDF organisational research committee) and would receive a report on the leadership behaviours exhibited in their interactions with the LBA.

Subordinate Participants. Each of the subordinate participants was sent an email asking them to participate in the study (see Appendix E) with an information sheet, a link to the electronic version of the consent form (see Appendix D) and a link to the MBQ. Subordinate participants received no incentive for their participation but were informed they could complete the requirements of the study during work hours (as stipulated by the NZDF organisational research committee).

All participants in the initial sample were given three weeks to complete the LBA and MBQ and were informed that any data received after this period may not be included in the study. Due to the impending deadline of the Christmas holiday period, the second group were only given two weeks to complete the LBA and MBQ with a reminder email being sent at the one-week mark. Once this period had elapsed the data was examined for completeness. Applicant participants were removed from the study if they had not fully completed the LBA (six participants) or if none of their subordinates had completed the MBQ on their behaviours (one participant). Applicant participants that were removed from the study received an email from the researcher with the following message in accordance with the direction of the NZDF Organisational Research Committee:

"Thank you for participating in the LBA validation study. Unfortunately, some of data we received was incomplete at the research deadline and therefore your responses have been removed from the study in accordance with the requirements of the research design. As a result of this we will be unable to provide you with a report on the leadership behaviours demonstrated in the LBA. We understand this can be disheartening but would ask that you do not raise this issue with any subordinates or any other participant of the study. If you have any questions or concerns about this decision, please contact the researchers or the Chair of the University of Canterbury Human Ethics Committee (contact details below)."

Materials

Applicant participants were given the Leadership Behaviour Assessment (LBA) to complete and subordinate participants were given the Manager Behaviour Questionnaire (MBQ). Both materials include a section asking for basic demographic information such as gender, ethnicity and age.

Manager Behaviour Questionnaire

The Manager Behaviour Questionnaire (MBQ) is an 87-item questionnaire comprised of eleven scales that were selected and adapted by the three people that were involved with the creation of the LBA. These scales were selected to be the criterion measures for each of the ten behaviour types considered when designing the LBA. Some items were adjusted from a self-report format to a subordinate report format, eg: "I share with others useful work experience and know-how" became "My manager shares with others useful work experience and know-how". A list of the adjusted items used in the MBQ is included in Appendix A.

Subordinate participants were asked to rate how well each item describes their superior's behaviour on an eight-point Likert scale with the following labels: strongly disagree (1), disagree (2), mostly disagree (3), slightly disagree (4), slightly agree (5), mostly agree (6) agree (7), strongly agree (8). During the data analysis it was discovered that a seven-point Likert scale had unintentionally been

used for the extent of collaboration scale, with the following labels: strongly disagree (1), disagree (2), somewhat disagree (3), neither agree nor disagree (4), somewhat agree (5), agree (6), strongly agree (7). Some items required reverse coding, and they are denoted by the indicator "(R)" next to the item in Appendix A and the paragraphs that follow. The scoring of the reversed items were adjusted during the data analysis with one becoming eight, two becoming seven etc.

The MBQ was developed into a Qualtrics survey along with questions on basic demographic details. The first question subordinate participants were asked was "what is the research code of your manager" (this was included in the email with the link inviting them to participate in the survey). This question was included to link the subordinate participants MBQ data to the LBA data completed by the applicant participant that nominated them for the study. The research code was randomly generated by the LBA and was allocated to each applicant participant and their respective subordinate participants by the researcher.

Manager Behaviour Questionnaire scales

The scales used in the MBQ are as follows:

The Knowledge Sharing Behaviour Scale (KSBS) (Lu, Leung & Koch, 2006) is a seven-item scale designed to measure an individual's knowledge sharing behaviour. An example item is "My manager keeps their work experience and never shares it out with others easily. (R)". This scale was included in the MBQ to measure the shares knowledge behaviour. An acceptable coefficient alpha of .87 was found for the scale during this research.

The Shares Knowledge Scale (SKS) is a two-item subscale of the Leader Empowering Behaviour Questionnaire (LEBQ) (Konczak, Stelly & Trusty, 2000) designed to measure empowering behaviour of leaders. An example item of the SKS is My manager shares information that I need to ensure high quality results". This scale was included in the MBQ to measure the shares knowledge behaviour. The two items from this scale were combined with the KSBS to provide an overall measure

of knowledge sharing behaviour. An acceptable coefficient alpha of .87 was found for the scale during this research.

The Delegation of Authority Scale (DAS) is a three-item subscale of the Leader Empowering Behaviour Questionnaire (LEBQ) (Konczak, Stelly & Trusty, 2000) designed to measure empowering behaviour of leaders. It was included in the MBQ to measure delegation. An example item is "My manager gives me the authority to make changes necessary to improve things". An acceptable coefficient alpha of .87 was found for the DAS during this research.

The Supervisor Decisiveness Scale (SDS) (Germeijs and De Boeck, 2002) is a 14-item scale designed to measure a supervisor's decisiveness. It was included in the MBQ to measure decisiveness. An example item is "My manager does not hesitate to make a decision". An acceptable coefficient alpha of .82 was found for the SDS during this research.

Feedback Environment Scale (FES) (Steelman, Levy & Snell, 2004) is a 19-item scale designed to measure a supervisor's feedback behaviours. It was included in the MBQ to measure feedback orientation. An example item is "I seldom receive praise from my manager. (R)". An acceptable coefficient alpha of .91 was found for the FES during this research.

The Innovative Behaviour Measure (IBM) (Scott & Bruce, 1994) is a 15-item scale designed to measure an individual's innovative behaviours. It was included in the MBQ to measure the supports innovation behaviour. An example item is "My manager promotes and champions ideas to others". An acceptable coefficient alpha of .90 was found for the IBM during this research.

The Perceived Managerial Support for Mentoring Scale (PMSMS) (Eby, Lockwood & Butts, 2006) is a six-item scale designed to measure a supervisor's support for mentoring. It was included in the MBQ to measure mentoring support behaviour. An example item is "My manager promotes mentoring opportunities". An acceptable coefficient alpha of .89 was found for the PMSMS during this research.

The Leader Inclusiveness Scale (LIS) (Nembhard & Edmondson, 2006) is a three-item scale designed to measure leader inclusiveness in teams of physicians. It was included in the MBQ to measure collaboration. An example item is “My manager asks for the input of other team members with different expertise”. An acceptable coefficient alpha of .71 was found for the LIS during this research.

Formal Procedures Scale (FPS) is a six-item subscale of the Formal Justice Scale (Niehoff, & Moorman, 1993) designed to measure perceptions of procedural justice. It was included in the MBQ to measure justness. An example item is “Job decisions are made by my manager in an unbiased manner”. An acceptable coefficient alpha of .93 was found for the FPS during this research.

Faith in Intentions Scale (FIS) (Cook, & Wall, 1980) is a five-item scale designed to measure perceptions of interpersonal trust at work. It was included in the MBQ to measure the develops trust behaviour. An example item is “I feel quite confident that my manager will always treat me fairly”. An acceptable coefficient alpha of .94 was found for the FIS during this research.

Extent of Collaboration Scale (ECS) (Greenwald, & Zukoski, 2018) is an eight-item scale designed to measure extent of collaboration within an organisation. It was included in the MBQ to measure collaboration. An example item is “My manager and the team share knowledge that promotes work progress”. During the data analysis it was discovered that a seven-point Likert scale had unintentionally been used for the ECS in the MBQ, ranging from strongly disagree (1) to strongly agree (7). It was the only scale that used this Likert scale. An acceptable coefficient alpha of .90 was found for the ECS during this research.

Leadership Behaviour Assessment

The Leadership Behaviour Assessment (LBA) is a tool designed to measure leadership behaviour for the purposes of personnel selection. The LBA requires an internet enabled computer and can be completed from anywhere in the world with internet access, via a link sent to the participant.

This dissertation does not describe the LBA in detail because the detailed workings and scoring of the assessment must be kept confidential to ensure its effectiveness as a selection tool, but a general description of its design follows. The supervisor letter sent to the marker of this dissertation also includes a one-time link to the assessment which can be used by the marker.

The LBA presents the applicant with an Online Management System (OMS) which simulates various activities that a manager may be expected to carry out in the course of their duties. The OMS presents the applicant with seven sections based around common themes and each containing a different number of activities for the participant to complete. The participant is free to navigate and complete each of these sections and their activities in any order. The participant is given 20 minutes to complete the OMS after which point the assessment will close and prevent any further answers. The time limit was determined by pilot testing as sufficient. The 20-minute time limit was chosen to ensure that the simulation included a degree of time pressure simulating workplace conditions, and to ensure a degree of efficiency in the data collection. The options that the participant prioritises and the actions that they take within the OMS are associated with each of the ten leadership behaviour types simulated by the LBA. These behaviour types are knowledge sharing, delegation, decisiveness, feedback orientation, support for innovation, support for mentoring, collaboration, justness, development of trust and inclusiveness.

The seven sections of the LBA were designed to present participants with information in different ways to allow a wide range of leadership behaviours to be observed. Two of the sections provide the participant with information about an incident or event and asks them to make a decision on how to resolve it. Each decision is restricted to four different options that require varying levels of investment from the participant, from doing nothing to scheduling time to resolve the issue themselves. Two other sections present the participant with information that is both positive and negative and ask the participant to determine either how widely the info is shared amongst their organisation or to prioritise which information is shared at all. Two sections invite the participant to schedule activities

for themselves or members of their team and ask them to prioritise which activities to focus over a limited period of time. The final section asks the participant to determine which personnel from different areas of an organisation should be involved in decision making. The activities within each section were designed by the researchers to cover a broad range of the ten leadership behaviour types. When designing each of the activities within the seven sections of the LBA the researchers mapped each activity onto one or more of the ten leadership behaviour types based on the researchers understanding of the behaviours. This was done to ensure there was an even spread of behaviour types across each of the sections of the LBA and to ensure each behaviour was linked to multiple activities.

The LBA was designed to measure leadership behaviours in two different ways based on what can be recorded by the web-based design of the LBA. These are classified by the researchers as overt and covert measurements. Overt measurements refer to the information that can be gained from the options participants choose when presented with information within the LBA. The LBA is designed to reduce transparency, but because the participant is aware of the choices, they are making within OMS it is possible for them to attribute (rightly or wrongly) behaviours to those decisions and select the most socially desirable answer. Covert measures refer to the data gathered by the LBA in a less transparent manner, so the participant is unlikely to be aware that their behaviour is being measured. Covert measures included in the LBA were defined in the LBA development section of the introduction.

Once the LBA was complete it was tested by 10 participants for bugs and to optimise the user interface. Participants were sent a unique link to the LBA by email or social media and were asked to complete it. Participants were debriefed by the researcher after they had completed the LBA and asked several planned questions including “were there any terms used in the LBA that you did not understand?” and “did you notice any errors in the program?”.

