https://dspace.org/

Consistency between raters in a 360-degree leadership feedback instrument.

Kara, Sarika

2011-09-20
http://hdl.handle.net/10539/10435

Downloaded from a DSpace repository. https://dspace.org/

1. Introduction

1.1 Introduction

In today's volatile, changing and competitive world of work, organisations are continuously looking for ways to increase their performance (Banu & Umamaheswari, 2009). In the light of the fact that most Fortune 500 companies are utilising 360-degree feedback methods to increase performance, many organisations use this method within their leadership development programmes (Atwater, Brett & Charles, 2007; Yogun, 2009). These feedback instruments are used by organisations in their performance management systems for the purpose of obtaining evaluations from supervisors, peers or colleagues, subordinates, and even customers and clients, in addition to the individual's own self-assessment of his or her performance(Carson, 2006; Hassan & Rohrbaugh, 2009). The success of these instruments is based on the premise that those people who have frequent interaction with an individual, who therefore are fairly knowledgeable about the individual's performance, and whose opinions are valued by that individual, provide the best information in terms of increasing individual performance (Hassan & Rohrbaugh, 2009). The advantage of 360-degree feedback questionnaires is that ratings are made anonymously, and therefore allow for candid feedback in a relatively non-threatening setting (Hassan & Rohrbaugh, 2009). These identified purpose, applications, and benefits also apply to the particular 360-degree questionnaire used in the current study.

The questionnaires that are used are standardised questionnaires looking at various dimensions of individual performance that are administered to different parties within organisations (Liviu, Salantă & Butilcă, 2009). Within the current study, a single questionnaire has been administered to both participants and respondents, yielding scores that assess various components of leadership performance of the participants in this study. These components are reflected by the subscales used in the design of the questionnaire and scores obtained for each component are thus analysed in the current context. In the current research, the term "participants" refers to the individuals who are being assessed on leadership performance, and the word "respondents" refers to the

peers or colleagues, subordinates and supervisors who rate each participant respectively (Cook, 2010). Research in this area has mainly focused on the psychometric properties of 360-degree feedback instruments and processes, with an increase in research examination of the ratings between and within groups of respondents (Hassan & Rohrbaugh, 2009). One possible reason for this observed trend is that a deeper understanding into the areas of consistency and inconsistency with an empirical approach is now required to improve the use and understanding of the 360-degree feedback methods (Hassan & Rohrbaugh, 2009). Therefore investigating relationships between groups of respondents, as well as between the ratings of respondents and participants from this particular 360 degree feedback questionnaire would contribute to knowledge regarding how different groups in organisations perceive each other, and the dimensions that contribute to performance.

Research has often investigated the relationships of ratings of general job performance between the different groups of respondents within 360-degree feedback (Beehr et al., 2004). The questionnaire that is used in the current study focuses on leadership performance. Therefore, the current research also focuses on the ratings made by different groups of respondents regarding the leadership performance of participants. The degree of congruence between these participant and respondent ratings has been attributed to many variables such as various biases and self-knowledge (Fletcher & Baldry, 2000). Therefore, this study has attempted to investigate the degree of consistency between the different groups of raters in a particular 360-degree feedback leadership questionnaire. Furthermore, by investigating the degree of consistency for each of the subscales of the questionnaire, the researcher will identify the areas of consistency and inconsistency within the construct of leadership which has not been adequately addressed in past research. The reason being, that in general 360-degree feedback questionnaires assess constructs in a more holistic way, yielding an average score of performance. However, in order to obtain a deeper view into leadership performance in a 360-degree feedback context, investigating areas of consistency and inconsistency at a subscale level might yield better information in terms of the components of leadership performance for each individual who is being evaluated. It is

also hoped that insight will be gained into the specific areas of leadership where individuals' self-perception correlates with that of others and vice versa.

1.2 Aims

The focus of this study planned to be on the ratings made by different groups of respondents regarding the leadership performance of participants. This was achieved by examining the degree of consistency between the various groups of raters of a particular 360-degree feedback leadership questionnaire. Furthermore, by investigating the degree of consistency for each of the subscales of the questionnaire, the researcher hopes to identify the areas of consistency and inconsistency within the construct of leadership.

The following chapter will proceed to review the existing literature to provide a theoretical overview and background to the current study's objectives and aims.

2. Review of literature

Given the emphasis organisations today place on feedback and measuring performance, this chapter will explore the existing literature and research in relation to the particular 360-degree leadership questionnaire used in the current study.

2.1 The 360-degree feedback method

Feedback has been cited as being the most powerful tool for influencing performance and development within organisations (Forte, 2008). Feedback refers to the information people receive regarding the quality of particular performance behaviours (Liviu et al., 2009). Ongoing and constructive feedback can foster continuous learning and improvement within organisations, which in turn may increase competitiveness and success (Forte, 2008; Liviu et al., 2009). Furthermore, feedback can reinforce positive behavioural patterns and assist in taking steps to correct or improve development areas for individuals (Liviu et al., 2009). This development and learning from feedback that was received will most likely occur if the recipient judges the value of the feedback and accepts it as being the truth (Liviu et al., 2009). Furthermore by using results for development purposes, recipients themselves gain more out of the feedback given (Toegel & Conger, 2003).

The concept of feedback forms the basis of 360-degree feedback methods regardless of whether they are used for development or performance appraisal purposes. This feedback process allows for recipients and managers in particular to see themselves as others do (Atwater et al., 1998). In particular, past research has shown that feedback on performance is important to recipients, and provides a source of motivation for improving and changing behavioural patterns (DeNisi & Kluger, 2000). However, research showing whether this motivation actually translates to increased performance is limited (DeNisi & Kluger, 2000). In contrast, there is evidence that shows a decline in performance due to poor-quality feedback and the systems thereof (DeNisi & Kluger, 2000). What is clear, however, is that to ensure the attraction and retention of high-performing and effective individuals, organisations need a sound and effective performance assessment and feedback system (Singh, 2010).

Traditionally organisations addressed the “harder side” of performance by using an individual’s financial and operational success to assess performance (Maylett & Riboldi, 2007). Performance appraisal systems are typically characterised as annual events. These are structured as a type of check-list against pre-established criteria. Usually a single evaluator is employed, and presented to recipients by a supervisor. This typically involves negative feedback, and is tied to remuneration and promotion. It deals with the formal aspects of business, and is typically not sufficiently detailed to yield information that can be used by the recipient as a basis for self-development (Liviu et al., 2009). However, this trend is changing, with organisations also taking into account the “softer side” of performance such as interpersonal relationships (Fletcher et al., 1998; Maylett & Riboldi, 2007). To assess this “softer side”, and adopt a more holistic view of performance, many organisations have utilised 360-degree feedback methods (Maylett & Riboldi, 2007).

The popularity of this type of assessment and feedback has grown since its inception in the late 1980s and early 1990s (Toegel & Conger, 2003). The 360-degree feedback methods are currently used by approximately 90 percent of Fortune-100 companies in the United States of America (Ghorpade, 2000; Smith & Fortunato, 2008). This trend has also been observed across organisations in other countries and cultures (Gillespie, 2005). Reasons for this popularity include the following:

The changing nature of organisations’ internal and external environment
The lack of effectiveness of traditional single-rater appraisals
The lack of holistic focus of traditional appraisals
Increased accuracy
The effects of globalisation (Carson, 2006; Ghorpade, 2000).

Furthermore, the use of 360-degree feedback methods gives organisations a competitive edge, and equips them with tools to deal effectively with turbulent business environments (Becton & Schraeder, 2004).

The 360-degree, or multi-source feedback method, is an innovative management development technique which can also be used for the development of an effective

performance management system within organisations (Hassan & Rohrbaugh, 2009). Compared with the traditional single-source assessments or performance appraisal systems, the 360-degree approach utilises evaluations from other sources as well as the individual's own assessments and evaluations (Hassan & Rohrbaugh, 2009). The "other" means of performance evaluation can come from a variety of sources, such as peers or colleagues, supervisors, subordinates, clients and customers (Hassan & Rohrbaugh, 2009).

There is some debate as to the difference, if any, between 360-degree feedback and other types of multi-source feedback (Foster & Law, 2006). Some theorists view multi-source and 360-degree feedback as the same, while others argue that 360-degree feedback is a type of multi-rater feedback (Foster & Law, 2006). The latter perceive multi-rater feedback as having a minimum of two raters and 360-degree feedback requiring a full circle of feedback (Foster & Law, 2006). A full circle of feedback refers to evaluations sourced from all levels of the organisation in relation to a particular individual's current level. In this way a picture of holistic skill and performance is provided by placing the individual being rated in the centre of a web of relationships rather than traditional hierarchical evaluation (Calhoun et al., 2010; Liviu et al., 2009). For the purposes of this study, the terms are used interchangeably, as this reflects their use in the research sourced for this study. However, the term "360-degree feedback" is used for the purposes of consistency.

The 360-degree feedback method is typically designed following an organisational analysis (Beehr et al., 2001). The organisational analysis yields higher-order competencies, extra-role behaviour, and contextual performance behaviours as a measure of performance (Beehr et al., 2001). The organisational analysis allows for the customisation of 360-degree feedback questionnaires to fit the particular context of use (Beehr et al., 2001). Furthermore, the organisational analysis yields competencies with greater depth, as well as information indirectly related to specified roles which also forms part of performance as a construct. Although being quite different to traditional performance appraisals, 360-degree feedback questionnaires also provide an indication

of broad-band performance (Beehr et al., 2001). Furthermore both can include measures of task performance referring to directly assigned tasks (Liviu et al., 2009).

In contrast with performance appraisals, 360-degree feedback questionnaires provide different perspectives on an individual's performance in the organisation (Hassan & Rohrbaugh, 2009; Liviu et al., 2009). The premise of these measures is that those who frequently interact with an individual are fairly knowledgeable regarding his or her performance, and their opinions are valued, since they would be in the best position to provide accurate and constructive feedback to increase the performance of that individual (Hassan & Rohrbaugh, 2009; Liviu et al., 2009). This justifies the need for obtaining ratings from the participants themselves. For the purpose of completing the questionnaire in this study, one or more peers or colleagues, and one or more subordinates and supervisors were used. Furthermore, in a socially-constructed world, positive or negative judgements about employees within organisations make up an important component of reality for individuals (Becton & Schraeder, 2004). Therefore, the use of more than one rater group is further justified.

There are a number of advantages, disadvantages and cautions related to the use of 360-degree feedback methods within organisations. One of the main advantages of using 360-degree feedback questionnaires is that rater anonymity is increased, where raters or respondents can give candid feedback in a secure context (Atwater et al., 2007; Conger & Toegel, 2002; Ghorpade, 2000). However, honouring this anonymity can be challenging depending on how many people are participating in the process (Wimer, 2002). Compared with single-source feedback methods, information from multiple sources, including the participant, provides more holistic information regarding an individual's performance, and increases the reliability of the feedback (Becton & Schraeder, 2004; Conger & Toegel, 2002; Liviu et al., 2009).

The 360-degree feedback methods have been shown to assist in developing leaders and have positive ripple effects throughout the organisation (Atwater et al., 2007; Singh, 2010). Furthermore, once managers and employees trust and accept the process, the feedback and insight gained can also be incorporated into personnel decisions (Atwater et al., 2007). One way that this trust is created and maintained is through increased

perceptions of fairness created by the 360-degree process by giving employees a chance to be heard, and actively seeking their opinions of managers (Becton & Schraeder, 2004; Carson, 2006; Kerin, 2010; Toegel & Conger, 2003). This feedback process can also increase “self” and “other” understanding of the behaviour and attitudes of individuals, and increase individual ownership for self-development and learning (Banu & Umamaheswari, 2009; Singh, 2010; Weigelt et al., 2004). The understanding and ownership have been shown to increase managerial effectiveness, improving external awareness, human-resource management, communication and interpersonal skills (Banu & Umamaheswari, 2009). In particular the feedback gained can provide a useful way of identifying key competencies and development areas for individuals and organisations (Singh, 2010). Increasing perceptions of fairness and objectivity can also result in improved organisational culture and morale, if 360-degree feedback methods are implemented properly (Atwater et al., 2007; Carson, 2006; Singh, 2010).

Although 360-degree feedback has many potential advantages, its success is related to the way in which it is implemented (Tosti & Addison, 2009). The 360-degree feedback methods are usually time-consuming and expensive (Conger & Toegel, 2002). At times the feedback from these methods is reduced to quantitative values only which miss some of the important qualitative inputs (Conger & Toegel, 2002). Another reason for the ineffectiveness of 360-degree feedback methods is that organisations often use the method without clearly defining its goals and scope. It also does not gain the support of the organisation (Ghorpade, 2000; Kerin, 2010). Furthermore, the methods used fail to focus on benefits for both the individual and organisation (Ghorpade, 2000). This lack of strategic planning, foresight and attention to the process whereby the questionnaires are implemented, may cause the 360 feedback questionnaires to be ineffective (Ghorpade, 2000). Due to the use of different raters in 360-degree feedback methods, rating errors such as particular biases, may also creep into ratings and confound results and the accuracy of feedback (Ghorpade, 2000). The feedback from the process has at times been found to be overloaded, confounding, inaccurate and difficult to interpret (Tosti & Addison, 2009). Using different sources, although useful, may also yield too much information to interpret and make sense of. The potential of errors and biases

increases the possibility of confounding results, and although being objective ratings, may not always be accurate.

The design of 360-degree feedback questionnaires initially responded to the need of many organisations for standardised work-unit performance and to validate constructs such as leadership, development and performance management (Spangenberrg & Theron, 2004). This is evident from the fact that approximately 22 percent of organisations in the USA use generic instruments (Craig & Clark, 2009). However, with this standardisation comes the risk of inappropriately generalising to other contexts which would require a tailor-made questionnaire. This explains the current loss of appeal of standardised generic feedback instruments (Craig & Clark, 2009). One way to ensure a unique instrument is to gather context-specific information when compiling questionnaires where items assess performance that is relevant to the organisation and culture in question (Ghorpade, 2000).

The 360-degree feedback questionnaires have generally been used for managers and supervisors with its use expanding to other members of organisations (Liviu et al., 2009). In theory the 360-degree feedback method should be used for developmental purposes, and not for reasons of purely evaluating that encompass many dimensions of performance (Liviu et al., 2009). This is to ensure that raters provide feedback without concerns of negative consequences to the individual being evaluated, especially if evaluation is linked to remuneration and rewards. Furthermore, recipients are more receptive to feedback if not tied to performance appraisals. The cultural context of the organisation and broader environment has also been shown to impact upon ratings in 360-feedback circumstances (Conger & Toegel, 2002). A study measuring the equivalence of a 360-degree feedback survey in the USA and two international locations (United Kingdom and Hong Kong) supported this assertion by showing that constructs have different levels of relevance for each culture (Gillespie, 2005). However, in the context of this study, this factor has not been addressed owing to a lack of demographic information on the sample. The South African context, however, could present its own unique influences on ratings.

The 360-degree feedback method was originally designed for individual development purposes, and not for performance appraisal purposes (DeNisi & Kluger, 2000; Hassan & Rohrbaugh, 2009; Singh, 2010). However, at times it is used as a form of performance appraisal, and this can be damaging by decreasing the likelihood of improvement, effectiveness and continuous learning (Conger & Toegel, 2002; DeNisi & Kluger, 2000; Hassan & Rohrbaugh, 2009). Furthermore, if used for performance appraisals or selection, the use of 360-degree feedback methods can lower morale of employees, and may even lead to possible litigation (Kanouse, 1998). Evidence of this claim was provided by a study looking at 396 leaders in a large public transportation company in Eastern America; where a 360-degree instrument was administered on two occasions (Zimmerman et al., 2008). The first occasion was for development purposes and the second for administrative purposes (Zimmerman et al., 2008). Results showed that 360-degree feedback for development purposes differed in perspective, purpose and performance when compared with feedback for administrative purposes such as remuneration (Zimmerman et al., 2008). Feedback has been shown to have special status if results are confidential and used for developmental purposes (Toegel & Conger, 2003). In some cases where feedback was utilised for performance appraisal purposes, the data yielded became “owned” by the organisation and not the recipient, and this might have affected the acceptance and success of the feedback process (Toegel & Conger, 2003). Organisations need to keep this in mind when deciding on implementation of 360-degree feedback methods. Furthermore, this phenomenon could also have impacted on the results of this study.

If used appropriately, 360-degree feedback can be used effectively to examine behavioural patterns that are associated with specific organisational outcomes (Zimmerman, Mount & Golf, 2008). Organisations need to ensure that this method is introduced and implemented in a way that is accepted by all those concerned to maximise the benefits associated with the process (Atwater et al., 2007). Utility of 360-degree feedback questionnaires may depend on the ability to develop consensus on the meaning of performance, regardless of whether it is used for developmental purposes or not (Diefendorff, Silverman & Greguras, 2005). Furthermore, organisations need to be cognisant of the fact that, depending on the source of ratings, the purpose of any

feedback differs, (Singh, 2010). Therefore it is important to examine the relationships between ratings provided by different respondents.

2.2 The nature of 360-degree feedback ratings

Past research initially focused on the psychometric properties, advantages and disadvantages of the 360-feedback method (Hassan & Rohrbaugh, 2009; Hooiberg & Choi, 2000). However, more recent studies have focused on investigating the relationships between each group of respondents and the accuracy of their ratings (Hassan & Rohrbaugh, 2009). Both approaches have been used in this study, with an emphasis on the latter. Various types of comparisons between respondents can be examined. One can either compare participant to respondent ratings ("self-other ratings", or S-O ratings), or other groups of respondents' ratings to each other ("other-other ratings" or O-O rating) (Hassan & Rohrbaugh, 2009).

Past research highlights the importance of distinguishing between sources of feedback which yield different responses and perspectives (Becton & Schraeder, 2004). Furthermore, when conducting analyses on 360-degree feedback data there are three dimensions to take into account (Toegel & Conger, 2003). The individual dimension refers to self-reflections or ratings which increase self-awareness, followed by the interpersonal dimension referring to the interactive process through others' perceptions and ratings, and concluding with the organisational dimension referring to the way in which ratings are processed by facilitators of 360-degree feedback methods (Toegel & Conger, 2003). The rationale behind the use of multiple sources of ratings is that any source in isolation, although having accurate components, is largely problematic and inaccurate (Fleenor et al., 2010). Ratings may be higher or lower than those of other sources of ratings or they may be the same. This further shows the need for looking at all sources (Fleenor et al., 2010).

One particular approach in past research has been to examine the congruence of S-O ratings in order to determine their relationships in terms of individual outcomes such as performance (Atwater et al., 2005; Fleenor et al., 2010). S-O agreement has also been referred to as a gap-analysis process that systematically compares others' ratings with

participants' self-ratings (Calhoun et al., 2010). There are many factors that impact upon S-O agreement that have differing consequences on outcomes (Atwater et al., 2005). The general consensus is that greater agreement is indicative of a mutual understanding of the rating process and common perceptions between raters (Atwater et al., 2005). The typical 360-degree feedback process and instrument is designed to highlight S-O agreement or disagreement in order to increase awareness regarding discrepant areas that require behavioural change or development (Atwater et al., 2005; Calhoun et al., 2010; Fleenor et al., 2010).

S-O ratings yielded the following four response types:

Over-estimators
In-agreement poor estimators
In-agreement good estimators
Under-estimators (Atkins & Wood, 2002; Fleenor et al., 2010).

Over-estimators are participants who consistently rate themselves higher than others do or higher than their actual performance (Atkins & Wood, 2002). Conversely, under-estimators rate themselves lower (Fleenor et al., 2010). In-agreement poor estimators are participants who give themselves low ratings, coupled with low ratings from others (Atkins & Wood, 2002). Participants who give themselves high ratings corresponding to the high ratings of others are in-agreement good estimators (Atkins & Wood, 2002). O-O ratings may also be in agreement or disagreement according the above four response types between groups of respondents in relation to the participant being evaluated.

The relationship between self-ratings, others' ratings, and specific outcomes is complex, and interpretations should be approached with caution (Atwater et al., 1998). Adding to this level of complexity, theorists argue that research on 360-degree feedback has not kept up with practice in organisations providing few and insufficient models and data to guide practitioners and organisations (Antonioni & Park, 2001). Insufficient models and implementation may lead to errors and biases in the 360-degree feedback process, thus impacting on the quality and accuracy of ratings. However, past research has agreed

upon the common finding that there are varying degrees of convergence between rater sources which are dependent upon many factors that are internal and external to the organisation and raters participating in any 360-degree feedback process. All of these will be elaborated on at a later stage (Atkins & Wood, 2002). In addition, this convergence is dependent upon the perspective and purpose of the feedback instrument, that is for development or administrative purposes (Zimmerman et al., 2008). Each purpose provides a unique perspective for raters, and has strong effects on the relationship between ratings and performance (Zimmerman et al., 2008). Therefore, both S-O and O-O agreements will be explored in following sections of this review.