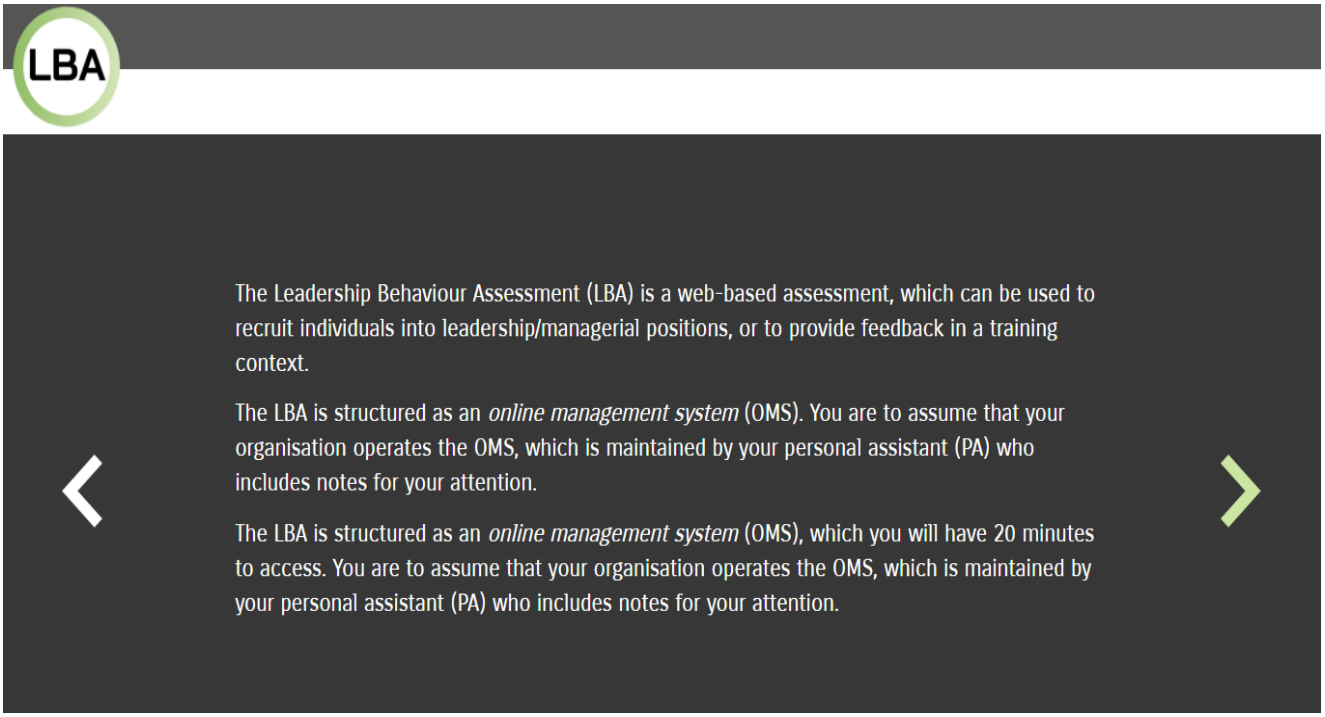
Procedure

After approval was received from the NZDF Organisational Research Committee, an email was sent to 706 personnel inviting them to participate in the study (see Appendix E), which included an information sheet and a link to an electronic version of the consent form hosted on Qualtrics (see Appendix D). Personnel that agreed to participate in the study, replied to the email with a list of the personnel numbers of all their direct report subordinates and their own personal email. The personnel numbers allowed the researcher to email the subordinates directly without requiring any further information. The personal email of the applicant participant was required because the link to the LBA could not be emailed to an organisational email address (due to unforeseen compatibility issues). Participants with incomplete data or other research requirements were sent an email (see Appendix H) at the one-week and two-week mark of the study to remind them of the research deadline.

Applicant participants procedure. Applicant participants were sent an email with a unique link to the LBA which asked them to ensure they had 20 minutes of uninterrupted time to complete it and made them aware that the link would only work once (see Appendix F). They were asked to complete the LBA on a personal device such as a laptop or tablet but to avoid using a smart phone if possible, as the LBA was not optimised for smart phone use. Applicant participants were allocated a randomised research code consisting of three letters and three numbers. This research code was used to identify the applicant participants in all data and no names or identifying information was used during the data analysis stage. On opening the LBA link received from the researcher, participants are asked to confirm the randomly assigned research code on the LBA matches the code given to them by the researcher in the accompanying email. They are then presented with the introduction shown below in Figure 1 which is a screenshot taken from the LBA.

Figure 1.

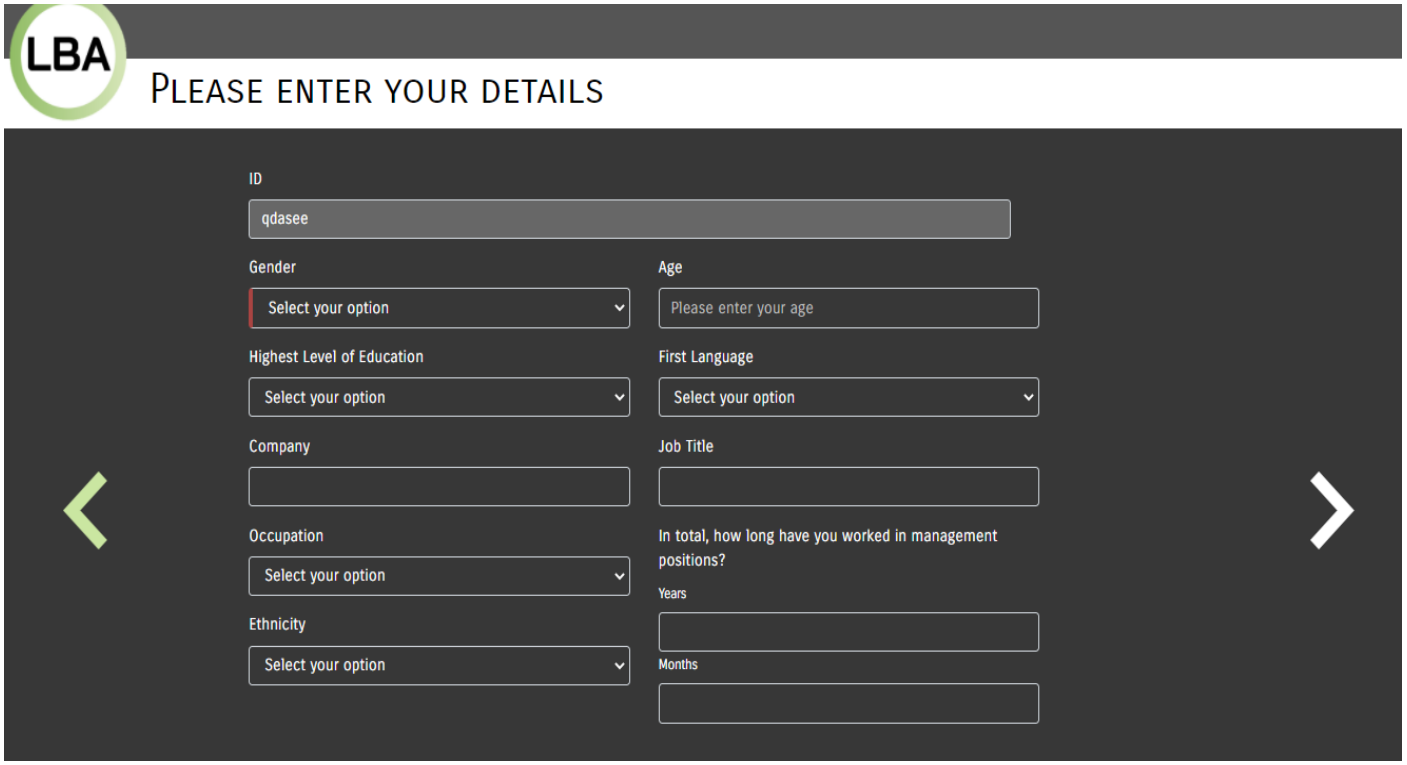
LBA Introduction



Participants were then asked to complete a data protection agreement, and a candidate information and release form asking them to comply with the conditions of the assessment before being allowed to continue. These conditions included understanding the length of time allowed by the assessment; ensuring cell phones are turned off; ensuring they have access to reading glasses or contacts if required; and that their data will be stored in a collective data base for statistical purposes. Participants were then asked to complete basic demographic questions which are shown in Figure 2 which is a screenshot taken from the LBA.

Figure 2.

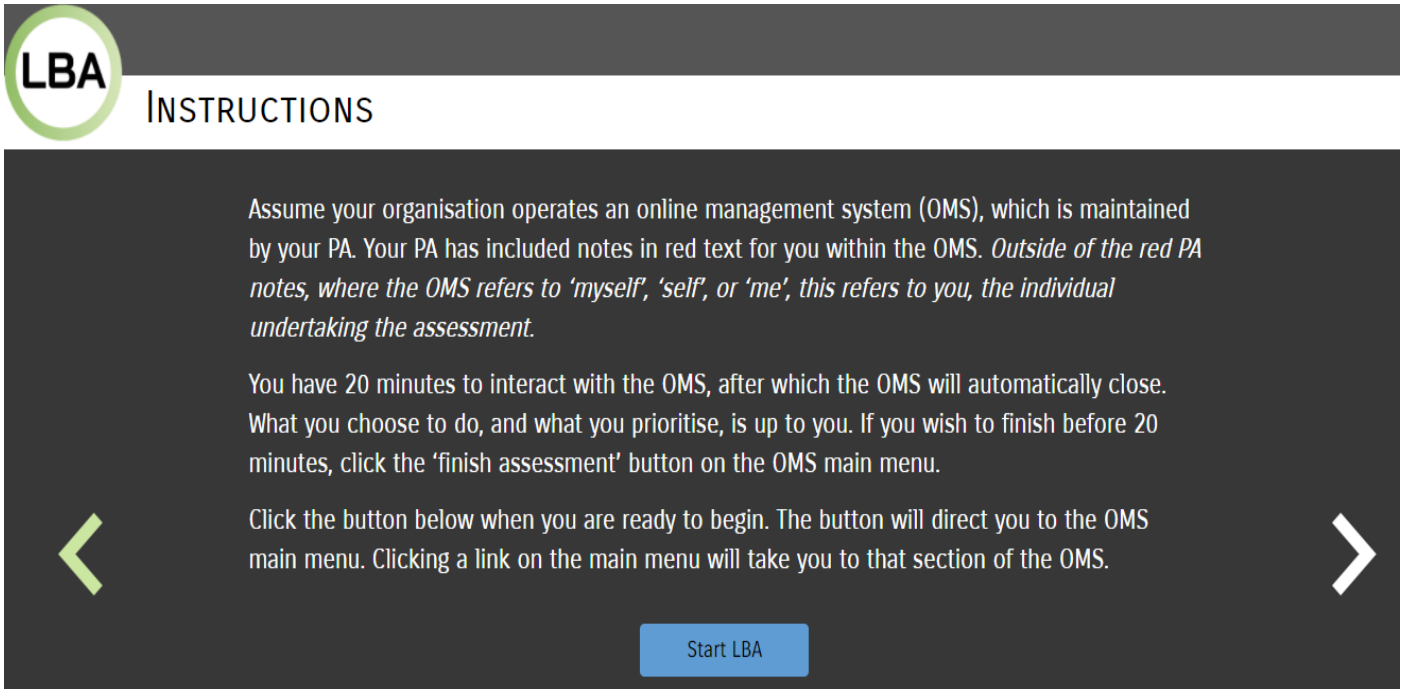
LBA Demographic Questions



Participants are then given the final set of instructions before being given the option to start the LBA. Instructions are shown below in Figure 3 which is a screenshot taken from the LBA.