Although it is beneficial to gather feedback and ratings from a variety of sources to determine an individual's performance, more information does not always ensure better feedback (Ghorpade, 2000). Furthermore, according to the results of a study conducted by using data from a multinational oil company in the USA, the intended objectivity that is built into the 360-degree process does not always ensure accuracy (Fletcher et al., 1998). Ratings from the different sources within a 360-degree feedback context may be congruent or differ, with differences observed in many different combinations (Atkins & Wood, 2002). Furthermore, although once regarded as errors, rating discrepancies are now seen to carry meaningful information such as providing insight into leaders' interpersonal worlds (Atkins & Wood, 2002; Fleenor et al., 2010).

Looking more closely at S-O ratings, research has shown self-ratings to be more lenient with increased congruency between other rater groups (Atkins & Wood, 2002; Heijden & Nijhof, 2004; Patiar & Mia, 2008; Yammarino & Atwater, 1993). This finding was further supported by the results of a study on 2 213 employees of a Midwest American insurance company who completed a 360-degree feedback questionnaire where self-ratings were found to be much higher than others' ratings (Beehr et al., 2001). Although self-ratings provide a unique perspective on performance, the ratings may be less valid considering the tendency to overrate oneself (Beehr et al., 2001). Research also shows that aside from possibly being less valid, self-ratings are also the least accurate (Eichinger & Lombardo, 2004). Furthermore, over-raters also tend to show lower subordinate ratings and low self-awareness (Atwater et al., 1998).

In contrast, individuals who tend to underrate themselves showed mixed results on performance in past research (Atwater et al., 1998). Research conducted on a sample of 63 participants in a large Australian service provider showed that under-estimators tend to be highly competent with corresponding higher supervisor and peer or colleague ratings in a 360-degree feedback context (Atkins & Wood, 2002). This effect, with higher than expected ratings from “others” is called the “halo effect” referring to the elevation of ratings by raters of the individual which does not reflect the accurate score of performance (Beehr et al., 2001). The halo effect falls under the broad category of biases (Fletcher et al., 1998

This bias points to the assertion that in a 360-degree feedback context there is less of a concern about accuracy and more concern regarding impression management (Antonioni & Park, 2001; Van der Heijden & Nijhof, 2004). The implication is that in general raters do not always have an accurate or clear indication of individuals’ actual performance (Patiar & Mia, 2008).

Apart from the accuracy of ratings, there are various other factors that contribute towards S-O rating incongruence within the typical 360-degree feedback method. In relation to self-ratings, the ratings by bosses were shown to be the most accurate due to a common frame of reference within the 360-degree feedback context as a consequence of prior training and experience. This finding was shown by a study conducted on 2 350 managers participating in a management skills profile, as well as from a sample of 433 managers from a midsized insurance company in the USA (Antonioni & Park, 2001; Sytsma & Hezlett, 1998). A study with a sample of managers across luxury hotels in Australia showed that the biggest difference lay between supervisor’s and self-ratings as supervisors tend to be more conservative in their ratings, and are more accustomed to the evaluation process (Patiar & Mia, 2008).

Method effects have also been shown to be the largest source of S-O disagreement and inaccuracy in performance ratings of a 360-degree feedback context (Sytsma & Hezlett, 1998). Referring specifically to others’ ratings, raters may evaluate individuals based on prototypical mental representations of leadership performance. They may have low intelligence and knowledge of the individual’s tasks or broader work environment, and

may have conflicting goals which affect the credibility and accuracy of ratings (Fleenor et al., 2010). Related to this, the context in which a rater observes, and then rates, a particular individual may differ, with any recent negative events influencing the overall rating in the feedback process (Fleenor et al., 2010). Having the appropriate knowledge and relevant exposure to observe the participant also influences the congruency and accuracy of S-O ratings in 360-degree feedback (Smith & Fortunato, 2008). Organisations also need to keep track of how often raters are asked to rate the same individuals on more than one occasion, as they may not be able to identify subtle improvements over time (Fleenor et al., 2010). Gender of raters and participants, as well as age and tenure within the organisation, can impact on others' ratings in the 360-degree feedback context (Fleenor et al., 2010). The mood of raters can also influence the ratings given where an overly positive mood may cause leniency (Fleenor et al., 2010). The personality of raters is another factor that most research studies or organisations do not take heed of, as personality is mostly impractical and difficult to isolate (Fleenor et al., 2010; Patiar & Mia, 2008).

Another interesting factor affecting S-O agreement is that the more discomfort a person has with the rating process, the less likely ratings will reflect reality or correspond to self-ratings (Fleenor et al., 2010). This discomfort may be due to the risk of accountability, low interdependence, fear and low trust in the process, all of which are governed by the context and culture within an organisation (Fleenor et al., 2010; Smith & Fortunato, 2008). If the organisation is shown over time to lack integrity; mistrust and cynicism can become entrenched into the culture, and in turn may decrease the honesty of 360-degree feedback ratings (Smith & Fortunato, 2008). This might even conceal valuable differences from different sources of raters or increase S-O and O-O disagreement. Following from this, the external environment and culture influence the way in which raters provide feedback (Atwater et al., 2005; Fleenor et al., 2010). A more collectivistic culture may show less leniency bias and increased S-O agreement, while, looking at results across 21 countries, more individualistic cultures may have inflated leadership-behavioural self-ratings (Fleenor et al., 2010). This is especially evident where western individualistic practices are enforced in organisations with differing cultural contexts (Atwater et al., 2005). In these situations S-O agreement or

incongruence is invalid, and thus not given enough importance when examining 360-degree feedback data (Atwater et al., 2005).

Another important reason accounting for the discrepancy between S-O ratings is that people's perspective of leadership performance behaviours differs according to their location within the organisation's hierarchy (Beehr et al., 2001; Fleenor et al., 2010; Hooijberg & Choi, 2000). Depending on the level of interaction, raters may have different exposure to behaviours of individuals' roles (Fleenor et al., 2010; Fletcher et al., 1998). By asking subordinates to rate their superiors, raters may feel uncomfortable in violating traditional organisational structures, and thus be reluctant to provide honest or accurate ratings (Smith & Fortunato, 2008). Therefore, the purpose and quality of feedback differ depending on the source (Antonioni & Park, 2001; Singh, 2010).

A further contributing factor influencing self and other rating agreement is the choice of a rater for the 360-degree feedback process. One finding of a study that was carried out on the validation of a 360-degree instrument demonstrated that when participants were given the option to choose their raters, validity may be compromised (Atkins & Wood, 2002; Carson, 2006). This is because individuals usually choose raters that they believe will rate them favourably resulting in artificially elevated ratings which may mask crucial differences in the process, or highlight inconsequential differences (Atkins & Wood, 2002; Becton & Schraeder, 2004; Toegel & Conger, 2003). This bias is particularly evident when individuals choose raters with whom they have friendly or good relationships (Hooijberg & Choi, 2000; Patiar & Mia, 2008). Their mutual liking which has been developed can influence what is observed and what is remembered, and hence, the attributions made by raters (Antonioni & Park, 2001; Fleenor et al., 2010). However, some theorists may argue that instead of influencing performance ratings, the good performance displayed by individuals result in good relationships and friendships, thus not elevating the results artificially (Patiar & Mia, 2008). However, a full understanding of context would clarify this matter.

Conversely, a general dislike or desire to retaliate against a participant can also negatively bias a rater against such an individual, and may result in giving him or her negative and inaccurate ratings (Fletcher et al., 1998; Smith & Fortunato, 2008). This is

referred to in the literature as “interpersonal affect” where either the like or dislike of an individual may influence ratings (Antonioni & Park, 2001). This emotional reaction towards participants dictates what is observed, remembered and the attributions raters may make (Antonioni & Park, 2001). A manifestation of this effect is the leniency or severity bias referring to the tendency to give higher or lower ratings than accurate performance. This may present many organisational challenges such as legal problems and perceived unfairness (Antonioni & Park, 2001).

One way to overcome these problems with the self-selection of raters is to adopt objective methods for selecting raters (Becton & Scraeder, 2004). However, if participants have no input into the selection process, feedback received may not be accepted, and may not promote effective development or change (Becton & Schraeder, 2004). Furthermore, it is important to select raters who share a common frame of reference to provide holistic observations of the same context and behaviours (Fleenor et al., 2010; Moses et al., 1993). On a more collaborative basis, organisations need to assist in selecting raters who have known the person long enough to make full, clear and accurate observations (Eichinger & Lombardo, 2004). The standard of best practice is raters who have known the person long enough to see past first impressions, but not too long to generalise (Carson, 2006; Eichinger & Lombardo, 2004). Furthermore, choosing raters that are untrained in correct procedure within the 360-degree feedback context may reduce the usefulness of feedback, and present the opportunity for personal attacks or praise to seep in (Kanouse, 1998). Participants also need to be trained in the correct selection of raters if allowed to be part of the selection process in 360-degree feedback methods (Conger & Toegel, 2002). By choosing appropriate raters, S-O and O-O comparisons can be made objectively and accurately.

Although the majority of past findings from research conducted in the 360-degree feedback area focused on the incongruence between the different sources of ratings, some studies show no differences between self and other ratings (Patiar & Mia, 2008). One possible explanation of this finding is that individuals being rated are high self-monitors, and can adapt behaviour according to the requirements of a situation, and thereby increasing consistency of results and S-O ratings (Fleenor et al., 2010). Rater

effects and biases may also conceal differences between S-O and O-O rater groups coupled with the statistical challenge of large standard deviations (Weigelt et al., 2004). However, rating congruence may not always be negative where self-other ratings that are positive and in agreement indicate good performance (Atwater et al., 1998; Atwater et al., 2005). In this case congruence is more advantageous than a lack thereof (Atwater et al., 2005). Conversely, self-other ratings that are negative and in agreement indicate poor performance (Atwater et al., 1998).

Other reasons for rating congruence are that raters may conserve accurate feedback and match perceptions of individuals to avoid risking relationships and jeopardising team synergy (Peiperl, 2001). Another trend in 360-degree feedback processes is to average out ratings from sources to make comparisons between and within groups (self, peer/colleague, supervisor and subordinate). This may conceal any variations that may be present (Conger & Toegel, 2002; Toegel & Conger, 2003). Other statistical methods used to analyse data may also cause this congruency of ratings (Fleenor et al., 2010).

Although it is important to acknowledge S-O agreement, the inherent S-O disagreement supports the need to examine the 360-degree feedback questionnaire of this study, in order to identify trends between self-raters and others.

Choosing a broad range of rater sources also improves the validity and accuracy of feedback to participants, thus ensuring at least one to three self-ratings, peer or colleague, supervisor and subordinate ratings (Carson, 2006; Fleenor et al., 2010). Corresponding to this, O-O ratings, although not covered in immense detail in past research is also an important area of investigation within the 360-degree feedback method context. One possible reason for the lack of focus on this aspect, is that in general O-O ratings tend to be parallel to each other, but as a whole different to self-ratings of 360-degree feedback questionnaires (Haasan & Rohrbaugh, 2009; Heijden & Nijhof, 2004; Patiar & Mia, 2008; Yammarino & Atwater, 1993). However, a study conducted on 2 213 employees of an American insurance company revealed that in particular supervisor and peer or colleagues ratings correlated more strongly with each other (Beehr et al., 2001). This finding was further supported by a study carried out on a

sample of managers from North America, the United Kingdom, Italy, France and Denmark who participated in a leadership development programme (Atwater et al., 2005).

Looking at subordinate ratings in comparison with the other groups of respondents, the above-mentioned halo affect that is present in S-O ratings, is particularly salient (Antonioni & Park, 2001; Smith & Fortunato, 2008). Subordinates have been shown to intentionally inflate ratings when compared with other groups of respondents and self-ratings (Smith & Fortunato, 2008). Possible reasons for this may be to avoid perceived retribution, limit interpersonal conflict, or to ingratiate themselves (Smith & Fortunato, 2008). A similar trend may also be observed for peer ratings (Antonioin & Park, 2001). However, subordinate ratings have still been shown on average to be higher (Antonioni & Park, 2001). One possible reason for this discrepancy is that the source of feedback determines the nature and quality of feedback, where subordinate feedback is viewed as a source of development guidance, and peer ratings are viewed in terms of providing a teamwork perspective (Singh, 2010). This teamwork perspective yielded by peer ratings may also be influenced by existing team dynamics and politics (Peiperl, 2001). Therefore, the selection of peer raters needs to be accompanied by organisational assistance in order to ensure objectivity, and decrease the potential effects of related biases (Becton & Schraeder, 2004).

Looking at O-O ratings as a whole, past research has shown that supervisor ratings tend to be the most accurate followed by peers or colleagues, and lastly subordinates (Eichinger & Lombardo, 2004). These being in relation to self-ratings. Among factors impacting upon 360-degree feedback ratings , research shows that when job security is high with no significant organisational changes forecasted, ratings from others are usually the most valid and effective (Kanouse, 1998). Furthermore it is important for organisations to have clear and open communication regarding the purpose of using the questionnaire, how data will be used, and how the feedback process will occur. It is also important to note that regardless of purpose, ratings from all groups are subject to the influence of personal judgement and bias (Fletcher et al., 1998). This implies that raters are not always honest, anonymity cannot always be guaranteed, ratings cannot

necessarily force behavioural change, and higher ratings do not always translate into higher performance on the job (Pollman, 1990).

In order to avoid repetition in this review, the factors described as affecting S-O ratings also apply to O-O ratings. Therefore, when considering the nature of and patterns within O-O ratings, these factors should also be considered.

Taking into account that the current study investigates the psychometric properties of the particular 360-degree feedback questionnaire that is used to gather data, these properties are also important to consider when investigating S-O and O-O ratings. The psychometric properties of 360-degree feedback instruments also impact on the quality and congruency of S-O and O-O ratings (Fleenor et al., 2010). If a measure shows equivalence across different rater groups, one may conclude that any differences between ratings are correct and meaningful (Fleenor et al., 2010). This equivalence refers to whether the instrument items, constructs and dimensions mean the same thing for different raters, and measure the same things across the entire process (Diefendorff et al., 2005). Furthermore operationalising the dimensions and performance to create common understanding is the key to establishing this equivalence (Fleenor et al., 2010). By comparing ratings between rater groups, some studies have attempted to validate their respective 360-degree feedback instruments, and to make comparisons within group ratings in order to establish inter-rater agreement (Atkins & Wood, 2002). However, due to inherent inaccuracies and biases presented above, this may not be a true reflection of validity, where comparison with another source of performance such as assessment centre ratings might be more effective (Atkins & Wood, 2002). In order to yield significant differences or valid findings, when designing 360-degree feedback instruments, only items with significant item discrimination and good factor loadings should be included (Fletcher et al., 1998). To further improve design, equivalence should be established to ensure the same factor structure across groups, as well as to ensure that raters all respond in the same manner (Diefendorff et al., 2005). A good way of checking this equivalence is to conduct confirmatory factor analyses (Diefendorff et al., 2005). This supports the approach that is used in the current study, and which will be detailed in the results chapter.

Following from this by training raters in the 360-degree feedback process, organisations equip raters with expert knowledge which may increase the acceptance, quality and accuracy of ratings (Carson, 2006; Fleenor et al., 2010; Liviu et al., 2009). Furthermore, establishing norms of performance and context will improve the quality and accuracy of ratings (Fleenor et al., 2010; Moses, Hollenbeck & Sorcher, 1993). This can aid interpretation when investigating the relationships and trends of S-O and O-O ratings. By designing 360-degree feedback instruments to be both trait and reference based, a common understanding is created which provides raters with the tools to provide complete descriptions in their ratings (Moses et al., 1993). Coupled with effective design, in order to ensure maximum effectiveness of 360-degree feedback methods, organisations can provide coaching, ensure understanding of results, and create a safe and supportive atmosphere (Peiperl, 2001; Smith & Fortunato, 2008; Toegel & Conger, 2003). This approach also leads to sustained leadership development and performance, especially if coupled with structured development plans (Singh, 2010).

Related to leadership performance specifically, S-O agreement has been shown to be an indication of leadership effectiveness (Fleenor et al., 2010). This rating congruence has been shown to be a reflection of high self-awareness which has shown to improve performance and other organisational outcomes (Fleenor et al., 2010; Fletcher & Baldry, 2000; Singh, 2010). Self-awareness refers to seeing the self as others do, and having a deep understanding of own strengths and development areas (Alimo-Metcalfe, 1998). This corresponds to one of the major aims of designing and using 360-degree feedback instruments in organisations, which is to enhance self-awareness and thus promote development (Calhoun et al., 2010; Fletcher & Baldry, 2000; Kerin, 2010). In particular research has found that if used for developmental purposes 360-degree feedback can, and does, increase self-awareness (Pollman, 1990). As mentioned already, self-ratings often suffer from self-enhancement bias, and thus the feedback from multiple sources can highlight this discrepancy, which in turn motivates a change in behaviour (Atwater et al., 2005). However, this change of behaviour is not always easy as self-awareness has been shown to be relatively stable over time (Calhoun et al., 2010). Furthermore, by highlighting the discrepancy between self-rating and others, insight can be gained into an individual's self-awareness regarding his or her strengths and areas requiring

improvement (Calhoun et al., 2010; Carson, 2006; Fletcher & Baldry, 2000; Levine, 2010; Van der Heijden & Nijhof, 2004; Woo, Sims, Rupp & Gibbons, 2008). This finding was determined in a study that examined the 360-degree feedback ratings of a medical team in Louisville, USA, where inflated self-ratings indicated blind spots in individuals' self-awareness (Calhoun et al., 2010). Through the use of 360-degree feedback methods, organisations can establish a formal leadership development programme to allow leaders and managers to engage in more self-reflection (Kerin, 2010).

From what has been addressed in terms of S-O and O-O ratings and the factors impacting on agreement at disagreement between ratings, the need for examining the particular 360-degree feedback questionnaire of this study is clear. In doing so, trends between raters and the implications thereof can be identified and explored. Considering the particular questionnaire of the current study that measures leadership performance, the ties between the 360-degree feedback process and leadership are also important to consider.

2.3 Leadership and the 360-degree feedback method

Taking into account that in the current study the instrument was used to measure leadership performance, it is important to provide a brief insight into leadership performance and development in the current context. In past research, the 360-degree feedback and action-learning have been cited as being the two most popular approaches used by organisations today. This is the result of a shift towards leadership development in the past decade as well as a shift towards flatter organisational structures (Conger & Toegel, 2002; Toegel & Conger, 2003). The reason for this is that the success and effectiveness of organisations depend on the quality and effectiveness of leadership (Conger & Toegel, 2002). This method has proved to be effective in monitoring and improving leadership performance, especially when used solely for developmental purposes in organisations coupled with conducive organisational and broader cultural context (Gillespie, 2005; Zimmerman et al., 2008).

There appears to be many and varied definitions of leadership-encompassing traits, behaviours, interactions, roles, follower perceptions, influence, tasks performed and

organisational culture (Kerin, 2010). Furthermore, cultural context can shape the perceptions of what leadership is, and what effective leadership performance constitutes (Gillespie, 2005). Typical management activities of planning, organising, staffing and problem-solving refer to what is called “transactional leadership” (Alimo-Metcalfe, 1998). However, with increasing environmental demands on organisations, greater complexity and ambiguity, a more transformational leadership style is required to ensure effectiveness and competitiveness (Alimo-Metcalfe, 1998; Kerin, 2010). Transformational leadership gained popularity in the 1980s and adopted a hybrid approach, encompassing the features of transactional leadership as well as a broader impact with the end goal of inspiring and empowering others (Alimo-Metcalfe, 1998; Kerin, 2010). The components of this style include the following:

Charisma
Inspirational motivation referring to creating optimism and drive to mission or goals
Intellectual stimulation referring to the encouragement of others to challenge ideas and seek feedback
Individualised consideration referring to an active belief in others and their abilities (Alimo-Metcalfe, 1998; Kerin, 2010).

Transformational leadership styles usually result in a strong identification with the leader, and a shared vision beyond self-interest (Kerin, 2010).

Looking at a study that measured transformational leadership performance in a 360-degree feedback context, others’ ratings were more congruent to each other than to self-ratings, where self-ratings were inflated, and reflected a misjudgement in recognising development needs (Alimo-Metcalfe, 1998). Looking at a sample of leaders across different countries in America and Europe, others’ ratings specifically were found to be more strongly related to leadership performance (Atwater et al., 2005).Following from this, it was also found that satisfaction and acceptance of feedback corresponded to levels of congruence to self-ratings (Alimo-Metcalfe, 1998). The premise of using 360-degree feedback methods in the context of leadership development is that the art of leadership is difficult to teach, but can be learned, and that development can be stimulated by shedding awareness into strength and development areas (Kerin, 2010).

Leadership style and traits greatly influence the ratings provided in feedback processes (Sytsma & Hezlett, 1998). Furthermore, the general trend in organisations using traditional evaluation methods shows a lack of negative feedback to leaders (Kerin, 2010). This is due to power relations, low acceptance of criticism, a history of successes, and little time for self-reflection by leaders (Kerin, 2010). Therefore formal leadership initiatives such as the 360-degree feedback process may not yield surprising contents, but can provide a holistic and safe way of motivating self-reflection and awareness to address development areas (Kerin, 2010). The general dissatisfaction with traditional measures of performance which focused mainly on financial measures and outputs has also supported the use of 360-degree feedback methods to accurately assess leader performance in organisations. This was corroborated by findings from a sample of 60 part-time MBA students who managed work units at the University of Stellenbosch (Spangenberrg & Theron, 2004).

Looking more specifically into leadership and the 360-degree feedback process, the insight gained from examining the congruency of ratings drives behavioural change and leadership development (Conger & Toegel, 2002). Discrepancies may, however, be inherent when rating leaders do not act in a way that is similar to rater groups that yield differing perceptions (Fleenor et al., 2010). This incongruence is also a reflection of a leader's self-awareness where congruency of 360-degree feedback ratings or lack thereof, allows leaders to see themselves as others do which may highlight gaps in accurate self-perception (Alimo-Metcalfe, 1998; Atwater et al., 2005). Leaders whose self-ratings are congruent with others' are shown to be more self-aware with higher levels of observed performance (Fleenor et al., 2010). In the light of these factors, feedback must be approached with caution as inconsistencies that are identified may not always be accepted or utilised for personal development if individuals are dissatisfied or have a low understanding of ratings (Conger & Toegel, 2002). Furthermore, if leaders do not feel there is a need for change, feedback is rejected with no gains from the 360-degree feedback process (Kerin, 2010). This is because leaders tend to be satisfied with feedback that is consistent with their own self-perceptions, even if negative (Fleenor et al., 2010). Evidence of this was found from the results of a study of 84 executives from an American government agency who found that feedback from

the 360-degree process was deemed not to be meaningful, and was thus rejected. This was due to poor feedback administration and cultural influences (Kerin, 2010). Therefore, one way of ensuring that feedback is accepted and translated into effective leadership development, is to couple feedback with ongoing coaching and follow-up as leaders do not always have superiors governing their progress (Carson, 2006; Kerin, 2010; Levine, 2010). Doing so ensures that feedback is coupled with deeper cognitive processes of leaders to create perceptions of credibility, usefulness and accuracy to promote acceptance and leadership performance development (Becton & Schraeder, 2004). It is also worth noting that assessing leadership performance and competence is complex, and it is not always effective to use one source of evaluation such as 360-degree feedback methods in isolation (Alimo-Metcalfe, 1998).

Having an understanding of the leadership dimension of the 360-degree feedback instrument that is used in this study provides a point of reference when analysing and interpreting the findings of this study, as well as an understanding of the implications of leadership development.

2.4 Research Questions

Keeping in mind the existing literature and research conducted within this area, the following two particular research questions were explored during this research:

Question 1

What is the degree of consistency between different groups of raters of a particular 360-degree feedback questionnaire?

Question 2

Where are the areas of consistency and inconsistency between groups of ratings in a particular 360-degree feedback questionnaire?

2.5 Conclusion

In conclusion, a vast amount of research into the 360-degree feedback area explored the nature and psychometric properties of the instruments that were used, S-O ratings,

O-O ratings, the impact on individual and organisational outcomes and the uses and implications of using these questionnaires. All of these have been explored in this review.

Delving deeper into this area requires an exploration into the relationships between ratings, S-O and O-O. This leads to a better understanding of the appropriate use of and value of 360-degree feedback questionnaires in organisations. Furthermore, by gaining this understanding, organisations and practitioners can make the appropriate links to organisational outcomes such as leadership performance. Being able to adequately develop leadership through its accurate assessment, can improve organisational effectiveness.

Taking all this into account, the need for examining a particular 360-degree feedback questionnaire in this study in relation to the research questions is supported. Furthermore, all the factors and points discussed in this chapter, facilitate interpretation of results in the study.

3. Methodology

3.1 Design

This study utilised a relational quantitative methodology with an *ex post facto* and correlational research design examining the relationships between variables in a statistical sense (Whitley, 2002). The *ex post facto* design involves a study where the researcher has no control over either the participants or events of interest (Neale & Liebert, 1986). A major benefit of this particular research design is that it is cost-effective, and involves no direct manipulation of independent variables (IVs) or participants (Neale & Liebert, 1986). Since there is no manipulation of the independent variable (IV), there is a non-experimental research design, and it is cross-sectional because it can analyse existing data in a relatively short period of time (Whitley, 2002). Archival data is used, eliminating the need for collecting data, and allowing the researcher only to investigate relationships between variables. Furthermore, this brings to attention one of the limitations of an *ex post facto* research design. This limitation is that, owing to the lack of manipulation by the researcher, there could be uncertainty about possible circumstances or extraneous variables that might have impacted on the results (Leedy, 1993). Archival data is used. This highlights the appropriateness of the correlational research design which refers to the research methodology where the relationship between two or more variables can be assessed (Whitley, 2002). Furthermore, the correlational design is appropriate for this study as variables cannot be manipulated (Whitley, 2002). Advantages of a correlational research design include the enablement of researchers to analyse a large number of variables within a single study, and the ability to explore information regarding the degree of these relationships (Stangor, 2010). The fact that inferences cannot be drawn about the causal relationships between and among the variables may be seen as a limitation, as may be the possibility that extraneous variables that are related to observed relationships cannot always be accounted for (Stangor, 2010).

3.2 Sample

Archival data was sourced for the current research. Archives are records or documents that describe the characteristics of individuals, groups or organisations that are collected and maintained for purposes aside from research, and are collected by individuals other than the researcher (Calantone & Vickery, 2009). Statistical archives were used in this study. This term refers to data that has been stored in summary form (Whitley, 2002). Specifically, data from the 360-degree feedback questionnaire(Leadership 3), was selected from the organisation's archives for analysis purposes in this study (Cook, 2010). The data was originally used as part of a leadership development intervention where the 360-degree feedback questionnaire identified strengths and development areas (Cook, 2010).

The advantages of using archival data include cost- and time-effectiveness, access to a larger population and increased generalisability owing to the increased sample size (Kluwin & Morris, 2006). Furthermore, the use of archival data reduces the number of ethical issues that are involved in research such as the following:

Addressing researcher biases
Obtaining individual consent
Providing feedback to participants on completion of the study (Whitley, 2002)

The following are disadvantages of the use of archival data:

Problems of validity
The possibility of available data not providing appropriate links to relevant theoretical constructs
Biases in the initial data collection
Data being out of date
Missing data (Kluwin & Morris, 2006)

The ratings for the entire questionnaire, and for each of the subscales will be investigated between different groups of respondents for the 360-degree feedback instrument. These groups refer to the participants (who rated themselves), peers,

supervisors and subordinates. As already mentioned, the success of this type of instrument is based on the idea that people who most frequently have interaction with an individual are best suited and able to rate that particular individual (Hassan & Rohrbaugh, 2009). In accordance with the above, this study selected peers (colleagues), supervisors (managers) and subordinates as respondents owing to the fact that they obviously would have more frequent interaction with the people under study than other possible respondents such as customers.

The term “participants” refers to those individuals who are being rated by making use of the 360-degree feedback questionnaire, and the word “respondents” refers to the people who are rating the participants (Cook, 2010). Respondents in the database include managers, customers, direct reporters, subordinates, regional managers, seniors, internal clients, team members, colleagues, research executives, broker consultants, key brokers and others (Appendix 3). However, the current research focuses only on subordinates, peers and supervisors or managers. These groupings correspond with the groupings of respondents that are generally found in the literature on 360-degree feedback processes within organisations (Carson, 2006). Therefore, ratings from managers, subordinates, self-raters and colleagues were selected to correspond with these groupings. For practical purposes, responses from other parties were excluded. Furthermore, the terms “peers” and “colleagues” have been used interchangeably, with the label of “colleagues” being used throughout the presentation of results. This also applies to the terms “supervisors” and “managers” with “managers” being used as a constant in following chapters and sections. The term “supervisors” was consistently used in the literature for this study. For analysis purposes and presentation in following sections, participants are referred to as “self” meaning self-ratings from the data set.

A purposive sampling strategy was used to select the groups required within the existing data set (Whitley, 2002). Only participants who had a self-rating, as well as ratings from at least one peer, subordinate or manager were selected for the sample. The original data set that was sourced was comprised of 3 133 participants with self-ratings, and 22 662 respondents who had provided ratings for the participants in

question (Appendix 3). The selected data set was comprised of 21 862 individuals making up 90% of all participants and respondents. As mentioned above, ratings from managers, subordinates, self and colleagues were selected to correspond with the groupings proposed above. Through this selection, 19 074 respondents were retained from the original sample of 22 662 constituting 84.17% of the total respondents from the original data set (Appendix 3). Furthermore, only participants who had at least one rating from each selected respondent group were included in the final sample. Therefore, 2 788 participants and self-ratings were selected from the original sample of 3 133. This accounted for 88.99% of the participants. However, since no match could be found during the analysis of the data, some individuals may be lost. As is clear from Appendix 3, on average each respondent received between four and ten ratings from respondents. However, the total range of the number of respondents for majority of participants is from 0 to more than 20. The total sample group for the current study is reflected in Table 1. This refers to the individuals who were selected from the original data set described in Appendix 3.

Table 1. Summary of total participants and respondents of selected sample

Rater group	Frequency	Percentage	Cumulative frequency
Self	2788	12.75	2788
Colleague	7710	35.27	10498
Manager	2938	13.44	13436
Subordinate	8426	38.54	21862
Total (N)	**21862**	**90**	**-**

The data was obtained from various organisations in South Africa. However, details regarding the organisations are not available in the database. Therefore the researcher cannot identify the number of organisations that are included, or participants who are from the same organisation. Furthermore, no biographical information is available, and therefore the sample cannot be described in detail. This lack of descriptive data is a limitation of the study. However, owing to the fact that the questionnaire focuses on

leadership, it is assumed that all participants occupy positions at managerial level. The nature of the data allows the researcher to link participants to their respective respondents, but does not allow the researcher to identify whether participants are also respondents for other participants in the data set.

3.3 Research procedure

The researcher was invited to participate and conduct research as part of a broader validation study of a 360-degree feedback instrument. Thornhill Associates (Pty) Ltd, an organisation that is based in Johannesburg, developed this questionnaire. This organisation Thornhill Associates (Pty) Ltd) offers 360-degree feedback services online to both individuals and organisations in South Africa (Cook, 2010). The online delivery system is available to all individuals who have access to the Internet, and is online twenty-four hours a day (Cook, 2010). Furthermore, consultants and coaches are available to provide support to clients (Cook, 2010). The questionnaire and instructions given to participants and other respondents are identical, and are provided via an email message to participants explaining the process (Cook, 2010). Participants are then asked to forward this email to respondents with an invitation to provide feedback (Cook, 2010). Both participants and respondents complete the same questionnaire online arriving at the same landing page (Cook, 2010). Participants and respondents are required to rate the participant on a 7-point scale for each question, where 1= seldom effective, 4= effective and 7= always a strength.

Access to the data was obtained from the Managing Director of Thornhill Associates. The researcher verbally explained the aims, rationale and procedures of the research to him, and provided him with an official letter in order to confirm access to the archival data, and to obtain the conditions of using the data (Appendix 4). He then provided the researcher with a letter granting access to the archival data (Appendix 1). Taking the nature of archival data into account, the researcher did not have to approach participants directly or gain their individual consent. However, after the participants had completed the questionnaire, they gave permission to Thornhill Associates to use the data for research purposes (Cook, 2010). On completion of the study the results of the current research will therefore be made available to Thornhill Associates.

3.4 Measuring instruments

The measuring instrument that was used to collect the archival data for this study was the 360-degree feedback questionnaire (Leadership 3) (Appendix 2)(Cook, 2010). The questionnaire was developed by the Managing Director of Thornhill Associates in 2005, by utilising various theoretical sources (Cook, 2010). It was developed with the aim of accurately assessing and improving leadership performance and effectiveness in the context of 360-degree feedback (Cook, 2010). The questionnaire utilises the three-dimensional head-heart-and-hands model of leadership performance (Cook, 2010). This model was found to encompass key contributions from various established theories and models that the developer has encountered over time (Cook, 2010). According to the developer, the model that was used to develop the questionnaire accounts for all aspects of effective leadership (Cook, 2010). The "head" dimension refers to the concern for direction, the "heart" includes the human element of organisations such as interpersonal relationships, and the "hands" cover elements required for getting the job done (Cook, 2010). An example of items in the head dimension is: "It identifies the 'big-picture' strategic context of problems and opportunities". An item in the hands dimension reads: "Takes active steps immediately to ensure the team obtains the resources needed to do the job". The heart dimension comprises three subscales: enabling others, enabling self and enabling groups and the organisation (Appendix 2). The "enabling others" subscale deals with the interpersonal aspects of leading with emotional intelligence with items such as: "Actively builds on the benefits of diversity to draw the best out of all people" (Cook, 2010). The "enabling self" subscale deals with the personal aspects of leading with emotional intelligence with an example of an item being: "Is aware of his/her strengths and weaknesses, and is not afraid to disclose them" (Cook, 2010). The enabling groups and the organisation subscale deal with the effective use of influence and team-work management with an item such as: "Sums up decisions and puts what the team is thinking into words or action steps" (Cook, 2010). In addition to the head-heart-and- hands model of leadership performance, the questionnaire includes a leadership impact subscale dealing with the manner in which respondents feel and react to the participants' leadership (Cook, 2010). An example of one of the items in this leadership dimension is: "When I am with this person I feel

respected". Each of the items required respondents to rate participants on a 7-point scale for each question, where 1= seldom effective, 4= effective and 7= always a strength. Therefore, high scores indicate high performance within the respective dimensions and the questionnaire as a whole.

When developing questionnaires to be used as instruments of measurement, research indicates that enough theory needs to be provided and used as basis to explain item use and construction, as well as to explain the statistical information related to the instrument (Monroe, 1995). This ensures a clear idea of what is to be assessed with a complete, solid, theoretically-based and thorough definition of the construct to be measured and its operationalisation (Carretero-Dios & Pérez, 2007). Thus, in the construction of tests or instruments, theoretical evidence, together with the adoption of systematic criteria and procedures, should be presented by test authors (Carretero-Dios & Pérez, 2007). Furthermore, items must be logically and theoretically connected to the construct that is being measured (O'Leary-Kelly & Vokurka, 1998). For the effective construction of tests and measurement instruments, the test developer of the instrument that is used in this study, although grounded in extensive experience and in-depth knowledge of theories and constructs in this field, made more extensive use of an intuitive method compared with the more systematic method outlined above. The developer mentioned the utilisation of points and concepts from various theories and approaches, coupled with experience to design the questionnaire (Cook, 2010). However, this might not be rigorous enough and might have impacted upon the rigour and quality of the 360-degree feedback questionnaire (Leadership 3). The head-heart-and-hands model has not been found to be a theoretically established model on examination of the literature available for this study. However, on careful examination of test subscales and items, one can make links to concepts of transformational leadership. As mentioned in detail in the chapter on literature review, transformational leadership refers to a hybrid approach that encompasses the features of transactional leadership as well as a broader impact with the end goal of inspiring and empowering others (Alimo-Metcalfe, 1998). The components of this style include the following:

Charisma or inspirational motivation, referring to the creation of optimism and drive to mission or goals
Intellectual stimulation, referring to the encouragement of others to challenge ideas and seek feedback
Individualised consideration, referring to an active belief in others and their abilities (Alimo-Metcalfe, 1998; Kerin, 2010)

The “head” dimension appears to correlate with inspirational motivation of transformational leadership, where subordinates are driven to goal achievement. The “heart” component can be related to the charisma and individualised consideration of those that are being led. Furthermore, the encouragement and development of others also correlate with the enablement of others and groups or subscales of the instrument. The “hand” subscale corresponds with the transactional component of transformational leadership that addresses the completion of tasks and objectives. Lastly, the leadership impact subscale corresponds with one of the major outcomes of adopting a transformational leadership style. This outcome refers to a strong identification with the leader, and a shared vision that goes beyond self-interest (Kerin, 2010).

In addition to providing a measure of leadership performance, the ratings can also provide a method of investigating congruence between the different groups mentioned above. Comparisons of ratings between the various groups of raters for the entire questionnaire provide a measure of the degree of consistency. Furthermore, by comparing different respondents’ ratings for each subscale, areas of consistency or inconsistency can be identified.

There have been no previous research studies that have assessed the reliability and validity of the 360-degree feedback questionnaire (Leadership 3) (Cook, 2010). Evaluating the reliability of a questionnaire or instrument refers to assessing the consistency of a particular scale, where the same result is obtained each time it is applied to the same person (Rosenthal & Rosnow, 1991). Internal consistency refers to a component of reliability where a particular scale is assessed by comparing the correlational matrix for item scores to see which item scores correspond with each other, and to what extent this takes place (Rosenthal & Rosnow, 1991). High internal

consistency means that a particular participant should score similarly on related items reflected by a high Cronbach's alpha (α) reliability coefficient (Rosenthal & Rosnow, 1991). Higher alphas, being close to 1.00, reflect stronger relationships between the items or scales (Rosenthal & Rosnow, 1991). Cronbach coefficient alphas have been computed for the 360-degree feedback questionnaire (Leadership 3) as a whole and across all rater groups. These will be presented in the results chapter. Furthermore, this study will take an initial step towards a full validation that can be undertaken by other researchers. This will be covered in detail in the results section under the scale-evaluation subsection.

3.5 Analysis

The results were analysed by making use of quantitative methods. Summary descriptive statistics for each rater group were calculated. Summary statistics include calculating means, medians, modes, standard deviations and ranges with minimum and maximum values (Howell, 2004).

Evaluating the validity of an instrument or scale involves assessing the constructs the scale purports to measure (Rosenthal & Rosnow, 1991). A complete validation study on any particular scale would need to address various dimensions of validity including content, face, convergent and construct validity. Due to time constraints, this study focuses on construct validity. Construct validity refers to the degree to which a particular scale or measure supports predictions or hypotheses that are based on the theory of the construct purportedly being measured (Sireci, 2007). This theory refers to the head-heart-and-hands model on which the scale is based (Cook, 2010). From the definition outlined above, the fact that the underlying theory forms the base of establishing construct validity, corresponds with the aim of the validation aspect of the current research study. The reason for this being that the test developer utilised various theories and models to construct the instrument. Our purpose was to find out whether this approach was appropriate for constructing a valid test that accurately measures the desired construct of leadership performance. In order to establish validity effectively, the test items (also referred to as empirical indicators) must be logically and theoretically connected to the desired construct (O'Leary-Kelly & Vokurka, 1998). In order to do so

effectively, past research indicated that this was a crucial step within the construct validation process (O'Leary-Kelly & Vokurka, 1998). This provides further support for the choice of establishing construct validity in this study. Furthermore, past research has shown that if construct validity is not established, there is a greater risk of potential confounding of statistical results, and can therefore lower the credibility of the research findings and conclusions (O'Leary-Kelly & Vokurka, 1998).

The construct validity in this study is assessed by using factor analyses to explore the link between the structure of the 360-degree feedback questionnaire (Leadership 3) and the underlying theory and constructs. Furthermore, factor analyses were conducted for each group of respondents, and analysed. The factor analysis seeks to identify the hidden, underlying constructs by looking at clusters of items which are mathematically similar (Rosenthal & Rosnow, 1991). There are two main types of factor analysis, namely confirmatory and exploratory (Rosenthal & Rosnow, 1991). Both of these were conducted in this study. An exploratory factor analysis attempts to indentify the underlying dimensions or constructs of a particular scale, and can be used to summarise the data observed in a meaningful way (Bartholomew et al., 2008). In contrast, a confirmatory factor analysis seeks to confirm an existing hypothesis with an existing, pre-established factor structure by typically using structural equation modelling (SEM) (Joreskog & Sorbom, 1993). The SEM estimates the factor loadings for the empirical data that was obtained, and calculates the goodness of fit between the underlying theory, proposed structure and observed relationships (Bartholomew et al., 2008).

In the current study this procedure is used in conjunction with the exploratory factor analysis to confirm the factor structure and constructs proposed by the test developer as discussed in preceding paragraphs of this section. Unidimensionality is one of the important components of construct validation. This term refers to the existence of a single construct underlying a set of measures or items (O'Leary-Kelly & Vokurka, 1998). Two commonly used methods of establishing unidimensionality is to conduct exploratory and confirmatory factor analyses (O'Leary-Kelly & Vokurka, 1998). The exploratory factor analysis has been shown to be more useful during the early phases of

test or instrument construction or theory development, with the confirmatory factor analysis shown to be a stricter analytic technique (O'Leary-Kelly & Vokurka, 1998). By using both approaches, the one accounts for the other's inherent problems, and provides a balanced view of the construct validity of the instrument. Furthermore, utilising both analyses enhances the statistical rigour of findings of this study.

Statistics were used to decide whether parametric or non-parametric tests were to be used. If a particular data set meets the requirements of normality and interval data, parametric tests can be conducted (Huck, 2004). Therefore, in order to meet these requirements, parametric tests such as Pearson's Correlation Coefficients were used in this study. Pearson's Correlation Coefficients were calculated for each pair of group comparisons for the whole questionnaire, as well as for each subscale. Correlational hypothesis testing was also done on each set of comparisons. When reporting correlation coefficients, it should be noted that according to Huck (2004), correlation coefficients between 0 and .30 are weak, .30 to .50 are moderate, .50 to .70 are strong and .70 to 1.00 are very strong. However, these are guidelines and should be used as such.