Figure 3.

LBA Instructions



After clicking “start LBA” participants were presented with the main menu which has a list of the seven sections, randomly ordered. Each section represents a task that a manager might be expected to carry out in the course of their duties. When the participant clicked on one of these options they were taken to another page where red text simulating a note from a personal assistant asked them to make decisions based on the information presented to them on the page. Participants were given a number of different options with each decision and are allowed to change their decisions at any point. Along with the actual option selected by the candidate, the LBA recorded the order in which the options were selected, the time spent in each section, and whether decisions were changed at any point. All of this information is recorded in the data log which is used to provide a measure of overall leadership effectiveness.

Subordinate participants procedure. The personnel numbers of direct report subordinates provided by the applicant participants were collated into a list of 94 personnel. These 94 personnel were sent a were sent an email inviting them to participate in the study (see Appendix G) which included an information sheet and a link to an electronic version of the consent form hosted on Qualtrics (see Appendix D). The email included an anonymous link to the MBQ which was hosted on Qualtrics. 46 subordinate participants responded to this email and completed both the consent form and the MBQ. On opening the MBQ, participants were asked to identify their manager in the Qualtrics survey by their unique research code which was also included in the email below the MBQ link (see Appendix G). They then completed some basic demographic questions before completing the 88 item MBQ. The research codes provided by the subordinate participants were used to match their responses to the LBA data from the applicant participants.

Results

Data Analysis

All data analysis was performed using IBM Statistical Package for the Social Sciences (SPSS) version 26. The data from the MBQ and LBA were downloaded into separate excel spreadsheets and converted into SPSS data files.

Accounting for the Varied Number of Subordinates per Applicant Participant

One stipulation of the NZDF Organisational Research Committee was that each applicant participant must invite all their subordinates to participate in the study rather than a specific subordinate or a designated number of subordinates. This meant that there was a large disparity in the number of subordinate participants that completed the MBQ for each applicant participant. The number of subordinate participants that successfully completed the MBQ for every applicant participant ranged from 1 to 4 with a median of 1. Table 4 shows the frequency distribution of subordinate participants per applicant participant.

Table 4

Frequency Table of Subordinate Participants per Applicant Participant

Number of Subordinate Participants *	Frequency
One	15
Two	7
Three	3
Four	2

Note: *Subordinate participants that completed the MBQ

Before examining the criterion related validity of the LBA the data were adjusted to account for the difference between the numbers of subordinate participants for each applicant participant. The mean MBQ scale results for each applicant participant with more than one subordinate participant were not used as this could result in inaccurate representations of the criterion scores. To solve the

issue one subordinate participant was randomly selected to use in the data analysis for applicant participants that had data from more than one subordinate participant. This was achieved by generating a random value for each subordinate participant and using the subordinate participant with the highest randomly generated value.

Data Investigation

The MBQ data were investigated for missing cases and seven were found. In each item there were only a maximum of two missing cases representing less than 10% of the total responses for each item. Therefore, missing cases were replaced with the item mean, and Table 5 describes the missing data and the mean it was replaced with.

Table 5

Manager Behaviour Questionnaire Missing Values

Scale name	Item #	Missing values	Item means
The Leader Inclusiveness Scale	2	1	6.38
The Perceived Managerial Support for Mentoring Scale	1	1	6.73
	2	1	6.38
	3	1	6.54
	4	1	5.77
	5	1	6.38
The Knowledge Sharing Behaviour Scale	5	2	5.36

Note: N = 26

The LBA data does not require a *missing values analysis* as completion of all aspects the LBA is not forced, nor expected. However, the LBA data were inspected for outliers by calculating the number of sections attempted by each applicant participants out of the total of seven. A section was defined as attempted if the applicant participant had completed at least one action within the section. The median number of sections attempted by applicant participants was 7, indicating most applicant participants had sufficient time to attempt all the sections. One participant was found to have completed only 2 sections which is greater than 3 standard deviations lower than the mean (< 2.39).

To confirm that this applicant participant was an outlier the time they spent in the main menu (which has no activities and only lists the 7 sections) was calculated as a percentage of the total time they spent conducting the LBA. It was found that this participant spent 51% of their time in the main menu compared to the participant who spent the next highest percent of their time in the main menu at 6%. It was therefore determined that this applicant participants data were an outlier, and it is likely they were interrupted while conducting the LBA resulting in excessive time spent in the main menu. Therefore, their data were removed from the analysis leaving a total *N* of 26.

Overall leadership effectiveness variable

To establish criterion related validity for the covert measures of the LBA, a composite score was created from the sum of all the MBQ scales. Leadership research (Cohn & Moran, 2011; Ogbonna & Harris, 2000) indicates that demonstrating a wide range of leadership behaviours and adapting them to suit the situation is crucial for effective leaders. Additionally, as mentioned in the introduction and in Table 2, the wide range of leadership behaviours that the LBA was designed to measure are part of many well-established leadership models. A leader that is rated highly across all behaviours measured by the ten MBQ scales by their subordinates is likely to be a more effective leader. Therefore, to form an overall leadership effectiveness variable the subordinate participants score on each of the MBQ scales was summed. The overall leadership effectiveness variable/score, along with the individual MBQ scale scores were then correlated with the covert measures of the LBA to determine which covert variables have criterion related validity for measuring overall leadership effectiveness, and specific leadership characteristics. Before conducting this analysis, data range issues were examined.

Manager Behaviour Questionnaire Distribution, Range Restriction and Reliability

The criterion validity analysis of the LBA is determined by correlating the LBA results with the criterion measures (MBQ). A critical assumption of correlation analysis is that the data being correlated have normal distributions (Hunter, Schmidt, & Le, 2006), therefore it is important to

investigate the distributions of both the LBA and the MBQ data. The distribution, range, skew, and kurtosis were examined for the MBQ, with the results shown in Table 6. Inspection of Table 6 indicates that all the scales show a ceiling effect with most of their values clustered towards the higher end of the response rating range. According to Kim (2013), for a data set with a sample size of 50 or less to be considered normally distributed, it should have an absolute Z value of less than 1.96 or greater than -1.96. Absolute Z values are calculated by dividing the skewness or kurtosis value by their standard error. Absolute Z values for both Skewness and Kurtosis are shown in Table 6. Only the Innovative Behaviour Measure (IBM) and the Supervisor Decisiveness Scale (SDS) meet the requirements for being considered to have a normal distribution, every other scale including the MBQ overall leadership effectiveness score, demonstrates a high level of skewness or kurtosis. To some extent it is undesirable that the criterion measures deviate from a normal distribution. The skewed data from the MBQ will likely mean that subsequent analysis will underestimate the correlations between the LBA and the criterion variables when establishing criterion-related validity evidence (Hunter, Schmidt, & Le, 2006).

Table 6

Manager Behaviour Questionnaire Scale Descriptive Statistics and Distributions

			Range				Absolute Z Value	
Scale	Mean	*SD*	Min	Max	Skew	Kurtosis	Skew	Kurtosis
The Innovative Behaviour Measure	6.35	1.01	4.07	7.93	-0.80	-0.03	-1.76	-0.03
The Delegation of Authority Scale	6.72	1.39	2.33	8.00	-1.91	3.48	-4.20	3.93
The Supervisor Decisiveness Scale	6.50	0.71	5.14	7.79	-0.06	-0.62	-0.14	-0.70
The Leader Inclusiveness Scale	6.63	1.20	3.67	8.00	-1.17	0.82	-2.56	0.93
The Perceived Managerial Support for Mentoring Scale	6.30	1.06	3.33	8.00	-1.03	1.49	-2.26	1.68
The Knowledge Sharing Behaviour Scale	6.65	0.91	3.89	7.78	-1.51	2.40	-3.32	2.70
Feedback Environment Scale	6.74	0.79	4.21	7.95	-1.32	2.94	-2.90	3.32
Faith in Intentions Scale	7.25	0.86	4.60	8.00	-1.68	2.90	-3.68	3.27
Formal Procedures Scale	6.81	0.83	4.83	8.00	-0.98	0.61	-2.16	0.68
Extent of Collaboration Scale	5.94	0.69	3.38	7.00	-2.00	7.25	-4.38	8.18
Overall Leadership Effectiveness Score	65.89	8.00	42.58	76.96	-1.25	1.94	-2.72	2.18

Note: N = 26, Scale descriptive statistics are based on mean scale scores of each participant. Possible ranges for all scales is 1- 8, except the Extent of Collaboration Scale, which is 1-7 and the MBQ summed scale score which is 10-79.

Reliability analysis was conducted on each scale used in the MBQ. A Cronbach's alpha value for each scale was calculated and these are included in the method section of this report.

Relationships Between Scales in the Manager Behaviour Questionnaire

Each of the scales in the MBQ were specifically chosen to measure an aspect of leadership behaviour deemed important for effective leadership by multiple leadership models. Table 7 shows the correlations between each of the MBQ scales. Each of the scales had positive, significant correlations with every other scale and all but one of these correlations were at a high level (r >.50). It is unsurprising that these scale correlate so highly with each other as managers with years of training and experience are likely to display many of the behaviours measured by the scales.

Table 7

Corelation Matrix of MBQ Scales and Overall Leadership Effectiveness Score

		1	2	3	4	5	6	7	8	9	10
1	The Innovative Behaviour Measure										
2	The Delegation of Authority Scale	.68**									
3	The Supervisor Decisiveness Scale	.65**	.56**								
4	The Leader Inclusiveness Scale	.81**	.50**	.50**							
5	The Perceived Managerial Support for Mentoring Scale	.79**	.66**	.72**	.72**						
6	The Knowledge Sharing Behaviour Scale	.76**	.48*	.60**	.77**	.68**					
7	Feedback Environment Scale	.77**	.57**	.77**	.72**	.84**	.71**				
8	Faith in Intentions Scale	.70**	.78**	.56**	.58**	.74**	.65**	.70**			
9	Formal Procedures Scale	.87**	.72**	.77**	.68**	.86**	.66**	.80**	.75**		
10	Extent of Collaboration Scale	.58**	.55**	.64**	.58**	.70**	.72**	.76**	.70**	.72**	
11	Overall leadership effectiveness score	.91**	.79**	.78**	.82**	.91**	.82**	.89**	.85**	.92**	.78**

*Note: N = 26, * = p < .05, ** = p < .01*

LBA Scoring Development

There are two possible avenues to developing a scoring method for the LBA. As noted above, it may be possible to develop a set of scores representing overt responses by combining selected options, decisions made, and activities undertaken in the LBA however, there are issues with this. Firstly, overt responses are somewhat open to bias responding. The applicant may make a decision

based on what they think is the expected decision (social desirability, impression management bias). Second the LBA is complex and knowing what to combine into overt scores is not an easy task (especially with low response rates). Therefore, at this early stage in the LBA development the focus was on the covert measures. The covert measures are simple recordings by the LBA system and are not subject to direct participant bias. Thus, the covert measures of the LBA are used to investigate criterion related validity by correlating them with the overall leadership effectiveness variable and the individual MBQ scale scores. As previously discussed, the covert measures are less transparent and thus more difficult, if not impossible for the applicant participant to manipulate. This makes the covert measures more objective and less vulnerable to bias.

Leadership Behaviour Assessment Distribution and Range Restriction

The distribution, range, skew and kurtosis were examined for the LBA covert measures with the results shown in Table 8. According to Kim (2013), for data set with a sample size of 50 or less to be considered normal, it should have an absolute Z value of less than 1.96 or greater than -1.96. Absolute Z values are calculated by dividing the skewness or kurtosis value by their standard error. Absolute Z values for both Skewness and Kurtosis are shown in Table 8. Only the *actions taken,* and the *sections opened* measures meet the requirements for being considered a normal distribution, every other measure demonstrates a high level of skewness or kurtosis. This is an indication of range restriction and is present in all but the aforementioned two measures (Hunter, Schmidt, & Le, 2006). It is undesirable for any of the data to deviate from a normal distribution as correlations require normally distributed data. The range restriction and skewed data from the LBA will likely mean that subsequent analysis will underestimate the correlations between the LBA and the criterion variables (Hunter, Schmidt, & Le, 2006).

Table 8

Leadership Behaviour Assessment Descriptive Statistics and Distributions

			Range				Absolute Z Value	
Scale	Mean	*SD*	Min	Max	Skew	Kurtosis	Skew	Kurtosis
Total time (minutes)	19.04	1.84	13.64	20.00	-1.84	2.28	-4.04	2.57
Actions taken	94.88	22.85	54.00	147.00	.41	0.27	.89	.30
Response changes	5.42	4.57	0	19.00	1.53	2.27	3.35	2.56
Sections attempted	6.42	1.03	3	7	-1.94	3.74	-4.26	4.22
Sections opened	8.42	2.48	3	14	.46	.46	1.00	.52
Opened but no response	1.31	1.76	0	5.00	1.20	0.12	2.64	.14
Mean time to decide (seconds)	11.35	4.70	2.20	30.03	2.36	10.19	5.18	30.03

Note: N = 26,

Criterion-Related Validation Analysis of the LBA

To investigate the criterion related validity of the LBA each of the covert measures (*total time taken, Actions taken, response changes, sections attempted, sections opened, sections opened but no response, mean time to decide*) were correlated with the overall leadership effectiveness variable and the individual MBQ scale scores. The correlations with the individual MBQ scales are shown in Table 9 and the corelations with the overall leadership effectiveness score are shown in Table 10.

Table 9

Correlations Between MBQ Scales and Covert Measures

	Total time (minutes)	Actions taken	Response changes	Sections attempted	Sections opened	Opened but no response	Mean time to decide (seconds)
The Innovative Behaviour Measure	0.34	-0.17	-0.05	-0.33	-0.06	0.06	0.14
The Delegation of Authority Scale	0.07	-0.05	0.02	-0.18	-0.19	-0.16	0.18
The Supervisor Decisiveness Scale	0.23	0.09	0.09	-0.11	0.01	0.23	0.34
The Leader Inclusiveness Scale	0.18	-0.07	0.10	-0.31	-0.15	-0.04	0.19
The Perceived Managerial Support for Mentoring Scale	0.21	-0.02	-0.10	-0.16	-0.21	-0.05	0.12
The Knowledge Sharing Behaviour Scale	0.26	0.12	0.19	-0.17	-0.01	0.15	0.28
Feedback Environment Scale	0.33	-0.04	-0.05	-0.25	-0.02	0.19	0.26
Faith in Intentions Scale	0.20	0.02	0.00	-0.17	-0.29	-0.18	0.34
Formal Procedures Scale	0.34	-0.04	-0.05	-0.21	-0.16	-0.06	0.28
Extent of Collaboration Scale	0.24	-0.10	-0.22	-0.17	-0.13	0.06	0.21

Note: N = 26, none of the correlations were significant at $p < .05$.

Table 10

Correlations Between Covert Measures and Overall Leadership Effectiveness Score

Covert Measure	*r*	*p*
Total time (minutes)	.27	.18
Actions taken	-.04	.86
Response changes	0	1.00
Sections attempted	-.25	.22
Sections opened	-.15	.46
Opened but no response	0	.99
Mean time to decide (seconds)	.27	.19

Note: N = 26

Inspection of Tables 9 and 10 shows that none of the covert measures had a significant correlation with the overall leadership effectiveness score or any of the individual MBQ scales. However, three of the covert measures have correlations that are approaching significant values. *Total time* had a positive correlation approaching significance with the overall leadership effectiveness variable, meaning applicant participants that spent more time completing the LBA were rated as more effective leaders by their subordinates. Inspection of Table 9 shows that *Total time* had positive correlations approaching significance with the individual MBQ scales chosen to measure innovation, feedback and trust. *Mean time to decide* had a positive correlation approaching significance with the overall leadership effectiveness variable, meaning applicant participants that spent more time making their decisions were rated as more effective leaders by their subordinates. Table 9 shows that *Mean time to decide* had positive correlations approaching significance with the individual MBQ scales chosen to measure decisiveness, and trust. *Sections attempted* had a negative correlation approaching significance with the overall leadership effectiveness variable, meaning applicant participants that attempted less sections were rated as more effective leaders by their subordinates. Table 9 shows that *Sections attempted* had negative correlations approaching significance with the individual MBQ scales chosen to measure innovation, and inclusiveness. Given the non-normal distributions of both the LBA and MBQ data indicating range restriction, it is likely that these correlations are underestimated. None of the other covert measures demonstrated significant correlations or correlations approaching significance with the overall leadership effectiveness scale or any of the individual MBQ scales.

Discussion

Study Summary

The aim of this research was to examine the criterion related validity for the Leadership Behaviour Assessment (LBA). Criterion related validity was investigated by examining the associations between the covert measures of the LBA and subordinate ratings of leadership related constructs. The LBA was devised using a leadership construct based on ten leadership behaviour types and designed to minimise, if not eliminate, the impact of social desirability bias in candidate assessment data. The validation process demonstrated the potential of the LBA as a tool for measuring leadership effectiveness for personnel selection in a manner that significantly reduces the impact of social desirability bias.

Summary of Findings

The initial investigation of the data involved examining the MBQ data for range restriction. Eight of the ten scales in the MBQ showed evidence of range restriction and consequently, so did the overall leadership effectiveness variable which was a sum of these scales. Each scale within the MBQ showed a ceiling effect with most of their values clustered towards the higher end of the response rating range. This was problematic for the correlation analysis, an assumption of which is that the variables being correlated have normal distributions. Therefore, the non-normal distributions of the criterion measure variables likely meant that the correlations between the criterion measures and the LBA covert measures were underestimated (Sackett, Lievens, Berry, & Landers, 2007).

For the LBA covert measures to demonstrate criterion relate validity they were expected to correlate significantly with the criterion measure of overall leadership effectiveness. The direction of the correlation (positive or negative) indicating which type of covert behaviours were indicative of leadership effectiveness. None of the covert behaviours had significant correlations with the criterion measures but three had correlations approaching significance. Given the non-normal distributions of

both the LBA and MBQ data, it is likely that the correlations between these three covert measures and the MBQ summed scale score are underestimated.

The three covert measures with correlations approaching significance were *total time* spent conducting the LBA (r = .27, p = .18), *mean time to decide* (r = .27, p = .19) and number of *sections attempted* (r = -.25, p = .22). These findings suggest that leaders who spent more time working through the activities in the LBA and gave their limited attention to fewer sections were also rated as more effective leaders by their subordinates. This conclusion is intuitive and paints the picture of conscientious and considered leaders being deemed more effective by their subordinates.

In addition to the overall leader effectiveness score the covert measures were correlated with each individual scale in the MBQ scales and the results, though not significant, revealed relationships that would be worth further exploration. *Total time* had positive correlations approaching significance with the individual MBQ scales measuring the leadership behaviours of innovation (r = .34, p = .09), feedback (r = .33, p = .10), and trust (r = .34, p = .09). *Mean time to decide* had positive correlations approaching significance with the individual MBQ scales chosen to measure decisiveness (r = .34, p = .09), and trust (r = .34, p = .09). *Sections attempted* had negative correlations approaching significance with the individual MBQ scales chosen to measure innovation (r = -.33, p = .10), and inclusiveness (r = -.31, p = .12). The intuitive relationships between the subordinate leadership behaviour ratings on particular scales and the behaviour demonstrated in the LBA suggest that the covert measures reveal authentic leadership behaviour related to real life outcomes judged by independent raters. For example, the data shows that leaders who spent longer on average making decisions in the LBA and more time on the LBA overall were rated higher by subordinates on the scale measuring trust in their leadership. This is intuitive in the sense that a leader who tends to demonstrate a more considered approach, taking in more information before making a decision would build more trust with their subordinates. The same can be said of all the other relationships between the covert measures and the individual MBQ scales mentioned above.

Explanation of Findings

The range restriction that is evident in the data is unsurprising given that the applicant participants have had years of leadership experience and training. They should be expected to demonstrate a range of the behaviours measured by the MBQ and the LBA. According to Sackett et. al. (2007), direct range restriction on one or more variables can cause an underestimation of the correlations between them. It is likely that range restriction was present in both the LBA and MBQ data was caused by the selection of leaders with years of experience and training. It is very difficult to avoid range restriction when using a sample of military leaders as most go through an extensive selection process for leadership abilities and all military leaders undergo extensive leadership training before being placed into leadership roles.