Correlation coefficients were converted to *z* scores by using Fisher's *z* transformations so that the sampling distribution would be normal to make comparisons between different groups (Miles & Banyard, 2007). This was followed by significance testing using Fisher's *z* scores to compare two sets of correlations to identify whether there was a significant difference between the two sets (Miles & Banyard, 2007). The two sets of correlations were used for comparisons between participants and supervisors, participants and peers, participants and subordinates, subordinates and supervisors, subordinates and peers and supervisors and peers. Comparisons using the *z* scores yield *p* values of significance (Miles & Banyard, 2007). If $p<.05$ there is a significant difference between the two correlation coefficients that are being compared (Huck, 2004). In this study, a significant result would indicate a difference between the respective groups of raters on a particular subscale, or for the whole questionnaire.

In order to further investigate the variance between the groups of raters in the 360-degree feedback questionnaire, a hierarchical linear model (HLM) analysis was

conducted (Cramer, 2003). The HLM technique is useful to investigate relationships on a particular level and between different levels in an organisation (Wech & Heck, 2004). In this case the levels are represented by the different groups of raters within the sample of this study. The technique takes into account that individuals who are on a particular level or in a particular group may not yield independent observations due to their similarity(Wech & Heck, 2004).This similarity is a product of group membership(Wech & Heck, 2004). Thus, outcomes or results observed are explained for individuals in groups as a function of the group's characteristics as well as members' characteristics (Arnold, 1992). The use of the HLM also addresses the challenge of investigating relationships between groups with non-identical numbers and in some cases, "many-to-one" comparisons (Bartholomew et al., 2008). "Many-to-one" comparisons refer to groups that are being compared, where one group may have more individuals than one who are being compared with another group comprising one individual only. Through this analysis, the variance within groups, as well as between groups or levels, can be computed (Wech & Heck, 2004). Using these variances, the intraclass correlation coefficient (ICC) is then calculated. This represents the measure of variance that could potentially be explained by the Level-2 predictor, which indicates how much variance is explained by group or level membership compared with individual variance (Wech & Heck, 2004). The ICC is calculated by dividing the variance between groups by the total variance (Wech & Heck, 2004). If this coefficient is high, the variance can be attributed to group membership, and less so to individual differences (Wech & Heck, 2004). In the context of this study, a high ICC will indicate that variance or differences in ratings can rather be attributed to the rater group the respondent belongs to, than to individual differences.

3.6 Ethics

This study is low-risk, non-invasive and has no adverse psychological impact on the individuals who are represented in the archival data. The use of archival data has eliminated ethical issues regarding gaining individual consent, and procedural matters related to data collection within organisations. The researcher was not involved in the data-collection process, and had no access to any documentation derived from

participants, or concerning them. The data does not include any biographical or personal data such as the employee's number, name, position or age. Therefore the data is anonymous and does not contain any details that could be used to identify participants indirectly. Hence, ethical issues regarding anonymity can be considered as having been addressed.

As outlined in the access letter from Thornhill Associates (Appendix 1), all participants provided consent to use the data for research purposes at the time of the assessment. The database has been designed so that the researcher can link participants to their respective respondents, but not to identify whether participants are also respondents for other participants. Furthermore, any information that might identify participants was removed from the report, and will remain confidential.

Confirmation regarding access to the data was carried out through the presentation of a formal letter (Appendix 4), and was subsequently confirmed by Thornhill Associates, the organisation that owns the questionnaire and database (Appendix 1). Prior to sourcing of the data, ethics approval was requested and received from the Committee for Research on Human Subjects of the University of the Witwatersrand (Appendix 6). Results are reported in this research report, and a report will be forwarded to Thornhill Associates. Results may also be reported in a journal article or conference paper. The Managing Director of Thornhill Associates has indicated in writing that he is comfortable with the name of the questionnaire and Thornhill Associates being mentioned in reports. In the event of a manuscript being submitted for publication, an agreement was negotiated between the Managing Director, Fiona Donald (research supervisor) and myself (Appendix 5). Data will be stored in a secure place by the supervisor for a period of approximately 5 years to allow for the possibility of a publication. Nobody, other than my supervisor and I, will have access to the data.

4. Results

In this section, all findings with regard to the specified research hypotheses are included. The research hypotheses involve investigating the areas of consistency and inconsistency between ratings of different groups in the 360-degree feedback context, as well as the strength of these observed relationships. These relationships are presented for each group of respondents, namely self-raters, subordinates, colleagues or peers, and managers or supervisors. Instrument reliabilities and evaluation for the questionnaire is provided first. This is followed by distribution analyses which were conducted to ascertain the nature of the data. Correlation coefficients and significance-testing results follows, and includes comparisons by using Fisher's *z* transformations. Hierarchical linear model analyses of ratings between and within all rater groups concludes the chapter. Both research hypotheses are addressed simultaneously under each set of analysis results.

4.1 Questionnaire evaluation

The summarised reliabilities for all groups who answer the questionnaire are presented in Table 2. Cronbach's coefficient alpha is a function of the mean correlation of all the items of the instrument with one another, and is interpreted as a correlation coefficient (Whitley, 2002). This coefficient provides an indication of how consistent the scale is in measuring the intended attribute (Whitley, 2002). According to Whitley (2002), an alpha of .80 or higher is an indication of very high reliability; .70 to .79 is high; .60 to .69 is moderate; and .60 or less is weak, and should not be used. However, it should be noted that these are only guidelines.

Table 2. Questionnaire reliabilities for each group

Group	Cronbach coefficient alpha	
	Raw	Standardised
All Groups	.98	.98
Self	.96	.97
Manager	.97	.98
Colleague	.98	.98
Subordinate	.98	.98

According to Table 2, the 360-degree feedback questionnaire (Leadership 3) has a very high reliability with the Cronbach's alpha being greater than .90 for raw scores and greater than .90 for standardised scores for all groups of raters.

Exploratory factor analysis of the 360-degree feedback questionnaire (Leadership 3)

Exploratory factor analyses were conducted by using the principal components extraction method and orthogonal Varimax rotation for each group of raters. A six-factor solution was selected which corresponded with the six factors proposed by Dr Cook (2010) when designing the questionnaire of Head (creating direction), Heart 1(enabling others), Heart 2(enabling Self), Heart 3 (enabling groups and the organisation), Hands (executing), and Leadership impact. The subscales were grouped according to these proposed underlying factors theoretically (Cook, 2010). This was applicable across all groups. The six-factor solution was selected after having analysed the Eigen values greater than 1.00 according to Kaiser's criterion and scree plots indicating the number of factors at the position the plot starts to level off (Bartholomew et al., 2008). Eigen values reflect the degree of variance of each factor within a particular scale or instrument before rotation (Rosenthal & Rosnow, 1991). In order to determine the percentage of variance, the Eigen value can be divided by the total number of items, in this case 62 (Rosenthal & Rosnow, 1991). This value is presented under the column

headed proportion (which must be multiplied by 100 to get a percentage), where each subsequent factor will account for less variance than previous factors (Rosenthal & Rosnow, 1991). The column headed “cumulative” refers to the variance of all factors up to that point in the table. Looking at Table 3 below, the cumulative proportion for 6 factors together is .6045 (60.45%) accounting for 60.45% of the variance.

The Eigen values and scree plots of each group of raters are presented below. Table 3 presents the Eigen values for the self-rater group showing 10 factors that have values greater than 1.00, Table 4 presents the values for the colleague group with six factors, Table 5 presents the values for the manager rater group revealing eight factors, and Table 6 shows the Eigen values for the subordinate rater group showing five factors. Corresponding to these tables, the scree plots for each rater group are presented below. Figure 1 refers to the scree plot for the self-rater group, Figure 2 for the colleague rater group, Figure 3 for the manager rater group, and Figure 4 refers to the scree plot for the subordinate rater group.

Table 3. Eigen values for 360-degree feedback questionnaire (self) (Leadership 3)

Eigen values of the correlation matrix: Total = 62 Average = 1				
Item	**Eigen value**	**Difference**	**Proportion**	**Cumulative**
1	**19.7112373**	16.1042736	.3179	.3179
2	**3.6069637**	1.5392456	.0582	.3761
3	**2.0677181**	.1519646	.0334	.4095
4	**1.9157535**	.4334093	.0309	.4403
5	**1.4823442**	.1657409	.0239	.4643
6	**1.3166033**	.0503399	.0212	.4855
7	**1.2662634**	.1420480	.0204	.5059
8	**1.1242154**	.0151320	.0181	.5240
9	**1.1090834**	.0943811	.0179	.5419
10	**1.0147023**	.0791256	.0164	.5583

11	.9355767	.0312176	.0151	.5734
12	.9043591	.0389275	.0146	.5880
13	.8654316	.0562984	.0140	.6019
14	.8091332	.0277502	.0131	.6150
15	.7813831	.0232413	.0126	.6276
16	.7581418	.0089833	.0122	.6398
17	.7491585	.0154047	.0121	.6519
18	.7337538	.0230340	.0118	.6637
19	.7107198	.0350244	.0115	.6752
20	.6756954	.0036064	.0109	.6861
21	.6720891	.0172673	.0108	.6969
22	.6548218	.0203238	.0106	.7075
23	.6344980	.0028999	.0102	.7177
24	.6315981	.0054064	.0102	.7279
25	.6261918	.0134156	.0101	.7380
26	.6127761	.0229328	.0099	.7479
27	.5898433	.0027549	.0095	.7574
28	.5870884	.0155549	.0095	.7669
29	.5715335	.0052042	.0092	.7761
30	.5663293	.0093417	.0091	.7852
31	.5569876	.0203713	.0090	.7942
32	.5366164	.0029137	.0087	.8029
33	.5337027	.0047632	.0086	.8115
34	.5289395	.0055047	.0085	.8200
35	.5234348	.0108817	.0084	.8285
36	.5125531	.0200576	.0083	.8367
37	.4924954	.0067696	.0079	.8447
38	.4857258	.0084347	.0078	.8525
39	.4772911	.0085841	.0077	.8602
40	.4687070	.0088939	.0076	.8678
41	.4598131	.0074039	.0074	.8752

42	.4524092	.0133413	.0073	.8825
43	.4390679	.0060963	.0071	.8896
44	.4329716	.0044823	.0070	.8965
45	.4284893	.0111157	.0069	.9035
46	.4173736	.0052499	.0067	.9102
47	.4121237	.0084688	.0066	.9168
48	.4036549	.0102882	.0065	.9233
49	.3933667	.0012550	.0063	.9297
50	.3921117	.0053481	.0063	.9360
51	.3867636	.0074455	.0062	.9423
52	.3793181	.0155442	.0061	.9484
53	.3637739	.0040663	.0059	.9542
54	.3597076	.0071255	.0058	.9600
55	.3525822	.0154238	.0057	.9657
56	.3371584	.0020270	.0054	.9712
57	.3351314	.0146148	.0054	.9766
58	.3205166	.0054398	.0052	.9817
59	.3150769	.0118485	.0051	.9868
60	.3032284	.0333582	.0049	.9917
61	.2698702	.0258395	.0044	.9961
62	.2440307		.0039	1.0000

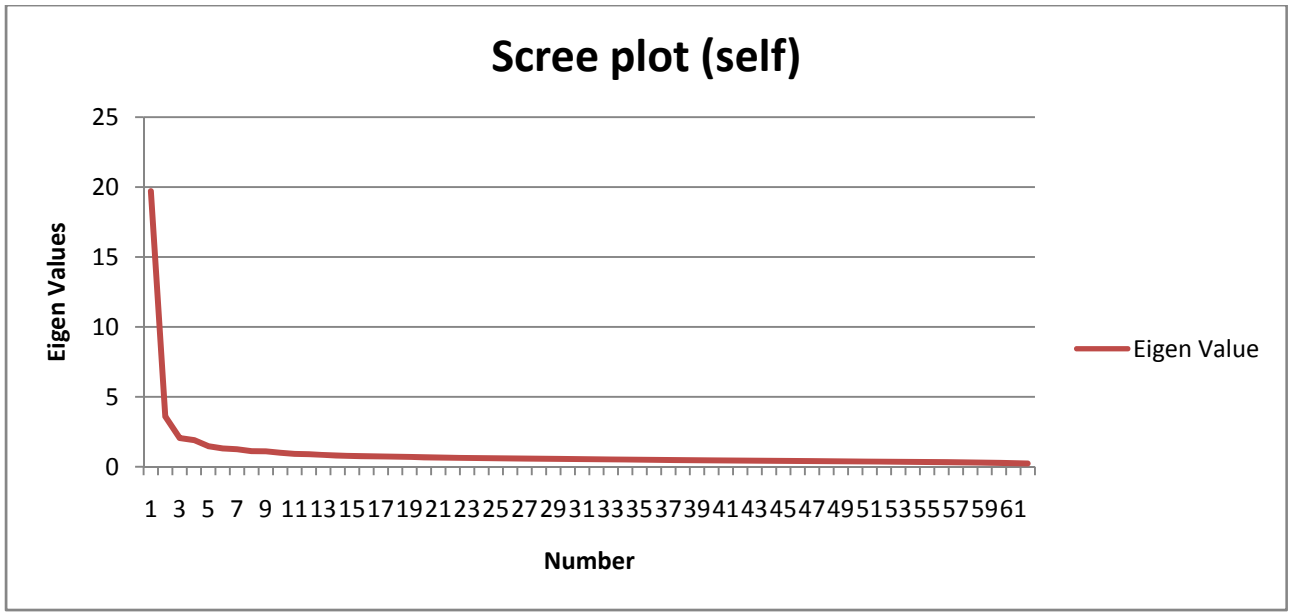


Figure 1. Scree plot of Eigen values for 360-degree feedback questionanire (self) (Leadership 3)

Table 4. Eigen values for 360-degree feedback questionnaire (colleague) (Leadership 3)

Eigen values of the correlation matrix: Total = 62 Average = 1				
	Eigen value	**Difference**	**Proportion**	**Cumulative**
1	**28.8977085**	25.7877980	.4661	.4661
2	**3.1099105**	1.2462749	.0502	.5163
3	**1.8636357**	.4983191	.0301	.5463
4	**1.3653165**	.2242918	.0220	.5683
5	**1.1410248**	.0424444	.0184	.5867
6	**1.0985803**	.1096532	.0177	.6045
7	.9889271	.0930443	.0160	.6204
8	.8958828	.0287410	.0144	.6349
9	.8671418	.1023964	.0140	.6488
10	.7647454	.0696245	.0123	.6612

11	.6951209	.0186250	.0112	.6724
12	.6764959	.0331284	.0109	.6833
13	.6433675	.0147356	.0104	.6937
14	.6286319	.0065493	.0101	.7038
15	.6220826	.0314436	.0100	.7138
16	.5906389	.0110025	.0095	.7234
17	.5796364	.0161981	.0093	.7327
18	.5634383	.0342971	.0091	.7418
19	.5291412	.0074989	.0085	.7503
20	.5216423	.0124765	.0084	.7588
21	.5091659	.0093304	.0082	.7670
22	.4998354	.0064563	.0081	.7750
23	.4933791	.0167210	.0080	.7830
24	.4766581	.0056823	.0077	.7907
25	.4709759	.0085041	.0076	.7983
26	.4624717	.0043820	.0075	.8057
27	.4580897	.0086681	.0074	.8131
28	.4494217	.0101887	.0072	.8204
29	.4392330	.0112193	.0071	.8275
30	.4280137	.0036750	.0069	.8344
31	.4243387	.0056074	.0068	.8412
32	.4187313	.0078793	.0068	.8480
33	.4108521	.0076701	.0066	.8546
34	.4031820	.0063841	.0065	.8611
35	.3967979	.0083514	.0064	.8675
36	.3884465	.0057810	.0063	.8738
37	.3826655	.0045612	.0062	.8799
38	.3781043	.0125291	.0061	.8860
39	.3655752	.0051290	.0059	.8919
40	.3604462	.0092128	.0058	.8977
41	.3512334	.0018742	.0057	.9034

42	.3493592	.0034709	.0056	.9090
43	.3458883	.0040654	.0056	.9146
44	.3418229	.0053992	.0055	.9201
45	.3364237	.0070882	.0054	.9256
46	.3293355	.0083356	.0053	.9309
47	.3209999	.0022205	.0052	.9360
48	.3187794	.0063285	.0051	.9412
49	.3124510	.0078469	.0050	.9462
50	.3046041	.0004845	.0049	.9511
51	.3041195	.0101006	.0049	.9560
52	.2940190	.0062257	.0047	.9608
53	.2877933	.0085309	.0046	.9654
54	.2792624	.0093746	.0045	.9699
55	.2698878	.0094006	.0044	.9743
56	.2604872	.0016549	.0042	.9785
57	.2588323	.0038701	.0042	.9827
58	.2549622	.0184809	.0041	.9868
59	.2364813	.0041942	.0038	.9906
60	.2322871	.0475944	.0037	.9943
61	.1846927	.0178682	.0030	.9973
62	.1668246		.0027	1.0000

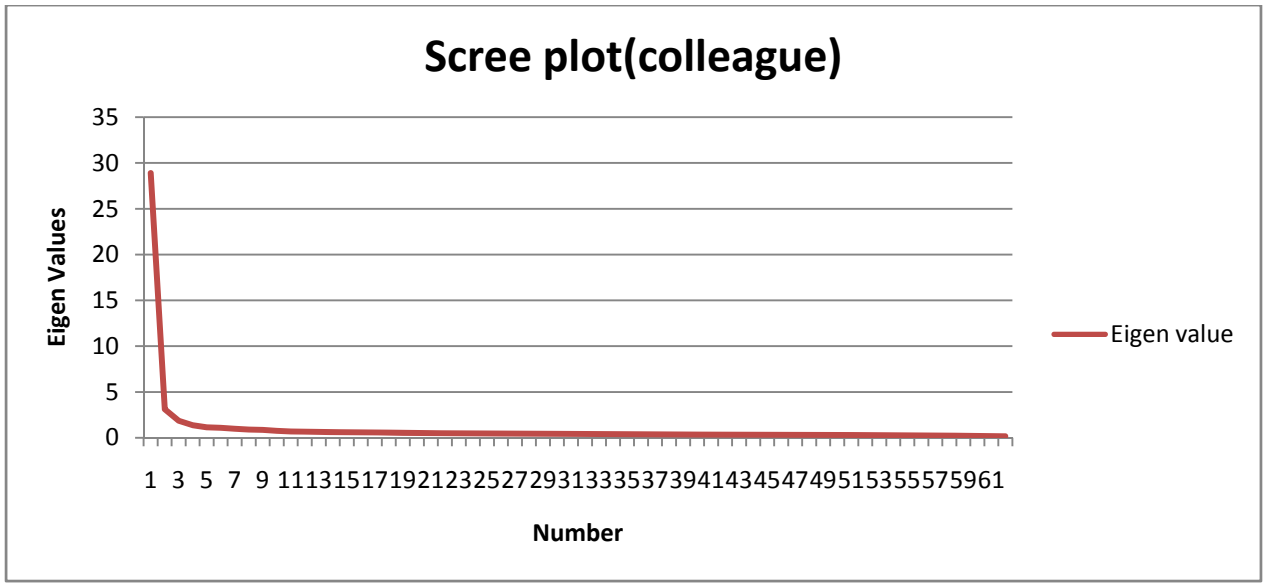


Figure 2. Scree plot of Eigen values for 360-degree feedback questionnaire (colleague) (Leadership 3)

Table 5. Eigen values for 360-degree feedback questionnaire (manager) (Leadership 3)

Eigen values of the correlation matrix: Total = 62 Average = 1				
	Eigen value	**Difference**	**Proportion**	**Cumulative**
1	**25.2519470**	21.8291034	.4073	.4073
2	**3.4228436**	1.1934481	.0552	.4625
3	**2.2293955**	.3181269	.0360	.4985
4	**1.9112687**	.5245211	.0308	.5293
5	**1.3867475**	.1350844	.0224	.5516
6	**1.2516632**	.0110528	.0202	.5718
7	**1.2406103**	.2150325	.0200	.5918
8	**1.0255779**	.0541219	.0165	.6084
9	.9714560	.1812237	.0157	.6241
10	.7902323	.0245173	.0127	.6368

11	.7657150	.0433836	.0124	.6492
12	.7223314	.0108502	.0117	.6608
13	.7114812	.0307968	.0115	.6723
14	.6806844	.0124931	.0110	.6833
15	.6681913	.0300860	.0108	.6940
16	.6381053	.0206851	.0103	.7043
17	.6174202	.0019292	.0100	.7143
18	.6154909	.0310226	.0099	.7242
19	.5844684	.0058514	.0094	.7336
20	.5786170	.0073898	.0093	.7430
21	.5712272	.0141865	.0092	.7522
22	.5570407	.0233052	.0090	.7612
23	.5337355	.0122535	.0086	.7698
24	.5214820	.0158667	.0084	.7782
25	.5056153	.0068383	.0082	.7863
26	.4987770	.0089222	.0080	.7944
27	.4898548	.0051465	.0079	.8023
28	.4847084	.0090289	.0078	.8101
29	.4756794	.0059248	.0077	.8178
30	.4697546	.0222184	.0076	.8254
31	.4475361	.0019745	.0072	.8326
32	.4455617	.0086949	.0072	.8398
33	.4368668	.0039074	.0070	.8468
34	.4329593	.0092527	.0070	.8538
35	.4237066	.0046971	.0068	.8606
36	.4190095	.0049383	.0068	.8674
37	.4140712	.0030685	.0067	.8741
38	.4110027	.0084954	.0066	.8807
39	.4025073	.0087954	.0065	.8872
40	.3937119	.0091215	.0064	.8935
41	.3845904	.0058334	.0062	.8997

42	.3787570	.0084226	.0061	.9058
43	.3703344	.0128222	.0060	.9118
44	.3575122	.0035120	.0058	.9176
45	.3540002	.0140942	.0057	.9233
46	.3399060	.0103038	.0055	.9288
47	.3296023	.0054279	.0053	.9341
48	.3241743	.0081311	.0052	.9393
49	.3160432	.0071643	.0051	.9444
50	.3088790	.0030621	.0050	.9494
51	.3058168	.0025907	.0049	.9543
52	.3032261	.0142729	.0049	.9592
53	.2889533	.0053805	.0047	.9639
54	.2835727	.0128964	.0046	.9685
55	.2706763	.0045444	.0044	.9728
56	.2661319	.0060427	.0043	.9771
57	.2600893	.0095240	.0042	.9813
58	.2505653	.0040012	.0040	.9854
59	.2465641	.0061440	.0040	.9893
60	.2404200	.0178975	.0039	.9932
61	.2225226	.0239151	.0036	.9968
62	.1986074		.0032	1.0000

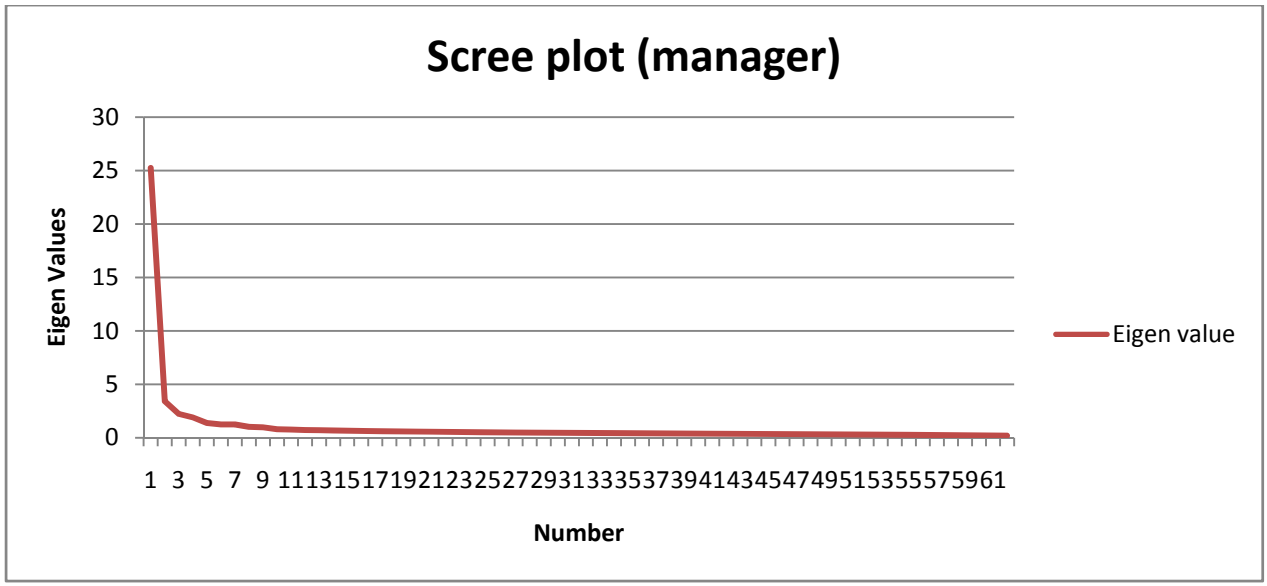


Figure 3. Scree plot of Eigen values for 360-degree feedback questionnaire (manager) (Leadership 3)

Table 6. Eigen values for 360-degree feedback questionnaire (subordinate) (Leadership 3)

Eigen values of the correlation matrix: Total = 62 Average = 1				
	Eigen value	**Difference**	**Proportion**	**Cumulative**
1	**31.8114628**	28.9180173	.5131	.5131
2	**2.8934456**	1.3081315	.0467	.5598
3	**1.5853141**	.3682004	.0256	.5853
4	**1.2171137**	.2165173	.0196	.6050
5	**1.0005963**	.0973475	.0161	.6211
6	.9032489	.0581595	.0146	.6357
7	.8450894	.0645408	.0136	.6493
8	.7805486	.0226328	.0126	.6619
9	.7579159	.0386893	.0122	.6741
10	.7192266	.0756378	.0116	.6857
11	.6435888	.0344469	.0104	.6961
12	.6091419	.0079666	.0098	.7059

13	.6011753	.0200320	.0097	.7156
14	.5811433	.0159138	.0094	.7250
15	.5652295	.0234774	.0091	.7341
16	.5417521	.0120719	.0087	.7428
17	.5296801	.0065464	.0085	.7514
18	.5231337	.0308818	.0084	.7598
19	.4922520	.0101490	.0079	.7678
20	.4821030	.0099850	.0078	.7755
21	.4721180	.0103884	.0076	.7831
22	.4617296	.0072000	.0074	.7906
23	.4545296	.0172844	.0073	.7979
24	.4372452	.0070649	.0071	.8050
25	.4301803	.0034245	.0069	.8119
26	.4267559	.0067516	.0069	.8188
27	.4200042	.0110965	.0068	.8256
28	.4089078	.0024028	.0066	.8322
29	.4065050	.0075802	.0066	.8387
30	.3989248	.0090531	.0064	.8452
31	.3898716	.0036353	.0063	.8515
32	.3862364	.0074727	.0062	.8577
33	.3787637	.0051115	.0061	.8638
34	.3736522	.0046700	.0060	.8698
35	.3689822	.0042450	.0060	.8758
36	.3647372	.0018461	.0059	.8817
37	.3628911	.0108821	.0059	.8875
38	.3520091	.0071033	.0057	.8932
39	.3449057	.0016950	.0056	.8987
40	.3432107	.0125447	.0055	.9043
41	.3306659	.0017236	.0053	.9096
42	.3289423	.0028768	.0053	.9149
43	.3260656	.0054905	.0053	.9202

44	.3205751	.0035374	.0052	.9253
45	.3170376	.0082941	.0051	.9305
46	.3087436	.0062476	.0050	.9354
47	.3024960	.0064527	.0049	.9403
48	.2960433	.0016586	.0048	.9451
49	.2943847	.0067711	.0047	.9498
50	.2876136	.0023219	.0046	.9545
51	.2852917	.0048683	.0046	.9591
52	.2804234	.0105262	.0045	.9636
53	.2698972	.0086263	.0044	.9680
54	.2612709	.0056599	.0042	.9722
55	.2556110	.0063805	.0041	.9763
56	.2492305	.0084687	.0040	.9803
57	.2407618	.0070051	.0039	.9842
58	.2337567	.0143071	.0038	.9880
59	.2194496	.0137874	.0035	.9915
60	.2056622	.0338837	.0033	.9948
61	.1717785	.0228012	.0028	.9976
62	.1489773		.0024	1.0000

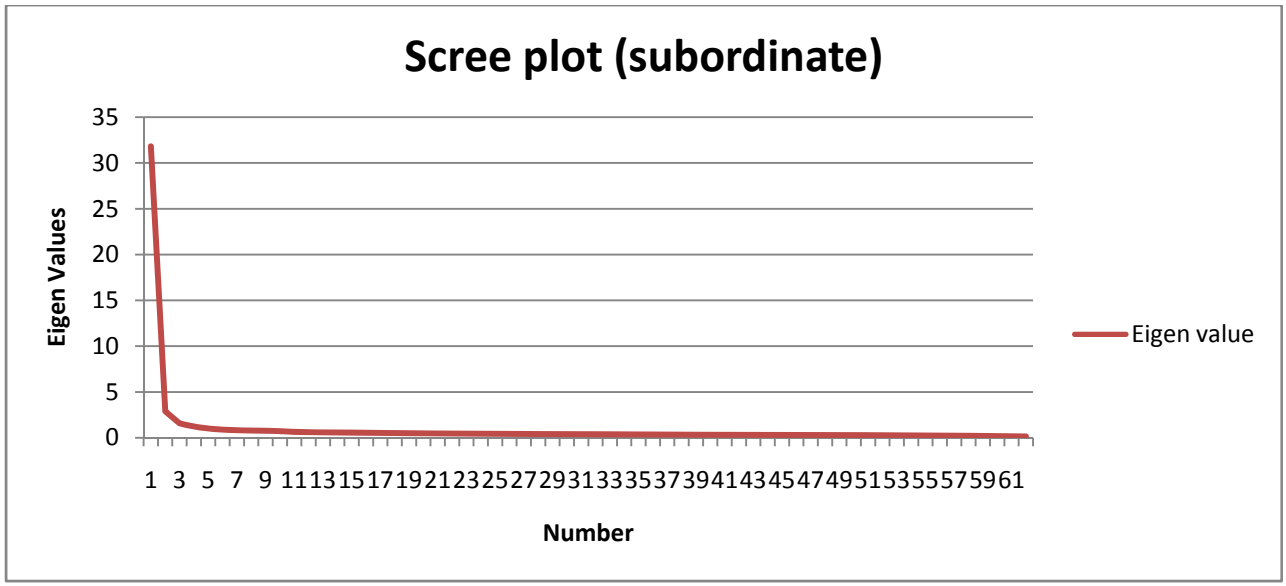


Figure 4. Scree plot of Eigen values for 360-degree feedback questionanire (subordinate) (Leadership 3)

Comparing Tables 3 to 6, as well as Figures 1 to 4, it is clear that there are some observed discrepancies between the number of principal components revealed by Eigen values greater than 1 and scree plots. Table 3 and Figure 1 reveals 10 factors according to Kaiser's criterion and 6 factors from the scree plot for self-ratings, Table 4 and Figure 2 revealed 6 factors for colleagues, Table 5 and Figure 3 revealed 8 factors according to the Eigen values and 6 factors from the scree plot for managers, and Table 6 and Figure 4 revealed 5 factors for subordinates. However, when averaged across the entire sample of raters, approximately 6 factors were obtained. Therefore, a six-factor solution was selected and supported. Therefore, looking at the analyses for each rater group and comparing them sheds light onto the structure of the questionnaire for all groups which can be inferred.

The rotated factor loadings for each group of raters are presented in the tables below, followed by analyses of these results with recommendations for possible future questionnaire modification. As the data used in this study was archival in nature, this scale was not modified and retested, but rather examined to determine a measure of construct validity.

The factor pattern for the sample of self-ratings revealed the variance explained by the six factors (rotated) to be 43.81% (Factor 1 = 7.05%; Factor 2 = 6.66%; Factor 3 = 5.81%; Factor 4 = 4.31%; Factor 5 = 3.21%; Factor 6 = 3.06%). The results of the six-factor solution (after Varimax rotation) are shown in Table 7 below with loadings of .40 or greater in bold. For self-ratings, all items which loaded on Factor 1 were classified under “Head”, all items loading on Factor 2 were classified under “Leadership” (on scale-leadership impact), all items loading on Factor 3 were classified under “Heart 2”, all items loading on Factor 4 were classified under “Heart 1”, all items loading on Factor 5 were classified under “Hands”, and all items loading on Factor 6 were classified under “Heart 3”. These classifications were assigned according to analysis of the factor loadings, as well as having taken the existing groupings by the test developer into account.

Table 7. Factor loadings – Varimax rotation (self)

Head items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
3	**.42945**	**.44330**	-.02058	.16326	.01495	.01736
7	**.58780**	.13446	.11895	.09551	.05860	.04366
9	**.51706**	.29664	.01713	-.09764	.30975	.05600
10	**.47376**	.38093	.18906	.18387	-.00361	.11731
13	**.58099**	.30173	.11720	.22200	-.02959	.01643
19	**.55300**	.17221	.28752	.12279	-.02313	.07082
20	**.60817**	.06913	-.02270	.07594	.08678	.29462
22	**.63390**	.04230	.07294	.19793	.10015	.21086
23	**.56816**	.11509	.08740	.05540	.24998	.10122
24	**.64242**	.17768	.18891	.09382	-.00224	-.00226
26	**.55944**	.20714	.23458	.12242	.13735	.08092
35	**.53001**	.05765	.03508	-.03257	.26980	.36100
36	**.45451**	.01758	.11057	.19246	.07418	.34828
45	**.50377**	.30671	.30076	.18584	.03414	-.04516
53	**.44040**	**.51737**	.07099	.15039	.15281	-.01739

56	**.46795**	**.51309**	-.01218	.08467	.12969	.02927
58	**.43003**	**.62199**	-.01983	.05493	.12619	-.03780
62	**.44508**	**.51924**	.01066	-.08086	.31695	.01934

Heart 1 items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
2	.20273	.02194	.22677	**.45106**	.18304	.18835
6	.07784	.11371	.11110	**.73430**	.00221	.20867
14	.22061	.20506	.16126	**.52270**	.31947	.00362
15	.05931	.06788	.25957	**.69630**	.10558	.21119
21	.37012	.17827	.13307	**.52073**	.17563	.08775
27	.28009	.17037	.21719	**.44967**	.01077	.25888
39	.25346	.17870	.22133	**.41664**	.22403	.03257

Heart 2 items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
1	-.03408	.20716	**.45840**	.20361	.17250	.27892
25	.16073	.16739	**.65098**	.08005	.07612	.15786
28	.11887	.11205	**.69027**	.19155	.20841	.09863
29	.05384	.06797	**.67953**	.23058	.05317	.11691
30	.07655	.07238	**.60510**	.26952	.25771	-.00941
32	.20792	.36509	**.42712**	.20136	.00898	.35144
40	.23936	.01102	**.63557**	.11779	.23577	.10309
42	.28975	.31256	**.42854**	.21334	.18807	.10340
44	.17074	.13712	**.57329**	.14441	.18095	.20471
49	.13580	.16442	**.66790**	.08564	-.00729	.23215

Heart 3 items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**

12	.22510	.19698	.28305	.15566	-.09978	**.45684**
31	.08951	.29172	.23714	.12786	.31851	**.45050**
37	.04069	.15199	.33230	.31957	.14010	**.54042**
38	.18898	.08801	.24329	.19177	.35566	**.48426**
43	.23340	.30106	.26728	.10361	.16171	**.54525**

Hands items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
8	.36759	.18618	.08409	.22898	**.45876**	.10018
11	.03024	.18299	.29813	**.40144**	**.45279**	.05024
17	.10961	.10595	.26118	.19434	**.64189**	.02792
46	.25510	.14371	.18751	.10833	**.61915**	.12411
52	.11341	**.43814**	.16781	.32683	**.42940**	.08008

Leadership items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
3	**.42945**	**.44330**	-.02058	.16326	.01495	.01736
18	.33520	**.44093**	.34822	.08577	.04516	.24231
41	.27985	**.49190**	.22969	.12676	.08648	.11113
51	.12179	**.52427**	.21009	.14415	.20312	.28845
52	.11341	**.43814**	.16781	.32683	**.42940**	.08008
53	**.44040**	**.51737**	.07099	.15039	.15281	-.01739
55	.26263	**.58979**	.09992	.04526	-.04391	.22759
56	**.46795**	**.51309**	-.01218	.08467	.12969	.02927
57	.02791	**.49055**	.29847	.34179	.11172	.10061
58	**.43003**	**.62199**	-.01983	.05493	.12619	-.03780
59	.18307	**.67875**	.13937	.22441	.27646	.11900
60	.09200	**.65663**	.17524	.15741	.14782	.34567
61	.13112	**.65760**	.23064	.17129	.22485	.18962
62	**.44508**	**.51924**	.01066	-.08086	.31695	.01934

Based on Table 7, Items 4, 5, 16, 33, 34, 47, 48, 50, and 54 do not load onto any factor, and should be removed in future versions of the instrument. Items 3, 52, 53, 56, 58 and 62 are cross-loading items, and can either be modified by rewording them or alternatively can be removed.

The rotated factor pattern for the sample of colleagues revealed the variance explained by the six factors (rotated) to be 37.47% (Factor 1 = 9.31%; Factor 2 = 8.16%; Factor 3 = 7.99%; Factor 4 = 4.71%; Factor 5 = 3.73%; Factor 6 = 3.57%). The results of the six-factor solution (after Varimax rotation) are shown in Table 8 below with loadings of .40 or greater in bold. For Colleagues: all items which loaded on Factor 1 were classified under “Hands”; all items loading on Factor 2 were classified under “Leadership” (on scale-leadership impact); all items loading on Factor 3 were classified under ‘Head’; all items loading on Factor 4 were classified under “Heart 3”; all items loading on Factor 5 were classified under “Heart 2”; and all items loading on Factor 6 were classified under “Heart 1”. These classifications were assigned according to analysis of the factor loadings, as well as taking into account the existing groupings by the test developer.

Table 8. Factor loadings – Varimax rotation (colleague)

			Head items			
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
7	.10366	.31208	**.55525**	.32580	.12389	.09674
8	.12466	.29096	**.48057**	.13634	.28097	**.42669**
9	-.02307	.33624	**.55584**	.32780	.01312	.28329
10	.19200	.26309	**.41419**	**.54370**	.15987	.13243
13	.17522	.27381	**.49334**	**.49729**	.14524	.05420
19	.31896	.26292	**.51696**	.34534	.14809	.00346
20	.21189	.21482	**.67194**	.19745	.13166	-.01399
22	.22652	.23710	**.67146**	.15910	.16077	.11906
23	.17987	.23335	**.64083**	.14238	.10768	.25579
24	.21320	.24333	**.56806**	.35471	.06306	.10336

26	.25587	.30175	**.54459**	.31062	.09565	.14583
35	.24740	.13245	**.70490**	.01963	.01463	.16972
36	.30773	.26975	**.57640**	.06414	.16902	.07433
45	.32517	.25679	**.41009**	.44767	.12172	.13936
50	.28124	.35265	**.41507**	.28013	.06055	.37100
56	.18360	**.64626**	**.41226**	.16295	.13154	.11549
62	.06932	**.67030**	**.43296**	.14401	.06685	.20625

Heart 1 items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
8	.12466	.29096	**.48057**	.13634	.28097	**.42669**
11	.38250	.23766	.06829	.20140	.39262	**.46272**
17	.31961	.26419	.18756	.15075	.17776	**.65170**
46	.28945	.30092	.34198	.11380	.09967	**.57869**

Heart 2 items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
2	.29357	.25549	.32997	.09011	**.46010**	.21002
6	.28090	.16116	.08428	.14634	**.72440**	.09978
14	.30440	.13751	.21789	.26342	**.46056**	.36730
15	**.40469**	.15864	.11315	.05932	**.69160**	.13624
21	.29154	.22465	**.44201**	.18527	**.41341**	.24285
27	.37636	.21449	.38576	.19646	**.42446**	-.01978

Heart 3 items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
3	.07799	.38715	.32325	**.51220**	.15339	.10155
4	.29501	.21433	.12261	**.46557**	.39060	.29161
10	.19200	.26309	**.41419**	**.54370**	.15987	.13243

13	.17522	.27381	**.49334**	**.49729**	.14524	.05420
16	.36426	.23434	.21114	**.46455**	.27823	.18437
34	.32596	.15009	.24769	**.44516**	.21502	.26343
41	.30420	.31627	.34593	**.48001**	.02846	.24978
45	.32517	.25679	**.41009**	**.44767**	.12172	.13936

Hands items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
1	**.47337**	.24200	.04135	.27636	.26793	.26784
12	**.52197**	.11119	.29576	.24472	.25414	-.12626
15	**.40469**	.15864	.11315	.05932	**.69160**	.13624
25	**.61647**	.20198	.16086	.23412	.13912	.15626
28	**.61607**	.20347	.20062	.12519	.26633	.27306
29	**.64174**	.18668	.08321	.10901	.28462	.05745
30	**.54563**	.12747	.10395	.15125	.28364	.34754
31	**.51690**	.31643	.12124	.16538	.03791	.36528
32	**.56223**	.27010	.18591	.36088	.20475	.09641
37	**.62974**	.20951	.17843	.05269	.26801	.05430
38	**.54706**	.20804	.33680	-.00320	.16527	.26503
40	**.56064**	.15667	.33527	.04392	.21096	.32977
42	**.47310**	.22118	.32738	.30652	.20006	.27573
43	**.55126**	.31828	.29378	.21328	.07047	.10696
44	**.58808**	.19040	.23696	.09052	.19553	.28188
47	**.52436**	.26250	.34851	.31633	.08647	.18431
48	**.52469**	.27432	.24417	.33500	-.08623	.05541
49	**.65601**	.23497	.24492	.17012	.15691	.12877
54	**.41239**	**.53075**	.07979	.14128	.27853	.25524

Leadership items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**

51	.38104	**.58688**	.21618	.17848	.14080	.18705
52	.33929	**.60731**	.21332	.12329	.23237	.30274
53	.18081	**.66699**	.32397	.15092	.13022	.13209
54	**.41239**	**.53075**	.07979	.14128	.27853	.25524
55	.28277	**.61464**	.29210	.29368	.01932	-.01312
56	.18360	**.64626**	**.41226**	.16295	.13154	.11549
57	.38607	**.61030**	.17535	.17710	.25011	.13758
58	.12714	**.72554**	.36839	.16447	.11930	.11134
59	.30695	**.72866**	.24703	.20003	.18928	.17739
60	.36222	**.70140**	.20560	.24079	.14915	.10709
61	.35158	**.68012**	.20814	.22970	.18723	.20674
62	.06932	**.67030**	**.43296**	.14401	.06685	.20625

Looking at Table 8, Items 5, 18, 33 and 39 do not load onto any factor, and should be removed in future versions of the instrument. Items 8, 10, 13, 15, 45, 54, 56 and 62 are cross-loading items, meaning that they load onto more than 1 factor (Rosenthal & Rosnow, 1991). These items can either be modified by rewording them, or can alternatively be removed.