What may also be present in the MBQ data is a halo effect. According to Frone, Adams, Rice, and Instone-Noonan (1986), if a subordinate believes a leader to be generally effective, they can tend to apply this judgement to all leadership qualities resulting in positive ratings in all categories. The stronger the halo effect, the greater the intercorrelations between supposedly distinct leadership behaviours (Behrendt, Matz, & Göritz, 2017). Intercorrelations between the scales of the MBQ were almost entirely high (r >.50) which was unexpectedly strong for distinct measures. This suggests that the MBQ data demonstrated a halo effect, causing the MBQ data have less variation between the different scales.

Given the non-normal distributions of both the LBA and MBQ data indicating range restriction, it is likely that the correlations between these three covert measures and the overall leader effectiveness variable are underestimated. It is also highly likely that the correlations between the three covert measures and the individual MBQ scales were underestimated. It is likely then that the correlations approaching significance were demonstrating real relationships between the criterion measure and the LBA, but these effects were masked by range restriction and small sample size. It could be said therefore that the LBA covert measures are measuring authentic leadership behaviours and the LBA

has potential as a tool for leadership selection. Further research to explore the criterion related validity of the LBA is justified based on these preliminary results.

Practical Implications

A common criticism of psychology and other social sciences is the amount of bias and distortion in the data collected to make real world decisions. The LBA was designed to reduce this bias by using authentic measures of leadership behaviour to minimise the impact of social desirability bias. For social desirability to impact the validity of a measure, the following three conditions must be met. (1) The applicant must be motivated to portray themselves in a positive light; (2) the measure must have a socially desirable response; and (3) this response must be obvious to the participant (Furnham, 1986; Villanova & Bernardin, 1991). It is impossible to reduce the motivation of a participant to portray themselves in a positive light and impossible to remove a socially desirable response from a measure designed to be used in personnel selection. So, the LBA achieves the latter criteria by using indirect measures of leadership behaviour that are not obvious to the participant. This reduces the impact of social desirability bias as the applicant is less likely to understand the method of assessment and adapt their responses to present themselves in a more positive manner (Reiners & Wood, 2014).

Using covert measures to provide an objective measure of leadership effectiveness is desirable not only for minimising susceptibility to social desirability bias but also for removing any potential bias introduced through perceptions of leadership behaviour. According to Behrendt, Matz, and Göritz (2017), evidence suggests that perceptions of leadership behaviour differ from the actual behaviour itself. This introduces a bias into leadership assessment that uses subjective measures of leadership assessment, such as self-report, subordinate, and superior ratings. These types of measures are influenced by subjective interpretations of leadership behaviour by the rater, which makes them more vulnerable to biases such as the halo effect.

Having the right people in leadership positions is crucial for organisational success and consequently leadership selection is an area that organisations look to invest significant resources into (Carnes, Houghton, & Ellison, 2015). Concerningly though, the most popular methods of leadership selection used by organisations often lack rigour (Howard, 2001). As previously discussed, many common leadership selection tools depend on self-report measures that are vulnerable to social desirability bias. Other tools that don't rely on self-report measures tend to be resource intensive or rely on perceptions of leadership (such as subordinate or supervisor ratings) which are vulnerable to other biases introduce through the raters subjective views on leadership.

This study has demonstrated the potential of the LBA for measuring authentic leadership behaviours. The covert measures of the LBA are non-transparent, meaning the applicant is unable to determine the most socially desirable response and select it. The covert measures that had correlations approaching significance with the overall rating of leadership effectiveness show the potential of the LBA as a leadership selection tool. With further research, LBA has the potential to be an innovative, rigorous, and practical measure of leadership behaviours in the high stakes area of personnel selection.

Ethics of Covert measurement

Covert measurement has always been a controversial topic in psychology, often criticised on ethical grounds (Spicker, 2011; Lugosi, 2006). The main complaint of detractors is that covert research is akin to deception and thereby, intrinsically unethical, or at least unprofessional (Beauchamp, Faden, Wallace, & Walters, 1982; Herrera, 1999; Warwick, 1982). Much of this criticism is focused on covert research conducted on unwitting participants but, covert measures may be open to criticism due to their non-transparent nature. Participants immediately understand what an overt item is designed to measure, whereas covert measures are less transparent making it difficult if not impossible for the participant to understand how their behaviour is being measured. As previously discussed, there are inherent advantages of covert measures in assessments due to this distinction and many authors argue the advantages of covert measures outweigh their perceived disadvantages. A study by Vidotto,

Anselmi, Filipponi, Tommasi, and Saggino, (2018), found that the covert items in integrity assessments were far less vulnerable to faking by participants than overt measures. Other advocates for covert measures argue that they offer access to data that may otherwise be impossible to measure in a manner that is free from bias (Calvey, 2000; Lauder, 2003; Miller, 2001).

Spicker (2011) argues that "…disclosure is not a dichotomous concept…" and there can be varying different levels of transparency for research participants. This spectrum of disclosure should be considered when evaluating the ethical standing of an assessment and covert measures should not be considered unethical by default. In this sense, the LBA measures leadership behaviour in a covert manner that is towards the more open end of the spectrum. Any applicant asked to complete the LBA as part of personnel selection will be broadly aware of the purpose of the assessment in measuring leadership behaviour but unaware of exactly how it will be measured. For this reason, it should be argued that the benefits of bias reduction in the high-stakes arena of personnel selection, justify the lack of transparency of covert measures. The history of stigma associated with covert research and a culture of hypersensitivity surrounding research leads to an over cautious approach to covert measures (Lugosi, 2006), and any criticism of their use should consider the utility they provide for personnel selection.

Limitations

One limitation of the study is the use of only one type of criterion measure to validate the LBA. According to Behrendt, Matz, and Göritz (2017), evidence suggests that perceptions of leadership behaviour differ from the actual behaviour itself. This is part of the fundamental need for a more objective measure of leadership behaviour that the LBA is designed to fulfil. The subordinate rating data that was used to validate the LBA however is still vulnerable to these biases. This was observed in the data with an obvious halo effect. Subordinate report data is also vulnerable to social desirability bias, especially where the subordinate suspects their manager may get access to the ratings they provide. This limitation was mitigated by making it abundantly clear to the subordinate participants

that their manager would not have access to the ratings they provide, nor to any statistics derived from the subordinate rating data (see Appendix G). There are very few practical options for measuring leadership behaviour rather than perceived behaviour which is part of the reason for creating the LBA. Other practical options for a criterion measure included supervisor rating and self-report measures, however, these types of measures are also perceptions of leadership behaviour and are vulnerable to the same biases. Other, more objective criterion measures (such as data from work samples) could be used in future research to explore the relationships discovered in this research.

Another limitation of this research was the small sample size. An a-priori power analysis (correlation; effect size = .55; α error probability = .05; power = .95) gave a required sample size of 33. To account for incomplete or incorrect data, the initial research plan was to continue recruiting applicant participants until 40 had completed the LBA. Unfortunately, the low response rate meant that this plan could not be completed in the time allocated for data collection and consequently sample size of only 26 was achieved. At the time data collection was occurring the NZDF was involved in a significant operation to protect the country from COVID 19 by managing isolation facilities for people returning from overseas. According to Stuff, at the time of the research, one in ten NZDF personnel were involved in the security of managed isolation facilities and "…replacements will rotate in every four to six weeks" (Block, 2020). This extra burden combined with the more ordinary stresses of daily life during a pandemic, likely had an impact on the number of personnel willing to volunteer to be involved in research. In fact, two applicant participants asked to be removed from the research due to unexpected work commitments stemming from the pandemic response. Despite this small sample size, the results are still useful as they demonstrate relationships that would be worth investigating further. Additionally, samples sizes as low as 30 have been used to validate other well-established measures such as the Wonderlic Contemporary Cognitive Ability Test.

Another practical limitation of the research is the low-stakes environment that the assessment was conducted in. Low-stakes assessment environments that have few real-life outcomes for

performance tend, to increase the variation in the effort among assessment takers (Schüttpelz-Brauns, Hecht, Hardt, Karay, Zupanic, & Kämmer, 2019). The LBA is intended for use in the high stakes environment of leadership selection, in which the wrong decisions can impact organisational effectiveness on a wide scale. During this study, the participants conducted the LBA in a relatively low stakes environment, where poor performance had little practical consequences for the participant. This makes it much more likely for participants to be distracted or perform sub-optimally, thus potentially impacting the results. This was mitigated by LBA link email which asked the applicant participants to ensure they are in a quiet environment that is free from distraction and have at least 20 mins of uninterrupted time to conduct the assessment, (see Appendix F). Additionally, as the LBA collects a wide range of data on how much time a participant spends in different sections of the LBA it was possible to identify and remove outliers from unusual patterns of behaviour in the assessment. For example, a participant spent over 50% of their time in the main menu of the LBA and completed an unusually low number of sections and thus, was removed as an outlier.

Future Research

The LBA would benefit from further research to explore its utility as a tool for personnel selection of leaders. This could be done by evaluating the criterion related validity of the LBA against another measure such as supervisor ratings of leadership behaviours. According to Frone et. al. (1987), both subordinate and self-ratings of behaviour are prone to the halo effect, however supervisor ratings of behaviour are less prone to the effect. The halo effect reduced the variation in the criterion measure data in this study making it much more difficult to determine individual differences between the individual scales and the leadership behaviours they measured. This made it very difficult to investigate the scoring of the LBA effectively. Future research using criterion measures that are less prone to the halo effect and to range restriction (through a more diverse sample) will make it much easier to determine individual differences. Supervisor ratings of leadership behaviours would be an excellent candidate for a criterion measure in future research.

This study and any future research on the LBA will contribute to a body of data that can be used to determine the various applications of the tool. This study has shown that there is potential in the covert measures that the LBA assesses and further research examining their criterion related validity would be extremely useful for the future of the assessment. In particular, further research should examine what other covert data the LBA can capture and whether this data is a valid predictor of leadership behaviours. As more data is collected it will also be possible to develop a scoring method for the LBA and expectancy tables that allow more informed selection methods to be made. Developing this scoring method and establishing its criterion related validity should be one of the mains aims of future research into the tool.

Conclusion

Despite the limitations of this study, the results demonstrate that the LBA has potential as a tool for leadership selection. Further research must be done to derive a scoring interpretation method to allow decisions to be made, such as the development of expectancy tables and further validation work is needed. The use of non-transparent measures in a web-based simulation offers a unique and objective method for measuring leadership effectiveness that has potential to mitigate social desirability bias, the greatest limitation of self-report measures. Thus, the LBA has potential to provide organisations with a more objective measure to identify individual differences in leadership effectiveness. Selecting the right people for leadership positions is crucial to organisational success and the LBA has demonstrated potential to provide added utility to this vital process.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179-211.

Alimo-Metcalfe, B., Alban-Metcalfe, J., Bradley, M., Mariathasan, J., & Samele, C. (2008). The impact of engaging leadership on performance, attitudes to work and wellbeing at work. *Journal of Health Organization and Management, 22*(6), 586-598.

American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational, & Psychological Testing (US). (1999). Standards for educational and psychological testing. *American Educational Research Association.*

Anderson, J. (2018). *Humble leadership model.* Retrieved from https://janeandersonspeaks.com/6-powerful-traits-humble-leader/humble-leadership-model/

Atwater, L. E., & Yammarinol, F. J. (1993). Personal attributes as predictors of superiors' and subordinates' perceptions of military academy leadership. *Human Relations, 46*(5), 645-668.

Avolio, B.J., & Gardner, W.L. (2005). Authentic leadership development: getting to the root of positive forms of leadership, *The Leadership Quarterly*, *16*(3).

Bass, B. M., & Avolio, B. J. (2000). Platoon readiness as a function of leadership, platoon and company cultures. Binghampton, NY: Research Foundation of the State University of New York, Office of Research and Sponsored Programs.

Bass, B. M., & Avolio, B. J. (Eds.). (1994). Improving organizational effectiveness through transformational leadership. Sage Publications, Inc.

Beauchamp, T., Faden, R., Wallace, J., & Walters, L. (Eds.). (1982). Ethical issues in social scientific research. Baltimore: John Hopkins University Press.

Behrendt, P., Matz, S., & Göritz, A. S. (2017). An integrative model of leadership behavior. The Leadership Quarterly, 28(1), 229-244.

Bititci, U. S., Mendibil, K., Nudurupati, S., Turner, T., & Garengo, P. (2004). The interplay between performance measurement, organizational culture and management styles. Measuring Business Excellence.

Blake, R. R., & Mouton, J. S. (1964). *The managerial grid: Key orientations for achieving production through people.* Houston, Tx: Gulf Pub. Co.

Blake, R. R., & Mouton, J. S. (1982). Theory and research for developing a science of leadership. *The Journal of Applied Behavioral Science*, *18*(3), 275-291.

Block, G. (2020). Coronavirus: One in ten Defence Force personnel deployed in Covid-19 battle. *Stuff.* Retrieved from: https://www.stuff.co.nz/national/300088508/coronavirus-one-in-ten-defence-force-personnel-deployed-in-covid19-battle

Calvey, D. (2000). Getting on the door and staying there: A covert participant observational study of bouncers. In G. Lee-Treweek & S. Linkogle (Eds.), Danger in the field: Risks and ethics in social research (pp. 43-60). London: Routledge.

Caputo, Andrea. (2017). Social Desirability Bias in self-reported well-being Measures: Evidence from an online survey. *Universitas Psychologica, 16*(2), 245-255.

Carless, S. A., Wearing, A. J., & Mann, L. (2000). A short measure of transformational leadership. *Journal of Business and Psychology, 14*(3), 389-405.

Carnes, A., Houghton, J. D., & Ellison, C. N. (2015). What matters most in leader selection? The role of personality and implicit leadership theories. Leadership & Organization Development Journal.

Cascio, W. F., & Aguinis, H. (2011). Applied psychology in human resource management (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Clarke-Midura, J., & Dede, C. (2010). Assessment, technology, and change. *Journal of Research on Technology in Education, 42*(3), 309-328.

Cohn, J., & Moran, J. (2011). *Why are we bad at picking good leaders? A better way to evaluate leadership potential*. New York City, NY: John Wiley & Sons.

Cook, J., & Wall, T. (1980). New work attitude measures of trust, organizational commitment and personal need non-fulfilment. Journal of occupational psychology, 53(1), 39-52.

Eby, L. T., Lockwood, A. L., & Butts, M. (2006). Perceived support for mentoring: A multiple perspectives approach. *Journal of Vocational Behavior, 68*(2), 267–291

Fiedler, F. E. (1996). Research on leadership selection and training: One view of the future. *Administrative Science Quarterly, 41*(2), 241-250.

Fiedler, F. E., & Garcia, J. E. (1987). *New approaches to effective leadership: Cognitive resources and organizational performance.* New York City, NY: John Wiley & Sons.

Frone, M. R., Adams, J., Rice, R. W., & Instone-Noonan, D. (1986). Halo error: A field study comparison of self -and subordinate evaluations of leadership process and leader effectiveness. Personality & Social Psychology Bulletin, 12(4), 454-461.

Frost, D. E., (1981). *The effects of interpersonal stress on leadership effectiveness.* Unpublished PhD dissertation. Retrieved from: https://digital.lib.washington.edu.

Furnham, A. (1986). Response bias, social desirability and dissimulation. *Personality and Individual Differences, 7*(3), 385-400.

Garg, S., & Jain, S. (2013). Mapping leadership styles of public and private sector leaders using blake and mouton leadership model. *Drishtikon: A Management Journal, 4*(1), 48-64.

Germeijs, V., & De Boeck, P. (2002). A measurement scale for indecisiveness and its relationship to career indecision and other types of indecision. *European Journal of Psychological Assessment, 18*(2), 113–122.

Greenwald, H. P., & Zukoski, A. P. (2018). Assessing collaboration: Alternative measures and issues for evaluation. American Journal of Evaluation, 39(3), 322–335.

Guyll, M., Spoth, R., & Redmond, C. (2003). The effects of incentives and research requirements on participation rates for a community-based preventive intervention research study. *The Journal of Primary Prevention, 24*(1), 25-41.

Herrera, C. D. (1999). Two arguments for "covert research" in social research.British Journal of Sociology,50(2), 331-341.

Howard, A. (1997). A reassessment of assessment centers: Challenges for the 21st century. *Journal of Social Behavior and Personality, 12*(5), 13.

Howard, A. (2001). Identifying, assessing, and selecting senior leaders. In Zaccaro, S. J., Klimoski, R. J., & Ebrary, I. (Eds.), *The nature of organizational leadership: understanding the performance imperatives confronting today's leaders (1st ed.)*, (pp. 305-346). San Francisco: Jossey-Bass.

Howard, A. (2007). Best practices in leader selection, in Conger, J.A. and Riggio, R.E. (Eds), *The Practice of Leadership: Developing the Next Generation of Leaders*, San Francisco, CA: Jossey-Bass. pp. 11-40.

Hughes, R. L., Ginnett, R. C., & Curphy, G. J. (1999). *Leadership: Enhancing the lessons of experience.* Boston: McGraw-Hill.

Hunsley, J., Vito, D., Pinsent, C., James, S., & Lefebvre, M. (1996). Are self-report measures of dyadic relationships influenced by impression management biases? *Journal of Family Psychology, 10*(3), 322-330.

Hunter, J. E., Schmidt, F. L., & Le, H. (2006). Implications of direct and indirect range restriction for meta-analysis methods and findings. Journal of Applied Psychology, 91(3), 594.

Jaffal, Y., & Wloka, D. (2015). Employing game analytics techniques in the psychometric measurement of game-based assessments with dynamic content. *Journal of E-Learning and Knowledge Society, 11*(3).

Kim, H. Y. (2013). Statistical notes for clinical researchers: assessing normal distribution using skewness and kurtosis. *Restorative Dentistry & Endodontics, 38*(1), 52-54.

Klimoski, R. & Hayes, N. (1980). Leader behavior and subordinate motivation. *Personnel Psychology, 33.*

Konczak, L. J., Stelly, D. J., & Trusty, M. L. (2000). Defining and measuring empowering leader behaviors: Development of an upward feedback instrument. *Educational and Psychological Measurement, 60*(2), 301–313.

Korn, J. H., & Hogan, K. (1992). Effect of incentives and aversiveness of treatment on willingness to participate in research. *Teaching of Psychology, 19*(1), 21-24.

Lajunen, T., & Summala, H. (2003). Can we trust self-reports of driving? Effects of impression management on driver behaviour questionnaire responses. *Transportation Research Part F: Traffic Psychology and Behaviour, 6*(2), 97-107.

Lashway, L. (1998). Measuring Leadership. *Research Roundup, 14*(2).

Lauder, M. (2003). Covert participant observation: Justifying the use of deception.Journal of Contemporary Religion,18(2), 185-196.

Leadership Mastery (2012). *Well-rounded leader model.* Retrieved from https://www.leadershipmastery.com.au/leadership-model/

Lee, J. (2005). Effects of leadership and leader-member exchange on commitment. *Leadership & Organization Development Journal, 26*(8), 655-672.

Lilienfeld, S. O., Alliger, G., & Mitchell, K. (1995). Why integrity testing remains controversial. *American Psychologist, 50*(6), 457-458.

Linkage Consulting Research Team. (2003). *Linkage Assessment Instrument Technical Report*, retrieved from http//:www.Linkageinc.com/research_products/assessment_instruments.aspx.

Lotus Leadership Institute (2011). *Transitional leadership model.* Retrieved from http://www.lotusleadership.net/the-difference/transitional-leadership-model/

Lu, L., Leung, K., & Koch, P. T. (2006). Managerial knowledge sharing: The role of individual, interpersonal, and organizational factors. *Management and Organization Review, 2*(1), 15–41.

Lugosi, P. (2006). Between overt and covert research: Concealment and disclosure in an ethnographic study of commercial hospitality. Qualitative Inquiry, 12(3), 541-561.

Mendes, F., & Stander, M. W. (2011). Positive organisation: The role of leader behaviour in work engagement and retention. *SA Journal of Industrial Psychology, 37*(1), 1-13.

Miller, J. M. (2001). Covert participant observation: Reconsidering the least used method. In J. M. Miller & R. Tewksbury (Eds.), Extreme methods: Innovative approaches to social science research (pp. 13-21). Boston: Allyn & Bacon.

Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European journal of social psychology, 15*(3), 263-280.

Nembhard, I. M., & Edmondson, A. C. (2006). Making it safe: The effects of leader inclusiveness and professional status on psychological safety and improvement efforts in health care teams. *Journal of Organizational Behavior, 27*(7), 941–966.

Niehoff, B. P., & Moorman, R. H. (1993). Justice as a mediator of the relationship between methods of monitoring and organizational citizenship behavior. *Academy of Management Journal, 36*(3), 527–556.

Ogawa, R. T., & Bossert, S. T. (1995). Leadership as an organizational quality. *Educational Administration Quarterly, 31*(2), 224-243.

Ogbonna, E., & Harris, L. C. (2000). Leadership style, organizational culture and performance: Empirical evidence from UK companies. International Journal of Human Resource Management, 11(4).

Oyinlade, A. O. (2006). A method of assessing leadership effectiveness. *Performance Improvement Quarterly, 19*(1), 25-40.

Paulhus, D. L. (1986). Self-deception and impression management in test responses. In A. Angleitner & J. S. Wiggins (Eds.), *Personality assessment via questionnaires: Current issues in theory and measurement* (pp. 143-165). New York; Berlin: Springer-Verlag.

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), Measures of personality and social psychological attitudes: Vol 1 (pp. 17–59). San Diego, CA: Academic Press.

Paulhus, D. L., & Reid, D. B. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality and Social Psychology, 60*, 307–317.