The factor pattern for the sample of managers revealed the variance explained by the six factors (rotated) to be 35.46% (Factor 1 = 7.83%; Factor 2 = 7.51%; Factor 3 = 6.01%; Factor 4 = 5.17%; Factor 5 = 4.66%; Factor 6 = 4.27%). The results of the six-factor solution (after Varimax rotation) are shown in Table 9 below with loadings of .40 or greater in bold. For managers: all items which loaded on Factor 1 were classified under “Leadership”, all items loading on Factor 2 were classified under “Heart 2” (on scale-leadership impact), all items loading on Factor 3 were classified under “Head”, all items loading on Factor 4 were classified under “Hands”, all items loading on Factor 5 were classified under “Heart 3”, and all items loading on Factor 6 were classified under “Heart 1”. These classifications were assigned according to analysis of the factor loadings, as well as taking into account the existing groupings by the test developer.

Table 9. Factor loadings – Varimax rotation (manager)

Head items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
1	.31606	.05282	**.45519**	.29446	.32402	.09253
2	.15544	.35616	**.53010**	.17786	.16863	.01291
4	.25943	.06421	**.46481**	.16990	.23172	.39934
6	.08210	.06799	**.72148**	.07911	.19873	.16946
8	.30625	**.49988**	**.42500**	.18331	-.02521	.10202
11	.32456	.05175	**.61641**	.28164	.02261	.15158
14	.17118	.09415	**.55069**	.32154	.14829	.35052
15	.09650	.09829	**.73480**	.19840	.18037	.11392
17	**.45489**	.16245	**.51502**	.30056	-.06748	.08335
21	.17105	.33599	**.48554**	.23569	.07902	.32302
31	.48723	.08552	**.41102**	.14364	.28845	-.01425
37	.17335	.05177	**.42849**	.25283	.55310	.01561
38	.31221	.35032	**.46269**	.14741	.24608	-.14241
46	**.50719**	.29755	**.44493**	.22661	-.03817	.07386

Heart 1 items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
3	.28288	.33561	.08570	-.08022	.30696	**.46532**
10	.21319	**.45680**	.09196	.12356	.24580	**.46928**
13	.23211	**.42699**	.15590	.04880	.30616	**.46741**
16	.21714	.18568	.34039	.27169	.24729	**.43272**
24	.15563	**.52741**	.02615	.18408	.18023	**.43729**
34	.21985	.15327	.29110	.22621	.07717	**.58629**
41	.30151	.32083	.18910	.16900	.25841	**.48273**
45	.16820	.32640	.14069	.23395	.22378	**.57392**

Heart 2 items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
7	.24277	**.57244**	.06583	.11345	.16639	.12847
8	.30625	**.49988**	**.42500**	.18331	-.02521	.10202
9	.26340	**.64921**	.03151	-.03096	.02579	.27266
10	.21319	**.45680**	.09196	.12356	.24580	**.46928**
19	.14248	**.50885**	.05597	.23550	.39084	.26371
20	.10993	**.71928**	.04522	.08657	.26774	.15951
22	.16049	**.69758**	.16674	.14844	.12859	.11173
23	.20699	**.64621**	.14811	.28134	-.00112	.12214
24	.15563	**.52741**	.02615	.18408	.18023	.43729
26	.32794	**.56209**	.12502	.14154	.15014	.23409
35	.08263	**.71814**	.09996	.10360	.15723	-.00970
36	.19616	**.59195**	.23288	.14750	.19593	.09207
62	**.65248**	**.44043**	.09121	.04097	-.01083	.14394

Heart 3 items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
5	.18171	.30620	.19669	.17481	**.54524**	.22914
12	-.00092	.25846	.19226	.30223	**.49785**	.18708
18	.30983	.26541	.12475	.21080	**.52663**	.35543
32	.23646	.17741	.19922	.32244	**.54349**	.25032
37	.17335	.05177	**.42849**	.25283	**.55310**	.01561
43	.26934	.27539	.22505	.14336	**.57361**	.05113
48	.24371	.16362	.01366	.19943	**.49744**	.31466
55	**.50939**	.27163	.00903	.07657	**.44071**	.19545
60	**.68287**	.11757	.17901	.13061	**.41568**	.11822

Hands items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**

25	.24119	.15940	.17639	**.58274**	.32130	.03513
28	.26888	.19999	.29699	**.64999**	.17864	.07634
29	.16249	.03784	.24419	**.61264**	.25274	.10197
30	.26036	.08104	.39040	**.52511**	.03608	.22378
40	.14716	.28079	.23750	**.70100**	.05962	.12688
42	.20410	.26291	.31948	**.42404**	.23470	.34631
49	.16367	.18036	.14280	**.62035**	.33163	.16836

Leadership items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
17	**.45489**	.16245	**.51502**	.30056	-.06748	.08335
31	**.48723**	.08552	**.41102**	.14364	.28845	-.01425
33	**.43444**	.38392	.22783	.24889	.16312	.28151
46	**.50719**	.29755	**.44493**	.22661	-.03817	.07386
50	**.40143**	.39985	.27914	.17141	.05042	.32884
51	**.59089**	.19569	.22973	.27518	.24751	.10056
52	**.57248**	.24212	.31690	.28443	.11003	.19983
53	**.61062**	.30569	.18313	.16401	.14480	.19311
54	**.55570**	-.01001	.31175	.29377	.26863	.11663
55	**.50939**	.27163	.00903	.07657	**.44071**	.19545
56	**.58765**	.39227	.15088	.17559	.11278	.20024
57	**.54876**	.11784	.22695	.30896	.25899	.22812
58	**.67409**	.34905	.08607	.03854	.11808	.23393
59	**.70571**	.24039	.16859	.19583	.20850	.24011
60	**.68287**	.11757	.17901	.13061	**.41568**	.11822
61	**.68592**	.22559	.22518	.25769	.20910	.16141
62	**.65248**	.44043	.09121	.04097	-.01083	.14394

Looking at Table 9, Items 27, 39 and 47 do not load onto any factor, and should be removed in future versions of the instrument. Items 8, 17, 24, 31, 37, 46, 55, 60 and 62

are cross-loading items, and can either be modified by rewording them, or can alternatively be removed.

The factor pattern for the sample of subordinates revealed the variance explained by the six factors (rotated) to be 39.41% (Factor 1 = 10.62%; Factor 2 = 9.89%; Factor 3 = 9.59%; Factor 4 = 4.15%; Factor 5 = 3.92%; Factor 6 = 1.24%). The results of the six-factor solution (after Varimax rotation) are shown in Table 10 below with loadings of .40 or greater in bold. For subordinates: all items which loaded on Factor 1 were classified under “Hands”, all items loading on Factor 2 were classified under “Head” (on scale-leadership impact), all items loading on Factor 3 were classified under “Leadership”, all items loading on Factor 4 were classified under “Heart 3”, all items loading on Factor 5 were classified under “Heart 2”, and all items loading on Factor 6 were classified under “Heart 1”. These classifications were assigned according to analysis of the factor loadings, as well as taking into account the existing groupings by the test developer.

Table 10. Factor loadings – Varimax rotation (subordinate)

			Head items			
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
3	.18246	**.45064**	**.44872**	.37276	.06347	.03727
7	.16499	**.62506**	.32451	.23890	.13390	-.03092
8	.18492	**.53257**	.29560	.20159	.33247	.32003
9	.18115	**.60907**	.34545	.13545	-.00662	.26021
10	.30554	**.50820**	.34466	.32363	.17313	-.00603
13	.23699	**.56755**	.40772	.25305	.12803	-.01752
18	**.42935**	**.46190**	.38090	.36111	.08013	-.04090
19	.29226	**.60940**	.26788	.22027	.16301	-.08124
20	.20913	**.69971**	.22642	.09012	.14894	.00833
21	.28938	**.43133**	.34116	.38427	.27445	.12013
22	.25630	**.63898**	.24556	.20143	.24156	.09708
23	.25206	**.58962**	.24937	.22134	.14372	.19963
24	.22777	**.60385**	**.40174**	.14970	.06622	-.05588

26	.30248	**.57761**	.35506	.23401	.15664	-.00683
33	**.42798**	**.45619**	.36084	.38644	.09166	.08189
35	.34960	**.64542**	.14157	-.09836	.09296	.25510
36	.30668	**.58973**	.23355	.07886	.23424	.03634
41	**.40987**	**.46210**	**.46068**	.15660	.08456	.05746
45	.28410	**.50950**	**.41267**	.25188	.12641	-.08307
46	.38078	**.41042**	.32885	.17850	.22941	.34294
50	.39515	**.47939**	**.40961**	.17660	.11967	.21165
53	.15638	**.40875**	**.69679**	.11437	.12208	.04985
55	.26519	**.40264**	**.63416**	.13291	.13659	-.11027
56	.15446	**.43945**	**.66518**	.15589	.17657	.02899
62	.16296	**.44867**	**.68075**	.12665	.09048	.15621

Heart 1 items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
17	.41988	.22885	.25441	.40165	.21699	**.42722**

Heart 2 items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
2	.27063	.33337	.28308	.14840	**.54786**	.11163
6	.31578	.13418	.15502	.17603	**.71501**	-.00415
15	.39329	.12378	.17487	.15553	**.69046**	.07435
27	.30296	.39920	.23743	.24571	**.46646**	-.07554

Heart 3 items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
4	.27680	.25901	.29664	**.53900**	.29730	.07378
14	.26394	.28105	.25022	**.47002**	.39899	.18557
17	**.41988**	.22885	.25441	**.40165**	.21699	**.42722**

30	**.51786**	.17973	.19560	**.45745**	.25810	.02165
34	.28025	.34596	.25731	**.51607**	.15316	-.01316
39	**.43301**	.28468	.21957	**.44019**	.19977	.16849

Hands items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
1	**.50921**	.16328	.26755	.24362	.31659	.13237
5	**.47461**	.38282	.27219	.10832	.21722	-.11854
11	**.46848**	.13520	.23564	.36564	.33021	.20704
12	**.56273**	.31226	.13699	.07579	.29472	-.20322
16	**.42820**	.30616	.31230	.33378	.18312	-.07126
17	**.41988**	.22885	.25441	**.40165**	.21699	.42722
18	**.42935**	**.46190**	.38090	.36111	.08013	-.04090
25	**.58661**	.24378	.21232	.30638	.17026	-.04109
28	**.60924**	.23995	.21221	.28857	.29122	.07009
29	**.54556**	.14211	.18517	.30484	.32343	-.14216
30	**.51786**	.17973	.19560	**.45745**	.25810	.02165
31	**.62285**	.22347	.25174	.19970	.05981	.26111
32	**.54456**	.28238	.29858	.39393	.12631	.00787
33	**.42798**	**.45619**	.36084	.38644	.09166	.08189
37	**.67172**	.19026	.16866	-.02397	.35724	.10472
38	**.63647**	.28532	.18597	.00532	.27747	.21545
39	**.43301**	.28468	.21957	**.44019**	.19977	.16849
40	**.58765**	.31521	.20920	.24039	.22311	.18770
41	**.40987**	**.46210**	**.46068**	.15660	.08456	.05746
42	**.48715**	.38650	.31636	.31639	.19242	.05683
43	**.62650**	.34288	.29436	.07667	.12717	.09568
44	**.64791**	.20815	.24949	.20086	.18991	.12225
47	**.54853**	.39131	.30323	.25261	.13656	.01776
48	**.60416**	.31451	.29590	.11939	-.02528	-.06022

49	**.67743**	.24823	.29306	.16931	.15387	.01544
52	**.41179**	.21232	**.60653**	.10520	.24175	.20680
54	**.45011**	.09570	**.53065**	.19868	.30759	.14051

Leadership items						
Item	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
3	.18246	**.45064**	**.44872**	.37276	.06347	.03727
13	.23699	**.56755**	**.40772**	.25305	.12803	-.01752
24	.22777	**.60385**	**.40174**	.14970	.06622	-.05588
41	**.40987**	**.46210**	**.46068**	.15660	.08456	.05746
45	.28410	**.50950**	**.41267**	.25188	.12641	-.08307
50	.39515	**.47939**	**.40961**	.17660	.11967	.21165
51	.38659	.23978	**.64380**	.14499	.15822	.07663
52	**.41179**	.21232	**.60653**	.10520	.24175	.20680
53	.15638	**.40875**	**.69679**	.11437	.12208	.04985
54	**.45011**	.09570	**.53065**	.19868	.30759	.14051
55	.26519	**.40264**	**.63416**	.13291	.13659	-.11027
56	.15446	**.43945**	**.66518**	.15589	.17657	.02899
57	.35558	.20302	**.65132**	.17410	.22754	.02134
58	.13868	.38092	**.73332**	.21798	.11418	.06223
59	.29417	.29383	**.74503**	.19187	.19275	.09319
60	.35488	.30586	**.70848**	.18115	.17805	.00612
61	.38504	.27387	**.69661**	.19632	.15925	.08005
62	.16296	**.44867**	**.68075**	.12665	.09048	.15621

Looking at Table 10, Items 3, 13, 17, 18, 24, 30, 33, 39, 41, 45, 50, 52, 53, 54, 55, 56 and 62 are cross-loading items, and can either be modified by rewording them or can alternatively be removed.

A comparison of factor loadings across the groups of raters, reveals inconsistency in the classification of items. Referring to the tables “Head” is shown to be classified as Factor

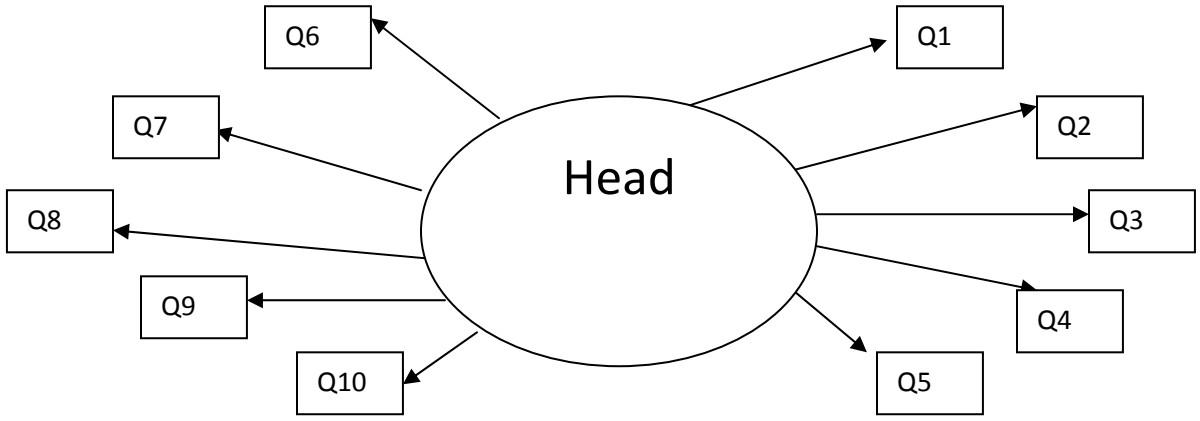
1 for self-ratings, Factor 3 for colleagues and managers, and Factor 2 for subordinates. “Heart 1” is shown to be classified as Factor 6 for colleagues, managers, and subordinates and Factor 4 for self-ratings. “Heart 2” is shown to be classified as Factor 5 for colleagues and subordinates, Factor 2 for managers and Factor 3 for self-ratings. “Heart 3” is shown to be classified as Factor 4 for colleagues and subordinates, Factor 5 for managers and Factor 6 for self-ratings. “Hands” is shown to be classified as Factor 1 for colleagues and subordinates, Factor 4 for managers and Factor 5 for self-ratings. “Leadership” is shown to be classified as Factor 2 for colleagues and self-ratings, Factor 1 for managers and Factor 3 for subordinates. On average, factor loadings for colleagues and subordinates are similar. This could imply that, depending on the rater group, different versions of the existing instrument should be constructed. Alternatively, depending on the rater group, data from different raters could be grouped and analysed differently. This is further supported by the variance in the cross-loaded and non-loading items between the four groups of raters as discussed in preceding paragraphs. Furthermore, for all groups the proposed groupings of items under the headings of “Head, Heart 1, Heart 2, Heart 3, Hands and Leadership” need to be revised and modified according to the factor loadings that are found for each group above. In summary Tables 7 to 10 show that items of the questionnaire load differently on the specified factors depending on the rater group. This means that future versions of the questionnaire might need to be designed for each rater group or alternatively items need to be modified to load on the same factors and in the same way for all rater groups.

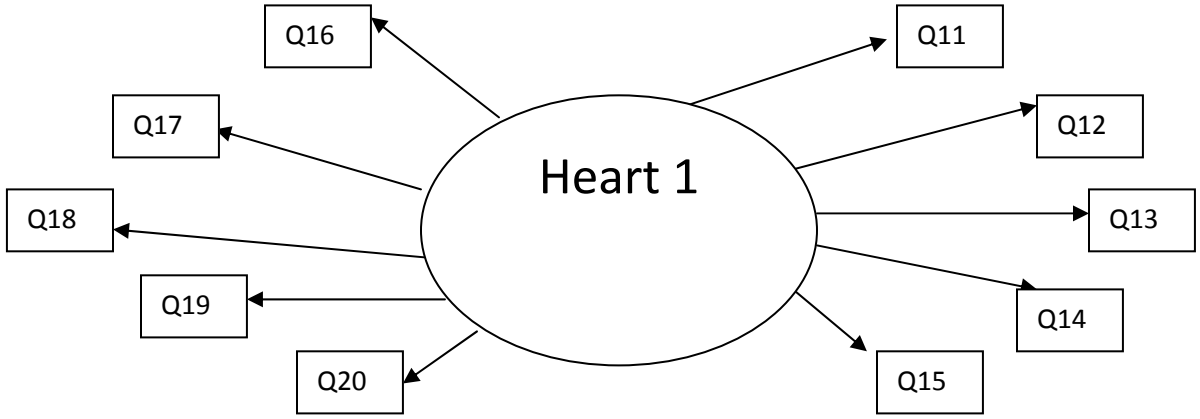
Confirmatory factor analysis of the 360-degree feedback scale (Leadership 3)

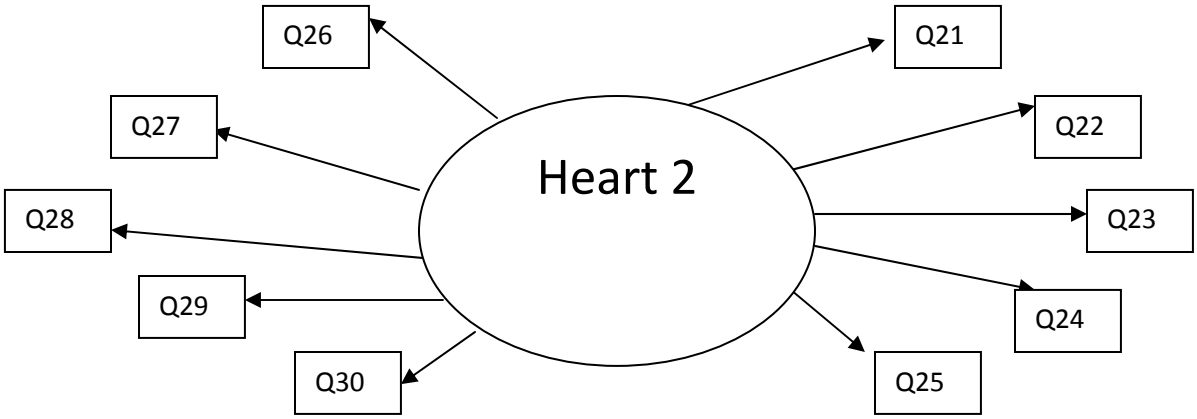
As outlined in the previous chapter, the confirmatory factor analysis tests certain relationships between variables, knowing in advance the number of that is factors required to explain the strength and nature of these relationships (Bartholomew et al., 2008). When this type of analysis is conducted, a hypothesised or existing model is compared to a covariance matrix in order to determine a population matrix (Schreiber et al., 2006). The confirmatory factor analysis has been calculated for the whole sample,

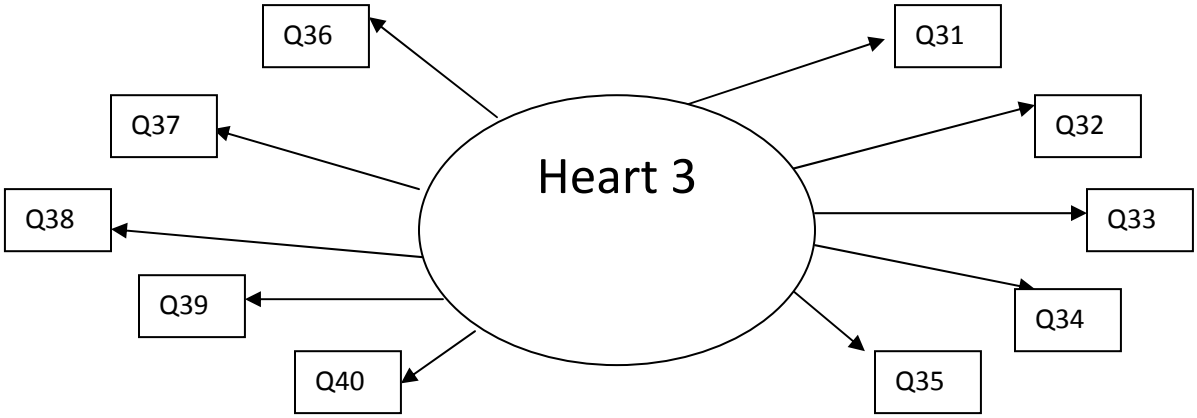
as well as for the four groups of raters where Model 3 refers to self-raters, Model 1 to colleagues, Model 2 to managers and Model 4 to subordinates.