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

Reiners, T., & Wood, L. C. (Eds.). (2014). *Gamification in education and business*. Springer.

Saari, L. M., Johnson, T. R., McLaughlin, S. D., & Zimmerle, D. M. (1988). A survey of management training and education practices in US companies. Personnel Psychology, 41(4), 731-743.

Sackett, P. R., Lievens, F., Berry, C. M., & Landers, R. N. (2007). A cautionary note on the effects of range restriction on predictor intercorrelations. Journal of Applied Psychology, 92(2), 538-544.

Schmidt, F. L., & Hunter, J. E. (1980). The future of criterion-related validity. *Personnel Psychology, 33*(1), 41-60.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. Psychological Bulletin, 124(2), 262-274.

Schüttpelz-Brauns, K., Hecht, M., Hardt, K., Karay, Y., Zupanic, M., & Kämmer, J. E. (2019). Institutional strategies related to test-taking behavior in low stakes assessment. Advances in Health Sciences Education, 1-15.

Scott, S. G., & Bruce, R. A. (1994). Determinants of innovative behavior: A path model of individual innovation in the workplace. *Academy of Management Journal, 37*(3), 580–607.

Scullen, S. E. (1997). When ratings from one source have been averaged, but ratings from another source have not: Problems and solutions. *Journal of Applied Psychology, 82*(6), 880-888.

Spicker, P. (2011). Ethical covert research. Sociology, 45(1), 118-133.

Steelman, L. A., Levy, P. E., & Snell, A. F. (2004). The Feedback Environment Scale: Construct Definition, Measurement, and Validation. *Educational and Psychological Measurement, 64*(1), 165–184.

Tutor Versal (2020). *Leadership performance model.* Retrieved from https://www.tutorversal.com/australia/performance-leadership-assignment-help.html.

Van de Mortel, T. F. (2008). Faking it: social desirability response bias in self-report research. *The Australian Journal of Advanced Nursing, 25*(4), 40.

Vidotto, G., Anselmi, P., Filipponi, L., Tommasi, M., & Saggino, A. (2018). Using overt and covert items in self-report personality tests: susceptibility to faking and identifiability of possible fakers. Frontiers in psychology, 9, 1100.

Villanova, P., & Bernardin, H. J. (1991). Performance appraisal: The means, motive, and opportunity to manage impressions. In R. A. Giacalone & P. Rosenfeld (Eds.), *Applied impression management: How image-making affects managerial decisions* (pp. 81-96). Beverly Hills, CA: Sage.

Warwick, D. P. (1982). Tearoom trade: Means and end in social research. In M. Bulmer (Ed.), Social research ethics: An examination of the merits of covert participant observation (pp. 38-58). London: Macmillan.

Weisberg, H. F., & Bowen, B. D. (1977). An introduction to survey research and data analysis. San Francisco: W.H. Freeman.

Yammarino, F., Spangler, W., & Bass, B. (1993). Transformational leadership and performance: A longitudinal investigation. *Leadership Quarterly, 4*, 81-102.

Appendix A

Manager Behaviour Questionnaire

Participants will be asked to rate how well each item describes their managers behaviour on an eight-point Likert scale ranging from strongly disagree (1) to strongly agree (8).

	Adapted Scale Items
Trust	
Faith in Intentions	My manager is sincere in their attempts to meet the workers' point of view.
	If I got into difficulties at work, I know my manager would try and help me out.
	I can trust my manager to lend me a hand if I needed it.
	I feel quite confident that my manager will always treat me fairly.
	My manager can be relied upon to do as they say they will do.
Decisive	
	My manager finds it easy to make decisions.
	It is hard for my manager to come to a decision. (R)
	My manager doesn't know how to make decisions. (R)
	My manager knows which steps to take when making a decision.
	I would characterize my manager as an indecisive person. (R)
	My manager doesn't hesitate much when they have to make a decision.
	It takes a long time for my manager to weigh the pros and cons before making a decision. (R)
	My manager makes decisions quickly.
	My manager delays deciding. (R)
	My manager doesn't postpone making decisions to a later date.
	My manager tries to avoid making a decision. (R)
	My manager tends to leave decisions to someone else. (R)
	Once my manager has taken a decision, they stick to that decision.
	My manager doesn't avoid situations where decisions have to be made.
Delegates	My manager gives me the authority I need to make decisions that improve work processes and procedures.
	My manager gives me the authority to make changes necessary to improve things.
	My manager delegates authority to me that is equal to the level of responsibility that I am assigned.
Shares Knowledge	My manager shares information that I need to ensure high quality results.
	My manager provides me with the information I need to meet customers' needs.
Supports Innovation	
	Creativity is encouraged by my manager.
	Our ability to function creatively is respected by my manager.
	My manager allows people to try to solve problems in different ways.
	The main function of members in this team is to follow orders which come down through my manager. (R)
	A person can get in a lot of trouble with my manager by being different. (R)
	My manager can be described as flexible and continually adapting to change.

	A person can't do things that are too different around here without provoking anger from my manager. (R)
	The best way to get along with my manager is to think the way the rest of the group does. (R)
	People around here are expected by my manager to deal with problems in the same way. (R)
	My manager is open and responsive to change.
	In this team, we tend to stick to tried and true ways. (R)
	My manager seems to be more concerned with the status quo than with change. (R)
	My manager's reward system encourages innovation.
	My manager publicly recognizes those who are innovative.
	My manager's reward system benefits mainly those who don't rock the boat. (R)
Shares Knowledge	
	In daily work, my manager takes the initiative to share their work-related knowledge with the team.
	My manager keeps their work experience and never shares it out with others easily. (R)
	My manager shares with others useful work experience and know-how.
	After learning new knowledge useful to work, my manager promotes it to let more people learn it.
	My manager never tells others their work expertise, unless it is required in the company. (R)
	My manager actively uses IT sources available in the company to share their knowledge.
	So long as others need it, my manager always tells whatever they know, without any hoarding.
Collaborates	
	My manager and the team provide each other with useful information that makes work progress.
	My manager and the team share knowledge that promotes work progress.
	My manager and the team understand each other when we talk about the work to be done.
	My manager and the team share resources that help perform tasks.
	My manager and the team communicate our ideas to each other about the work to be done.
	My manager and the team make progress reports.
	My manager and the team exchange information on 'who does what'.
	My manager and the team discuss work deadlines with each other.
Supports Mentoring	
	My manager serves as a role model for mentors.
	My manager encourages employees to be mentors.
	My manager promotes mentoring opportunities.
	There are few rewards available from my manager for mentoring others. (R)
	Mentors receive little recognition from my manager for their efforts. (R)
	Mentoring relationships are not reinforced by my manager. (R)
Provides Feedback	
Feedback Quality	My manager gives me useful feedback about my job performance.

	The performance feedback I receive from my manager is helpful.
	I value the feedback I receive from my manager.
	The feedback I receive from my manager helps me do my job.
	The performance information I receive from my manager is generally not very meaningful. (R)
Feedback Delivery	My manager is supportive when giving me feedback about my job performance.
	When my manager gives me performance feedback, he or she is considerate of my feelings.
	My manager generally provides feedback in a thoughtless manner. (R)
	My manager does not treat people very well when providing performance feedback. (R)
	My manager is tactful when giving me performance feedback.
Source Availability	My manager is usually available when I want performance information.
	My manager is too busy to give me feedback. (R)
	I have little contact with my manager. (R)
	I interact with my manager on a daily basis.
	The only time I receive performance feedback from my manager is during my performance review. (R)
Promotes Feedback Seeking	My manager is often annoyed when I directly ask for performance feedback. (R)
	When I ask for performance feedback, my manager generally does not give me the information right away. (R)
	I feel comfortable asking my manager for feedback about my work performance.
	My manager encourages me to ask for feedback whenever I am uncertain about my job performance.
Just (Promotes Justice)	
	Job decisions are made by my manager in an unbiased manner.
	My manager makes sure that all employee concerns are heard before job decisions are made.
	To make job decisions, my manager collects accurate and complete information.
	My manager clarifies decisions and provides additional information when requested by employees.
	All job decisions are applied consistently by my manager across all affected employees.
	Employees are allowed to challenge or appeal job decisions made by my manager.
Inclusive	
	My manager encourages team members to take initiative.
	My manager asks for the input of team members that belong to other areas of the company.
	My manager does not value the opinion of others equally. (R)

Appendix B

Information Sheet for Applicant Participants

Department of Psychology
Email: michael.heslop@nzdf.mil.nz
30/10/20
HEC Ref: HEC 2019/10/BL

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Validation of the Leadership Behaviour Assessment: Criterion-Related Validity Evidence

Information Sheet for Applicant Participants

My name is Michael Heslop, and I am a master's student at the University of Canterbury. I am conducting research into the development of the leadership behaviour assessment (LBA), which is an online assessment designed to measure leadership behaviours for personnel selection.

You have been approached to take part in this study because you are a manager in the New Zealand Defence Force at the lead systems/lead capability level.

If you choose to take part in this study, your involvement in this project will involve the conduct of the Leadership Behaviour Assessment (LBA) which is an online assessment that will take 20 minutes of your time and can be completed whenever is most convenient for you.

Additionally, you will be asked for the details of your direct report subordinates who will then be asked to conduct a short survey on your leadership behaviour. Subordinates must freely volunteer to participate in the study and are not to be ordered, instructed or otherwise coerced into participation. They will provide confidential ratings on your leadership behaviours that will be used to determine the predictive ability of the LBA. Their ratings will be held in the strictest confidence, to be viewed only by the research team. You will not have access to their ratings. Please do not consent if you are uncomfortable with this arrangement.

Participation for both you and your subordinates is voluntary and any participant has the right to withdraw at any stage without penalty. You may ask for your raw data to be returned to you or destroyed at any point. If you withdraw, I will remove information relating to you. However, once analysis of raw data starts on 23 Dec 20, it will become increasingly difficult to remove the influence of your data on the results.

The results of the project may be published, but you may be assured of the complete confidentiality of data gathered in this investigation: your identity will not be made public without your prior consent. To ensure confidentiality, you will only be identified by a randomly generated research code on any data submitted. All data will be stored electronically in a password protected file on a password protected computer in a locked room. Personnel data will be stored on DIXS. No person outside the research team will have access to the data. The final product of this research will be a thesis, which is a public document and will be available through the UC Library.

The project is being carried out in partial fulfilment of the requirements for a Master of Science in Applied Psychology by Michael Heslop under the supervision of Associate Professor Christopher Burt, who can be contacted at christopher.burt@canterbury.ac.nz. He will be happy to discuss any concerns you may have about participation in the project.

This project has been reviewed and approved by the University of Canterbury Human Ethics Committee, and the NZDF Org Research Committee. Participants should address any complaints to The Chair, Human Ethics Committee, University of Canterbury, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz).