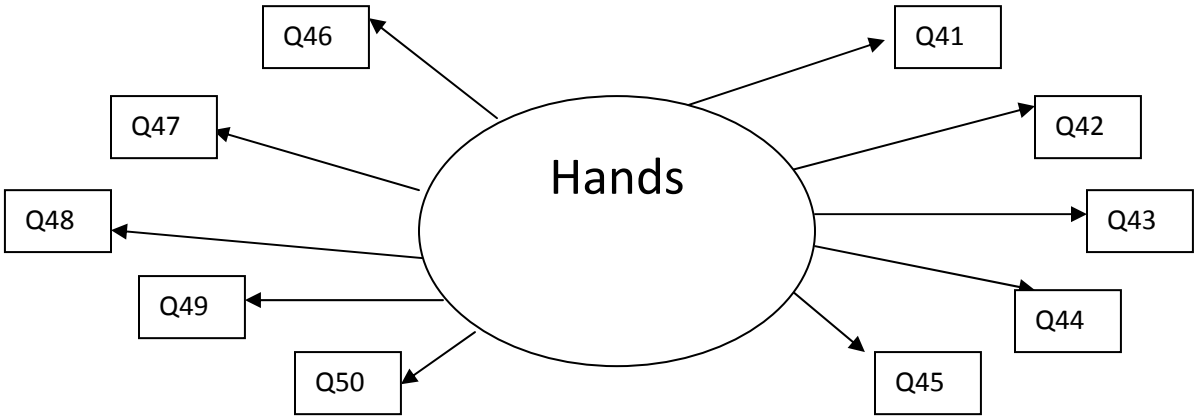
In order to begin a confirmatory analysis, a path diagram or structural model needs to be constructed in order to represent the way the different variables entering the model are related (Bartholomew et al., 2008). In this study, the path diagram/structural model below shows how the questionnaire items are grouped together, how they relate to the proposed questionnaire dimensions, and whether they relate to the gathered data making up the 360-degree feedback questionnaire (Leadership 3).






Heart 2
Q26
Q27
Q28
Q29
Q30
Q21
Q22
Q23
Q24
Q25


Heart 3
Q36
Q37
Q38
Q39
Q40
Q31
Q32
Q33
Q34
Q35


Hands
Q46
Q47
Q48
Q49
Q50
Q41
Q42
Q43
Q44
Q45

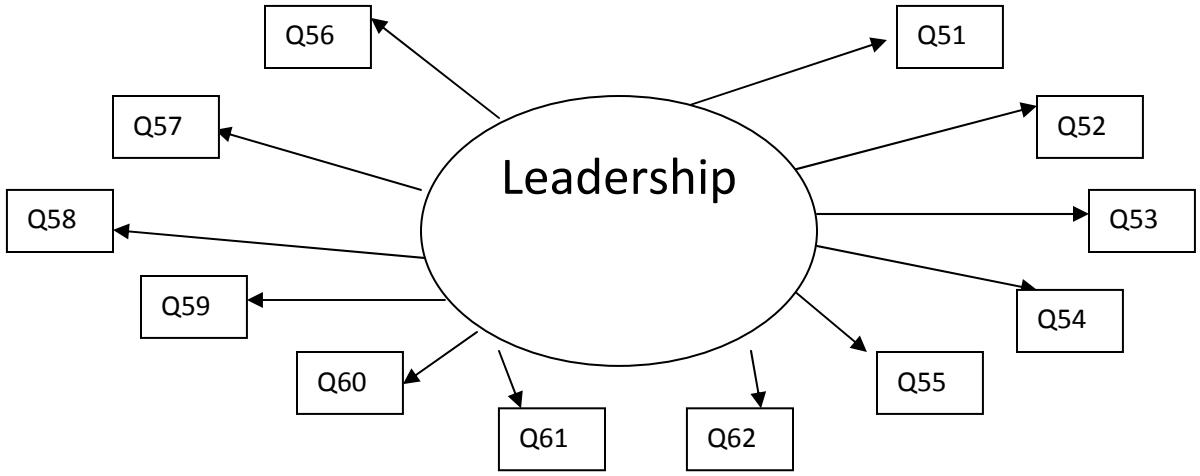


Figure 5. Structural model of the 360-degree feedback questionnaire (Leadership 3)

In Figure 5 the variables enclosed in ovals are called latent variables which are the "hidden" variables that cannot be measured, but form the theoretical dimensions of the questionnaire (Bartholomew et al., 2008). The variables enclosed in rectangles are called manifest variables which are the variables that can be directly measured (Bartholomew et al., 2008). In this case the latent variables refer to the theoretical dimensions of the questionnaire, and the items are the manifest variables grouped under each factor within the existing questionnaire. The use of single-headed arrows indicates direct relationships between the latent variables and their respective manifest variables (Bartholomew et al., 2008).

There are also pre-analyses needed to be calculated in order to determine whether criteria have been met to conduct the factor analysis (Schreiber et al., 2006). The sample size needs to be large enough so that there are at least 10 participants per

estimated parameter (Schreiber et al., 2006). In this study, 21 862 participants are a sufficiently large sample to calculate the analysis. Furthermore, analyses were also calculated for each group of raters with sample sizes of 2 788 self- raters, 7 710 colleague raters, 2 938 manager raters, and 8 426 subordinate raters. These subsample sizes are also adequate to meet specified criteria. Furthermore analyses denoted that convergence criteria had been met, and that the distribution of residuals was normal (Figures 6 and 7).

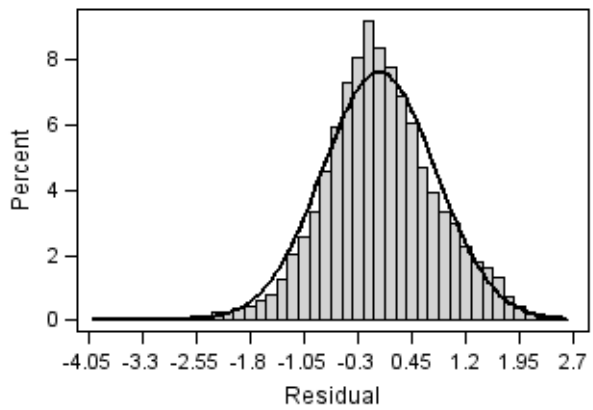


Figure 6. Distribution of residuals

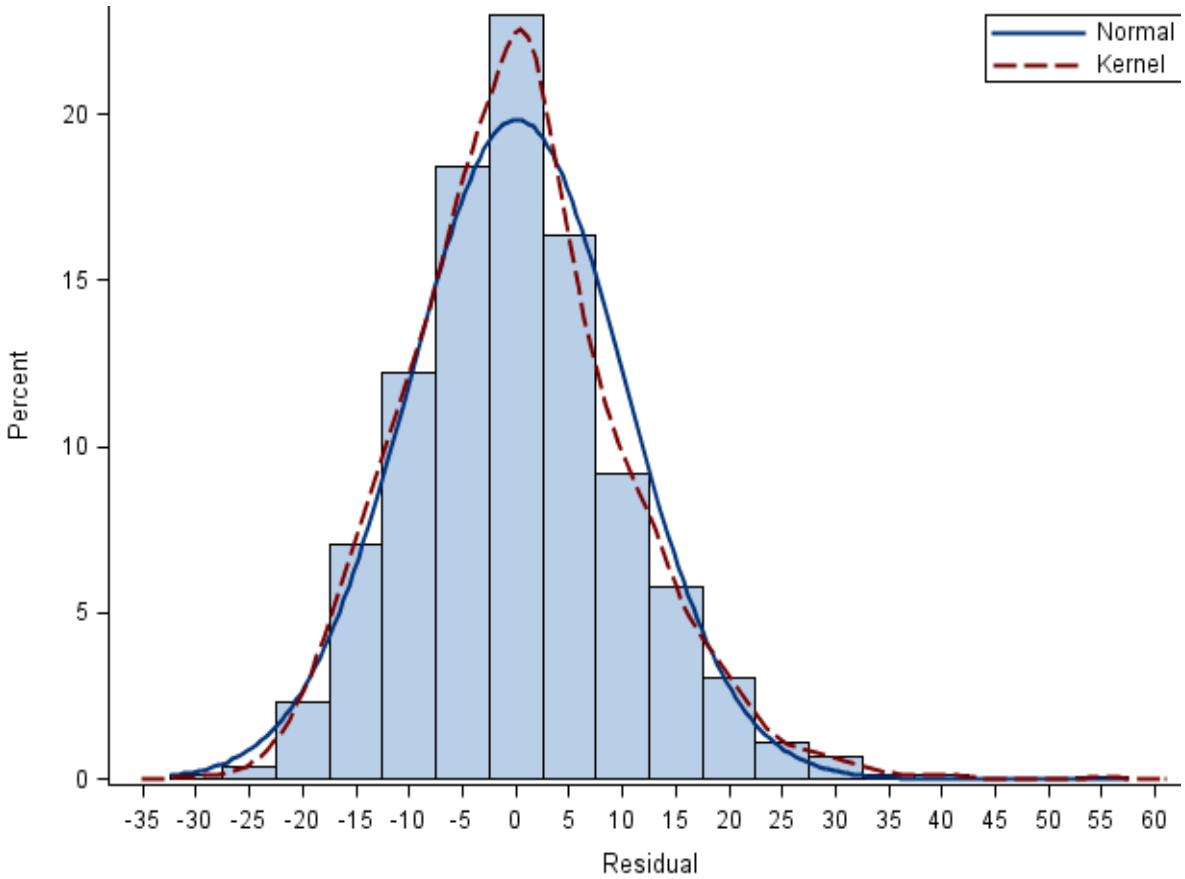


Figure 7. Distribution of standardised residuals

To test the adequacy of the specified model, the goodness of fit indices are calculated and analysed (Bartholomew et al., 2008). The Chi-square Goodness of fit test (χ^2) is very sensitive to large sample sizes, and therefore mostly significant which indicates a poor fit (Bartholomew et al., 2008). The goodness of fit index (GFI), and adjusted goodness of fit index (AGFI), which is an indication of how parsimonious the model is, should be greater than .9 to indicate a good fit (Schreiber et al., 2006). The Bentler-Bonnett normed fit index (NFI) compares the model with some baseline model, and should be greater than .9 to indicate a good fit (Bartholomew et al., 2008). The non-normed fit index (NNFI) and comparative fit index (CFI) are relative fit indices comparing the fitted model to the baseline model, and should be greater than .9 to indicate a good fit of the model (Bartholomew et al., 2008). The root mean square error of

approximation (REMSEA) should be less than .05, and thus be significant to indicate a good fit, and can test whether a model is particularly inadequate (Schreiber et al., 2006). Hoelter's critical sample size (CN) indicates the adequacy of the sample size, where a sample size greater than 200 indicates a good fit (Schreiber et al., 2006). The standardised root mean square residual (SRMSR) is a standardised summary of the average covariance residuals, and should be less than .05 to indicate a good fit (Bartholomew et al., 2008). The fit indices for the whole sample, as well as for each model are approximately equal and are thus presented as one value per index in the table 11 below.

Table 11. Goodness of fit indicators

Indicator	Index	Good/Poor Fit
χ^2	104,042.512	Poor fit
	$p<0.0001$	
Adjusted GFI	.645	Poor fit
GFI	.67	Poor fit
NFI	.817	Poor fit
NNFI	.828	Poor fit
CFI	.828	Poor fit
REMSEA	.064	Poor fit
CN	945	Good fit
SRMSR	.052	Poor fit

The goodness of fit indices indicate a poor fit of the hypothesised model to the data. One possible way of improving this fit, is to modify the existing model. One possibility is to respecify the model in the structural model to improve the fit.

Referring to Figure 6, it is evident that the latent variables are exogenous with paths extending to designated test items only, being the manifest variables, as the purpose of the analysis is to determine validity of the instrument in question. Therefore, results of the standardised parameter estimates, together with their significance level, are

presented in tables for each model below. The standardised parameter estimates are path coefficients from the specified path diagram/structural model, indicating whether the model is a good fit to the data obtained, and the strength of the proposed paths outlined in the path diagram (Schreiber et al., 2006). This gives an indication of which paths are adequate, and which can be rearranged or possibly removed. One should also note that for all models, all paths specified in Figure 6 were significant with a moderate strength of relationships, and should thus remain in the model. Evidence of this assertion is presented in Tables 12 to 15 below:

Table 12. Standardised parameter estimates/Standardised path coefficients – Model 3 (self)

Head		
Item	**Coefficient**	**t-value**
1	.48	-
2	.50	17.78
3	.52	18.24
4	.60	19.96
5	.51	18.01
6	.49	17.49
7	.49	17.59
8	.57	19.40
9	.50	17.92
10	.64	20.50
Heart 1		
Item	**Coefficient**	**t-value**
11	.53	-
12	.49	19.14
13	.57	21.28
14	.58	21.41
15	.53	20.32
16	.58	21.39

17	.49	19.05
18	.65	23.17
19	.55	20.79
20	.46	18.28
	Heart 2	
Item	**Coefficient**	**t-value**
21	.60	-
22	.54	22.03
23	.50	20.36
24	.53	21.50
25	.58	23.23
26	.60	23.89
27	.57	22.85
28	.61	24.19
29	.54	21.90
30	.55	22.29
	Heart 3	
Item	**Coefficient**	**t-value**
31	.57	-
32	.67	24.88
33	.65	24.35
34	.55	21.66
35	.47	19.15
36	.49	19.61
37	.58	22.54
38	.58	22.55
39	.58	21.51
40	.60	22.46
	Hands	
Item	**Coefficient**	**t-value**
41	.60	-

42	.67	25.72
43	.65	25.12
44	.59	23.54
45	.60	23.78
46	.54	21.848
47	.63	24.63
48	.49	20.33
49	.56	22.65
50	.66	25.63
Leadership		
Item	**Coefficient**	**t-value**
51	.65	-
52	.65	27.19
53	.65	27.16
54	.57	24.17
55	.59	24.86
56	.62	26.11
57	.61	25.63
58	.65	27.10
59	.80	32.13
60	.74	30.19
61	.76	30.90
62	.62	26.30

Looking at the results in Table 12 for the self-model; all paths are significant with most being moderate in strength (.40 to .80). In particular the paths between Items 59, 60 and 61 and Leadership show a strong relationship (.70 to .80). This indicates a good fit.

Table 13. Standardised parameter estimates/Standardised path coefficients – Model 1 (colleague)

Head		
Item	**Coefficient**	**t-value**
1	.63	-
2	.66	37.21
3	.65	36.87
4	.69	38.50
5	.68	38.10
6	.55	31.89
7	.66	37.19
8	.69	38.68
9	.63	35.90
10	.72	39.89
Heart 1		
Item	**Coefficient**	**t-value**
11	.66	-
12	.59	25.31
13	.69	40.53
14	.66	39.19
15	.60	35.68
16	.70	41.09
17	.66	39.28
18	.76	44.17
19	.70	41.34
20	.63	37.53
Heart 2		
Item	**Coefficient**	**t-value**
21	.73	-
22	.69	44.15

23	.66	42.55
24	.67	43.09
25	.68	43.42
26	.71	46.01
27	.67	42.94
28	.72	46.67
29	.61	39.21
30	.64	41.28
	Heart 3	
Item	**Coefficient**	**t-value**
31	.66	-
32	.73	42.44
33	.77	44.22
34	.67	39.02
35	.58	34.76
36	.64	37.87
37	.34	37.47
38	.68	39.88
39	.68	40.01
40	.71	41.28
	Hands	
Item	**Coefficient**	**t-value**
41	.73	-
42	.76	49.49
43	.71	45.58
44	.69	44.45
45	.72	46.64
46	.69	44.69
47	.77	49.90
48	.65	41.49
49	.73	46.86

50	.74	47.60
	Leadership	
Item	**Coefficient**	**t-value**
51	.77	-
52	.80	56.07
53	.76	52.48
54	.72	49.46
55	.73	50.11
56	.78	24.20
57	.79	55.02
58	.80	55.94
59	.89	63.88
60	.86	60.63
61	.86	61.70
62	.78	52.56

Looking at the results in Table 13 for the colleague model, all paths are significant with most being moderate to strong (.50 to .89). These results indicate that, compared with the model for self-ratings, this model has a better fit. Item 37 to Heart 3 in particular show a weak relationship, although they are significant. Changing or removing this path may improve fit for the model.

Table 14. Standardised parameter estimates/Standardised path coefficients – Model 2 (manager)

	Head	
Item	**Coefficient**	**t-value**
1	.63	-
2	.60	21.96
3	.58	21.46
4	.65	23.48

5	.66	24.01
6	.53	19.91
7	.57	21.03
8	.65	23.64
9	.54	20.34
10	.67	24.10
	Heart 1	
Item	**Coefficient**	**t-value**
11	.59	-
12	.56	20.39
13	.66	23.16
14	.65	22.85
15	.57	20.60
16	.67	23.41
17	.61	21.80
18	.71	24.43
19	.65	22.88
20	.59	21.13
	Heart 2	
Item	**Coefficient**	**t-value**
21	.67	-
22	.62	23.91
23	.62	24.07
24	.62	24.17
25	.63	24.48
26	.66	25.49
27	.62	24.07
28	.70	26.78
29	.58	22.64
30	.63	24.38
	Heart 3	

Item	Coefficient	t-value
31	.58	-
32	.68	23.18
33	.73	24.27
34	.61	21.54
35	.50	18.37
36	.61	21.56
37	.58	20.64
38	.59	20.97
39	.64	22.18
40	.64	22.27
Hands		
Item	**Coefficient**	**t-value**
41	.71	-
42	.73	29.86
43	.62	25.40
44	.66	26.89
45	.66	27.11
46	.65	24.42
47	.71	28.87
48	.58	23.37
49	.65	26.28
50	.69	28.12
Leadership		
Item	**Coefficient**	**t-value**
51	.73	-
52	.77	32.37
53	.76	32.01
54	.68	28.25
55	.66	27.54
56	.75	31.82

57	.75	31.46
58	.75	31.86
59	.85	36.27
60	.78	33.08
61	.83	35.15
62	.70	29.33

Looking at the results for the manager's model in Table 14, all paths are significant with most being moderate to strong (.40 to .85). The path between Item 61 and Leadership is particularly strong.

Table 15. Standardised parameter estimates/Standardised path coefficients – Model 4 (subordinate)

Head		
Item	**Coefficient**	**t-value**
1	.67	-
2	.69	45.76
3	.71	46.97
4	.71	47.08
5	.68	45.19
6	.58	39.13
7	.70	46.47
8	.73	48.13
9	.66	44.22
10	.77	50.45
Heart 1		
Item	**Coefficient**	**t-value**
11	.67	-
12	.62	42.25
13	.74	49.10

14	.70	46.81
15	.62	41.10
16	.69	46.55
17	.70	47.22
18	.79	52.55
19	.72	48.00
20	.66	44.45
Heart 2		
Item	**Coefficient**	**t-value**
21	.77	-
22	.74	56.91
23	.70	53.22
24	.69	52.09
25	.70	53.36
26	.76	58.67
27	.70	53.16
28	.75	57.79
29	.65	48.71
30	.70	53.39
Heart 3		
Item	**Coefficient**	**t-value**
31	.70	-
32	.77	53.01
33	.81	56.02
34	.67	46.31
35	.62	42.92
36	.68	47.21
37	.68	47.03
38	.71	49.48
39	.72	50.16
40	.76	52.97

Hands		
Item	**Coefficient**	**t-value**
41	.77	-
42	.79	62.04
43	.75	57.48
44	.73	55.82
45	.72	55.56
46	.74	57.09
47	.79	61.53
48	.68	51.72
49	.76	58.97
50	.78	61.01
Leadership		
Item	**Coefficient**	**t-value**
51	.79	-
52	.78	63.18
53	.79	64.54
54	.72	56.90
55	.78	64.10
56	.81	66.51
57	.79	64.13
58	.84	69.25
59	.89	76.03
60	.88	74.90
61	.87	73.70
62	.81	66.75

Looking at the results for the subordinate model in Table 15, all paths are significant with most being moderate to strong (.60 to .89). Corresponding to the model for colleagues, these results indicate that, compared with the model for self and manager ratings, this model has a better fit.