If you agree to participate, you are asked to complete the consent form via a link in the attached email.

Appendix C

Information Sheet for Subordinate Participants

Department of Psychology
Email: michael.heslop@nzdf.mil.nz
10/11/20
HEC Ref: HEC 2019/10/BL

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Validation of the Leadership Behaviour Assessment: Criterion-Related Validity Evidence

Information Sheet for Subordinate Participants

My name is Michael Heslop and I am a master's student at the University of Canterbury. I am conducting research into the development of the leadership behaviour assessment (LBA), which is an online assessment designed to measure leadership behaviours for personnel selection.

You have been approached to take part in this study because you are a subordinate of another participant in this study. I have located your contact details through a list given to me by your manager.

If you choose to take part in this study, your involvement in this project will involve the conduct of a short online questionnaire about the leadership behaviours of your manager. It will take about 10 minutes of your time and can be completed whenever is most convenient for you. The ratings of your manager's leadership behaviour that you provide in this questionnaire will be held in the strictest confidence. Only the research team will have access to your ratings and no one else in the NZDF will be able to access them. Your manager will not have access to your data.

Participation is voluntary and you have the right to withdraw at any stage without penalty. You may ask for your raw data to be returned to you or destroyed at any point. If you withdraw, I will remove information relating to you. However, once analysis of raw data starts on 30 Nov 20, it will become increasingly difficult to remove the influence of your data on the results.

The results of the project may be published, but you may be assured of the complete confidentiality of data gathered in this investigation: your identity will not be made public without your prior consent. To ensure anonymity and confidentiality, you will only be identified by a randomly generated research code on any data submitted. All data will be stored electronically in a password protected file on a password protected computer in a locked room. No person outside the research team will have access to the data. The final product of this research will be a thesis, which is a public document and will be available through the UC Library.

The project is being carried out in partial fulfilment of the requirements for a Master of Science in Applied Psychology by Michael Heslop under the supervision of Associate Professor Christopher Burt, who can be contacted at christopher.burt@canterbury.ac.nz. He will be happy to discuss any concerns you may have about participation in the project.

This project has been reviewed and approved by the University of Canterbury Human Ethics Committee, and participants should address any complaints to The Chair, Human Ethics Committee, University of Canterbury, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz).

If you agree to participate in the study, you are asked to complete the consent by using the link in the attached email.

Appendix D

Consent form for all Participants

Department of Psychology

Email: mrh160@uclive.ac.nz

Validation of the Leadership Behaviour Assessment: Criterion-Related Validity Evidence

Consent Form for all applicants

- ☐ I have been given a full explanation of this project and have had the opportunity to ask questions.
- ☐ I understand what is required of me if I agree to take part in the research.
- ☐ I understand that participation is voluntary and I may withdraw at any time without penalty. Withdrawal of participation will also include the withdrawal of any information I have provided should this remain practically achievable.
- ☐ I understand that any information or opinions I provide will be kept confidential to the research team and that any published or reported results will not identify the participants or their job titles. I understand that a thesis is a public document and will be available through the UC Library.
- ☐ I understand that all data collected for the study will be kept in locked and secure facilities and/or in password protected electronic form and will be destroyed after five years unless a publication outlet requires extended archiving of the data.
- ☐ I understand there are no risks associated with taking part in this research.
- ☐ I understand that I can contact the researcher Michael Heslop (mrh160@uclive.ac.nz) or supervisor Christopher Burt (christopher.burt@canterbury.ac.nz) for further information. If I have any complaints, I can contact the Chair of the University of Canterbury Human Ethics Committee, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz).
- ☐ By signing below, I agree to participate in this research project.

Name:_Signed:_____Date:_____

Appendix E

Participation Invitation Email for Applicant Participants

Tēnā koutou Ladies and Gentlemen,

My name is Michael Heslop, and I am currently completing a Master of Science at the University of Canterbury. As part of this I am conducting a validation study of an innovative new tool called the **Leadership Behaviour Assessment (LBA),** which is an online assessment designed to measure leadership behaviours for personnel selection.

You have been approached to take part in this study because you are a commander/manager in the New Zealand Defence Force at the lead systems/lead capability level. This research has been approved by the University of Canterbury Ethics Committee (Ref: HEC 2019/10/BL) and the NZDF Organisational Research Ethics Committee (Ref: Organisational research minute 500/PB/5/3).

Participation in this research is voluntary. If you agree to participate in this study, you will be asked to conduct a short online leadership assessment based on a simulation of an online management system. This will take 20 minutes and can be completed on any internet capable personal device. This link will not work on the internal NZDF internet, so we need to send it to your personal email. The LBA is not optimised for mobile phone use so we would prefer you open the link on a PC, laptop or tablet if possible. The LBA can be completed during work hours as stipulated by the NZDF ethics committee. Your interactions with the system will be measured and after the assessment has been validated, a report will be developed for you based on 10 leadership factors. This report will be useful tool for your development and self-awareness as a leader.

Additionally, your direct subordinates will be asked to complete a questionnaire on your leadership behaviour which takes 10 minutes and can be completed on the DIXS internet portal or any internet capable device. **Subordinate participation is completely voluntary, and subordinates cannot be ordered, instructed, or otherwise coerced into participation.** The subordinate ratings will be compared with the online assessment (LBA) report by the researchers and this will be used to validate the tool.

If you agree to be included in this study, please complete the three actions below in the following order:

1. Read the attached information sheet (attached),

2. Complete the consent form: Consent form link.

3. Please send a **list of the serial numbers** of all your **direct report subordinates** in an email to the researcher (michael.heslop@nzdf.mil.nz). Please send the researcher your personal email and a link for the LBA will be sent to it.

After you have completed the consent form and emailed me a list of subordinates, I will email you a link to the LBA and you will be able to complete this at your convenience (please allow 20 mins of uninterrupted time to complete the LBA). Your subordinates will be approached to participate in the study and complete the MBQ on your leadership behaviours (this will only take 10 minutes). The deadline for data collection in this study is **Monday 14 Dec 20**, any data received after this date may not be included in the research.

Your participation in this study will help develop an innovative tool that may be of use to the NZDF in the development and selection of future leaders.

If you have any questions about the process please feel free to contact me.

Thank you for your time.

Ngā mihi

Appendix F

LBA Facilitation Email for Applicant Participants

Tēnā koe,

Thank you for agreeing to participate in the validation study of the Leadership Behaviour Assessment (LBA). Please find your link to the LBA below. This link **will not** work on the DIXS ITD so please conduct the LBA on a personal device. The LBA is not optimised for mobile phone use so we would prefer you open the link on a PC, laptop or tablet. If you have no other option but to use a mobile device, please make sure that after you click on the link you change the phone internet browser setting to "view as desktop site" See attached power point for instructions on how to do that).

The assessment will only work once so please ensure the following before you click the link:

- You are in a quiet environment that is free from distractions,
- You have at least 20 mins of uninterrupted time to conduct the assessment,

Once you have completed your assessment the data will be used to validate the LBA. Once the LBA has been validated you will be provided with a report on your leadership behaviours based on your results from the LBA. This report will be a useful tool for personal leadership development. It may take **over a month** to validate the tool and create the report. The report will be sent to your DIXS email address.

If the link does not work or you have any problems conducting the assessment, please contact me and I will troubleshoot or provide another link.

The research deadline for this is **Monday 14 Dec 20.** Any assessments completed after this research deadline may not be able to be included in the study. Leadership reports may not be generated for assessments received after this deadline.

LBA details:

Your unique research code: XXX111

Your Unique link to the LBA from the University of Canterbury: Link

Thank you for your participation, it means so much. This represents over a year worth of research for a master's thesis and your participation will allow me to complete this qualification.

Ngā mihi,

Appendix G

Participation Invitation Email for Subordinate Participants

Tēnā koe,

My name is Michael Heslop, and I am currently completing a master's at the University of Canterbury. As part of this I am conducting a validation study of a new tool called the Leadership Behaviour Assessment (LBA), which is an online assessment designed to measure leadership behaviours for leadership selection. This research has been approved by the University of Canterbury Ethics Committee (Ref: HEC 2019/10/BL) and the NZDF Ethics Committee.

Your Manager has agreed to participate in the study, and they have given me your contact details as part of the requirements of the research. This research will assist your managers leadership development as they will receive a leadership behaviour report based on their behaviour in the Leadership Behaviour Assessment that they have completed already.

Participation in this research is voluntary and you cannot be ordered to participate. If you agree to participate in this study, you will be asked to complete the manager behaviour questionnaire where you will rate your manager's leadership behaviours. **Your ratings are completely confidential and will only be seen by the researcher.** Your manager **will not** be given access to your ratings or the statistics generated from your ratings. The questionnaire can be conducted on the DIXS ITD internet or on personal devices (including mobile phones) via the link below. It will only take **10 mins** and can be done during work hours as stipulated by the NZDF ethics committee.

The answers you provide will be compared by the researchers to the results of the Leadership Behaviour Assessment completed by your manager and they will be used to confirm the validity of the tool in measuring leadership behaviours.

If you agree to be included in this study, please complete these steps in the following order:

1. Read the attached information sheet (1 page).
2. Complete the consent form at the following link (takes less than 1 minute)
3. Complete the Manager Behaviour Questionnaire at the following link (approximately 10 minutes).

The manager that nominated you for this study is **X** You should base your ratings on their behaviour. Your managers research code is: **XXX111.**

Please conduct the questionnaire by **14 Dec 20.** Any questionnaires completed after this research deadline will not be included in the study. If the link does not work or you have any problems conducting the assessment, please feel free to contact me and I will troubleshoot or provide another link.

This study is the culmination of two years' worth of work towards a master's degree and I want to thank all those who participate. I cannot overemphasize how much your support will assist me in completing this study and qualification. Any assistance you provide is instrumental in the creation of this innovative tool for leadership selection.

Ngā mihi,

Appendix H

Research Deadline Reminder Email for Applicant Participants

Tēnā koutou Ladies and Gentlemen,

The research deadline for the LBA validation study is **14 Dec 20** and any responses received after this date may not be included in the study. All those that participate will receive a personalised leadership style report based on their interactions with the Leadership Behaviour Assessment. If you would like to participate, please complete the consent form included in the email below.

This study is the culmination of two years' worth of work towards a master's degree and I want to thank all those who participate. I cannot overemphasize how much your support will assist in the completion of this study and my master's degree.

Ngā mihi

Research Deadline Reminder Email for Subordinate Participants

Tēnā koutou Ladies and Gentlemen,

The research deadline for the LBA validation study is **14 Dec 20** and any responses received after this date may not be included in the study. If you would like to participate, please complete the consent form included in the email below.

This study is the culmination of two years' worth of work towards a master's degree and I want to thank all those who participate. I cannot overemphasize how much your support will assist in the completion of this study and my master's degree.

Ngā mihi