The tables that are presented, show that for each model (self, colleague, manager, subordinate), all specified paths from Diagram 6 are significantly greater than 2.00 with relationships being moderate to strong (approximately .40 to .90). Therefore, existing paths should not be removed for any of the above-specified models.

Although the confirmatory factor analysis procedure tests the adequacy of a proposed existing model, it is also necessary to measure and investigate additional paths that might improve the fit of the model (Bartholomew et al., 2008). To do this modification indices, such as those used by Legrange in this study, are calculated to determine the expected decrease in the chi-squared test statistic by the inclusion of each respective extra parameter (Bartholomew et al., 2008). Best practice indicates that only large values of modification indices would be considered (Bartholomew et al., 2008). This approach has been adopted in this study, with Legrange's modification indices presented in Table 16 below. It is important to note that the Chi-square statistic for the current model was 104 042.512.

Table 16. Modification indices for extra paths

Path to	Path from	LM stat
q6	q15	1024
q15	q6	934.82471
q62	q58	719.23516
q58	q62	718.52195
q17	q46	527.82072
q46	q17	526.62636
q30	q29	519.05207
q29	q30	519.05066
q9	q62	468.89682
q23	q22	433.41436
q17	q11	290.31230
q11	q17	288.70625
q35	q20	278.26305

q20	q35	268.21675
q54	Heart3	109.89756
q58	Heart3	104.28875
q54	Hands	101.11850
q46	Leadership	77.90713

According to Table 16, the following extra paths were specified with sufficient reductions in Chi-square statistics:

From Item 6 to Item 15; and from Item 15 to Item 6

Item 6 ("Is open and receptive to the new ideas of others and different ways of doing things") loads onto the Head dimension and Item 15 ("Contributes supportively and expertly to the success of peers and superiors") on the Heart 1 dimension of the existing model. Both items refer to the support of collaboration and thoughts of others as a leader. This reflects a possible conceptual similarity between these two items. However, on examination of the questions, each appears to fit more appropriately its original position; where Item 6 has more to do with the way in which decisions are made through collaborating with others, and Item 15 concentrates more on the support and development of others. Therefore, the recommendation is not to add this path in modified versions of the instrument, but possibly rather to reword both items to be more specific to the dimension each loads on.

From Item 58 to Item 62; and from Item 62 to Item 58

Both items 58 ("When I am with this person I feel the urgency of what we need to do") and 62 ("When I am with this person I feel proud to be part of a winning team") load onto the Leadership dimension of the existing model. Both items refer conceptually to aspects of motivation in the context of a team and the achievement of goals which are objectives for work-based teams. Therefore, the recommendation is the possible combination of these two items into one question which will encompass the inspiration and goal-orientation component of motivation.

From Item 17 to Item 46; and from Item 46 to Item 17

Item 17 ("Supports others' decisions") loads onto the Heart 1 dimension and Item 46 ("Has the courage to implement tough consequences for poor performance") onto the Hands dimension of the existing model. However, on examination, these two items do not appear to have conceptual similarity. Possible reasons for the observed modification indices here are that items have been poorly specified, or alternatively, that the raters experienced some bias in terms of method during completion of the questionnaire. Therefore, the recommendation is not to add this path in modified versions of the instrument.

From Item 29 to Item 30; and from Item 30 to Item 29

Both Items 29 ("Is consistently dependable in all situations") and 30 ("Keeps going with resilience in the face of difficulties or opposition") load onto the Heart 2 dimension of the existing model. Conceptually these items are similar as they both refer to effective self-management through consistent action towards goal achievement in any context. Therefore, the recommendation here is to combine the two questions into one encompassing question that deals with self-management in relation to goal achievement resilience.

From Item 62 to Item 9

Item 62 (When I am with this person I feel proud to be part of a winning team") loads onto the Leadership dimension and Item 9 ("Takes tough decisions when necessary, even when they are unpopular") onto the Head dimension of the existing model. However, on examination these two items do not appear to have conceptual similarity. Possible reasons for the observed modification indices here are that items have been poorly specified, or alternatively, that the raters experienced some bias in terms of method during completion of the questionnaire. Therefore, the recommendation is not to add this path in modified versions of the instrument.

From Item 22 to Item 23

Both Items 22 ("Is aware of his/her strengths and weaknesses and is not afraid to disclose them") and 23 ("Obtains and attends to feedback from other people without defensiveness") load onto the Heart 2 dimension of the existing model. Conceptually both items attend to aspects of self-awareness, which is the understanding of one's strengths and weaknesses, as well as accepting feedback about them. Therefore, the recommendation here is to combine the two questions into one encompassing question that deals with self-awareness relating to identification of strengths and weaknesses, and the willingness to accept related feedback.

From Item 11 to Item 17; and from Item 17 to Item 11

Both Items 11 ("Listens attentively and accurately") and 17 ("Supports others' decisions") load onto the Heart 1 dimension of the existing model. Although Item 17 deals with the support of others' decisions, leaders need to be able to listen effectively to do so. Item 11 addresses this. Therefore, the recommendation here is to combine the two questions into one encompassing question that deals with listening effectively, and (listening to?) the support of others' decisions. Alternatively, both items can be reworded to be more specific and distinct from each other conceptually. Item 11 can be reworded to address communication of others while Item 17 can be reworded to address collaboration during decision-making.

From Item 20 to Item 35; and from Item 35 to Item 20

Item 20 ("Inspires support and enthusiasm in others for a shared vision and mission") loads onto the Heart 1 dimension and Item 35 ("Is effective at winning people over to his/her point of view") onto the Heart 3 dimension of the existing model. Both items, conceptually relate to the inspiration of others, and gaining support from them. Therefore, the recommendation here is to combine the two questions into one encompassing question that deals with gaining the support of others, and inspiring those others as a leader. Furthermore, this question can be categorised under the Heart 3 dimension, as individuals constitute teams that make the item applicable to the individual and at team level.

From Heart 3 dimension to Item 54

Item 54 (“When I am with this person I feel that I want to do things right all the time”) loads onto the Leadership dimension of the current model. However, the item deals with the ability of a leader to motivate team members. This corresponds with the Heart 3 dimension which has the purpose to enable groups and the organisation. Considering that Item 40 deals with loyalty to the team, and, by implication, the need to perform well, the recommendation is either to assimilate Item 54 into the Heart 3 dimension, and remove it from the questionnaire, or reword it to fit distinctly into the Leadership dimension.

From Heart 3 to Item 58

Item 58 (“When I am with this person I feel the urgency of what we need to do”) loads onto the Leadership dimension of the current model. However, the item addresses the motivation provided by a leader to complete goals and tasks which correspond to the Heart 3 dimension, which has the purpose of enabling groups and the organisation. Considering that Item 40 deals with loyalty to the team, and, by implication the need to perform well, the recommendation is either to assimilate Item 58 into the Heart 3 dimension, and remove it from the questionnaire, or reword it to fit distinctly into the Leadership dimension.

From Hands to Item 54

Item 54 (“When I am with this person I feel that I want to do things right all the time”) loads onto the Leadership dimension of the current model. However, the item addresses the need for the leader to inspire others to perform tasks in the most effective way. This corresponds to the Hands dimension, which has the purpose of executing tasks and objectives. Considering that Item 41 deals with high standards of performance, the recommendation is either to assimilate Item 54 into the Hands dimension, and remove it from the questionnaire, or reword it to fit distinctly into the Leadership dimension.

From Leadership to Item 46

Item 46 ("Has the courage tom implement tough consequences for poor performance") loads onto the Hands dimension of the current model. However, the item addresses the ability of leaders to make tough decisions and implement them. This corresponds with aspects of the Leadership dimension which refer to the impact leaders make on others. However, there appears to be no similarity between Item 46 and any existing items of the Leadership dimension. Furthermore, Item 46 is more suitable when positioned in the Hand dimension, and therefore the recommendation is not to add this path into future modifications of the instrument.

4.2 Distribution analyses

To enable an interval variable to be distributed normally with the expectation of yielding parametric data, the skewness coefficient should be between -1 and 1, the kurtosis coefficient should be between -1 and 2, and the Kolmogorov-Smirnov test should be non-significant at the p<.05 level of significance (Huck, 2004). Meeting the requirements of normality and interval data, parametric tests can thus be conducted (Huck, 2004).The summarised results of the distribution analyses are reported in Table 17 below.

Table 17 Distribution analyses of ratings for all rater groups

Model	Mean	Median	Mode	Standard Deviation	Skewness	Kurtosis	Kolmogorov-Smirnov test
Self	4.72	4.66	4.18	.67	.38	.72	Significant (*p*<.01)
Colleague	4.71	4.66	4.65	.86	.09	.34	Significant (*p*<.01)
Manager	4.71	4.69	4.32	.77	-.05	.53	Significant (*p*<.01)
Subordinate	4.93	4.93	5.00	1.01	-.33	.42	Significant (*p*<.01)

In summarising Table 17, self-ratings were interpreted as being normal owing to the skewness and kurtosis coefficients that fell in the normal range, even though the

Kolmogorov-Smirnov test was significant. The same result was found for ratings by colleagues, managers and subordinates. Therefore, parametric statistical analyses, such as Pearson's correlation coefficient, were carried out.

4.3 Research questions 1 and 2

The results of relationships concerning the areas of consistency, or the lack thereof, are presented for each group of raters. Table 18 shows the correlation matrix for participants (self-raters), Table 19 for colleagues, Table 20 for managers and Table 21 for subordinates.

Table 18. Correlations for participants (self-ratings) (*N*= 2788)

	Head	Heart1	Heart2	Heart3	Hands	Leadership	Total360
Head	-------	.79 *p*=<.0001*	.73 *p*=<.0001*	.72 *p*=<.0001*	.72 *p*=<.0001*	.71 *p*=<.0001*	.87 *p*=<.0001*
Heart1		-------	.80 *p*=<.0001*	.78 *p*=<.0001*	.76 *p*=<.0001*	.70 *p*=<.0001*	.90 *p*=<.0001*
Heart2			-------	.78 *p*=<.0001*	.76 *p*=<.0001*	.66 *p*=<.0001*	.88 *p*=<.0001*
Heart3				-------	.80 *p*=<.0001*	.69 *p*=<.0001*	.89 *p*=<.0001*
Hands					-------	.74 *p*=<.0001*	.90 *p*=<.0001*
Leadership						-------	.86 *p*=<.0001*
Total360							-------

*= Significance at p<0.05

Table 19. Correlations for colleagues (*N*= 7710)

	Head	Heart1	Heart2	Heart3	Hands	Leadership	Total360
Head	-------	.85 *p*=<.0001*	.82 *p*=<.0001*	.80 *p*=<.0001*	.81 *p*=<.0001*	.78 *p*=<.0001*	.91 *p*=<.0001*
Heart1		-------	.85 *p*=<.0001*	.85 *p*=<.0001*	.85 *p*=<.0001*	.77 *p*=<.0001*	.94 *p*=<.0001*
Heart2			-------	.85 *p*=<.0001*	.84 *p*=<.0001*	.76 *p*=<.0001*	.93 *p*=<.0001*
Heart3				-------	.86 *p*=<.0001*	.76 *p*=<.0001*	.93 *p*=<.0001*
Hands					-------	.79 *p*=<.0001*	.94 *p*=<.0001*
Leadership						-------	.90 *p*=<.0001*
Total360							-------

*= Significance at p<0.05

Table 20. Correlations for managers (*N*= 2938)

	Head	Heart1	Heart2	Heart3	Hands	Leadership	Total360
Head	-------	.85 *p*=<.0001*	.79 *p*=<.0001*	.79 *p*=<.0001*	.80 *p*=<.0001*	.75 *p*=<.0001*	.91 *p*=<.0001*
Heart1		-------	.84 *p*=<.0001*	.83 *p*=<.0001*	.84 *p*=<.0001*	.75 *p*=<.0001*	.93 *p*=<.0001*
Heart2			-------	.83 *p*=<.0001*	.81 *p*=<.0001*	.74 *p*=<.0001*	.91 *p*=<.0001*
Heart3				-------	.85 *p*=<.0001*	.76 *p*=<.0001*	.92 *p*=<.0001*
Hands					-------	.77 *p*=<.0001*	.93 *p*=<.0001*

Leadership						-------	.89 *p*=<.0001*
Total360							-------

*= Significance at p<0.05

Table 21. Correlations for subordinates (*N*= 8426)

	Head	**Heart1**	**Heart2**	**Heart3**	**Hands**	**Leadership**	**Total360**
Head	-------	.88 *p*=<.0001*	.86 *p*=<.0001*	.83 *p*=<.0001*	.84 *p*=<.0001*	.81 *p*=<.0001*	.93 *p*=<.0001*
Heart1		-------	.89 *p*=<.0001*	.87 *p*=<.0001*	.87 *p*=<.0001*	.80 *p*=<.0001*	.95 *p*=<.0001*
Heart2			-------	.87 *p*=<.0001*	.87 *p*=<.0001*	.80 *p*=<.0001*	.94 *p*=<.0001*
Heart3				-------	.88 *p*=<.0001*	.76 *p*=<.0001*	.93 *p*=<.0001*
Hands					-------	.83 *p*=<.0001*	.94 *p*=<.0001*
Leadership						-------	.91 *p*=<.0001*
Total360							-------

*= Significance at p<0.05

Tables 18 to 21 show significant relationships between all dimensions of he questionnaire across all rater groups. The strength of all these relationships ranged from moderate to strong.

Furthermore, comparisons were made by using Fisher's z transformations in order to determine whether there were significant differences between groups for the dimensions of the 360-degree feedback instrument.

Table 22. Comparisons using Fisher's z transformations (*N=21862*)

	Self and manager	Self and subordinate	Self and colleague	Manager and subordinate	Manager and colleague	Subordinate and colleague
Head to Heart1	-5.73 *p*=<.05*	-14.15 *p*=<.05*	-7.24 *p*=<.05*	-6.80 *p*=<.05*	**-0.27** **Ns**	8.96 *p*=<.05*
Head to Heart2	-5.60 *p*=<.05*	-16.24 *p*=<.05*	-9.48 *p*=<.05*	-8.99 *p*=<.05*	-2.60 *p*=<.05*	8.67 *p*=<.05*
Head to Heart 3	-6.21 *p*=<.05*	-12.28 *p*=<.05*	-7.93 *p*=<.05*	-4.44 *p*=<.05*	**-0.40** **Ns**	5.50 *p*=<.05*
Head to Hands	-7.10 *p*=<.05*	-14.69 *p*=<.05*	-9.82 *p*=<.05*	-5.80 *p*=<.05*	**-1.29** **Ns**	6.13 *p*=<.05*
Head to Leadership	-3.02 *p*=<.05*	-10.60 *p*=<.05*	-6.18 *p*=<.05*	-6.68 *p*=<.05*	-2.45 *p*=<.05*	5.75 *p*=<.05*
Head to Tot360	-5.72 *p*=<.05*	-14.78 *p*=<.05*	-9.11 *p*=<.05*	-7.13 *p*=<.05*	-2.04 *p*=<.05*	6.78 *p*=<.05*
Heart1 to Heart2	-4.68 *p*=<.05*	-14.56 *p*=<.05*	-8.32 *p*=<.05*	-8.41 *p*=<.05*	-2.55 *p*=<.05*	7.89 *p*=<.05*
Heart1 to Heart3	-5.91 *p*=<.05*	-12.51 *p*=<.05*	-8.80 *p*=<.05*	-4.99 *p*=<.05*	**-1.58** **Ns**	4.56 *p*=<.05*
Heart1 to Hands	-8.48 *p*=<.05*	-15.09 *p*=<.05*	-10.97 *p*=<.05*	-4.37 *p*=<.05*	**-0.66** **Ns**	5.03 *p*=<.05*
Heart1 to Leadership	-3.79 *p*=<.05*	-10.52 *p*=<.05*	-6.92 *p*=<.05*	-5.70 *p*=<.05*	-2.27 *p*=<.05*	4.62 *p*=<.05*

Heart1 to Tot360	-6.53 *p*=<.05*	-13.42 *p*=<.05*	-9.51 *p*=<.05*	-4.91 *p*=<.05*	**-1.49 Ns**	4.54 *p*=<.05*
Hands to Leadership	-2.71 *p*=<.05*	-10.27 *p*=<.05*	-5.46 *p*=<.05*	-6.76 *p*=<.05*	-2.14 *p*=<.05*	6.24 *p*=<.05*
Heart2 to Heart3	**-1.58 Ns**	-3.63 *p*=<.05*	-2.56 *p*=<.05*	**-1.61 ns**	**-0.63 Ns**	**1.31 ns**
Heart2 to Hands	**-1.58 Ns**	-4.69 *p*=<.05*	-3.22 *p*=<.05*	**-0.53 ns**	**0.79 Ns**	**1.81 ns**
Heart2 to Leadership	**-1.68 Ns**	-4.31 *p*=<.05*	-2.68 *p*=<.05*	-2.19 *p*=<.05*	**-0.64 Ns**	2.10 *p*=<.05*
Heart2 to Tot360	**-1.68 Ns**	-5.00 *p*=<.05*	-3.25 *p*=<.05*	-2.71 *p*=<.05*	**-1.22 Ns**	**1.80 ns**
Heart3 to Hands	**-1.76 Ns**	-4.02 *p*=<.05*	-2.64 *p*=<.05*	**-1.75 ns**	**-0.49 Ns**	**1.70 ns**
Heart3 to Leadership	**-1.73 Ns**	-2.62 *p*=<.05*	-2.01 *p*=<.05*	**-0.47 ns**	**0.07 Ns**	**0.75 ns**
Heart3 to Tot360	-2.02 *p*=<.05*	-3.11 *p*=<.05*	**-0.58 ns**	**-0.53 ns**	**1.78 Ns**	3.19 *p*=<.05*
Leadership to Tot360	**-1.35 Ns**	-3.19 *p*=<.05*	-2.33 *p*=<.05*	**-1.41 Ns**	**-0.65 Ns**	**1.00 ns**
Hands to Tot360	**-1.82 Ns**	-4.33 *p*=<.05*	-3.17 *p*=<.05*	**-1.92 Ns**	**-0.91 Ns**	**1.33 ns**

*= Significance at p<0.05; **ns**=not significant at p>0.05

In summary of the findings from Tables 18 to 22 above, the following trends were observed. On average, there was a significant difference (*p*=<.05) between self-ratings and manager ratings, with correlation coefficients of weaker strength for self raters than managers. A similar result was found between self and colleague ratings across most dimensions. There is also a significant difference between self-ratings and subordinate ratings, with correlations for participants being weaker than subordinate ratings across all measured dimensions. Looking at the relationship between manager and

subordinate rater groups, significant differences were found across most dimensions with managers' correlations being on average weaker than subordinates'. However, on some dimensions no significant differences were found between managers and subordinates. This finding was also found between subordinates and colleagues. On average, managers and colleagues were found to show no significant difference between ratings on dimensions of the 360-degree feedback instrument.

Hierarchical linear model (HLM) analysis results

The HLM analysis was calculated to conduct cross-level analyses across the different rater groups of the 360-degree feedback process. This technique has simultaneously assessed relationships within a particular level, as well as between the different levels (Wech & Heck, 2004). The aim is to identify whether results, differences, or effects on the dependent variable occur as a consequence of group or level membership or whether it is due to individual variance (Wech & Heck, 2004). The descriptions of the components of the analysis are specified in Table 23 below.

Table 23. Components of the HLM analysis

Component	Description
Dependent variable	Tot360
Independent variable	Type of rater group
Level 1	Individual respondents
Level 2	Participant to whom ratings apply
Specified model	$Tot360_{ij} = \beta_{oi} + \beta_1 \text{ Type} + \varepsilon_{ij}$
	Fixed-effects model
Variance between groups (γ^2)	.116

Variance within groups (σ^2)	.672
Total variance ($\sigma^2 + \gamma^2$)	.788

Calculation of the intra-class correlation coefficient (ICC) represents the measure of variance that could potentially be explained by the Level 2 predictor, which indicates how much variance is explained by group or level membership compared with individual variance (Wech & Heck, 2004). The ICC is calculated by dividing the variance between groups by the total variance (Wech & Heck, 2004).

The ICC was found to be .147, which represents 14,7 percent of the variance being explained by respondent type and 85,3 percent being explained by individual respondents.

The post hoc analyses yielded insight into some of the differences between groups, and are presented in Tables 24 and 25 below.

Table 24. HLM post hoc analyses

Type	Type	Estimate	Standard Error	DF	t Value	*p* value
Colleague	Manager	-.00477	.02168	14E3	-.22	.8258
Colleague	Self	-.01042	.02020	14E3	-.52	.6060
Colleague	Subordinate	-.2144	.01598	14E3	-13.41	<.0001*
Manager	Self	-.00565	.02451	14E3	-.23	.8177
Manager	Subordinate	-.2096	.02109	14E3	-9.94	<.0001*
Self	Subordinate	-.2039	.01954	14E3	-10.44	<.0001*

*= Significance at p<0.05

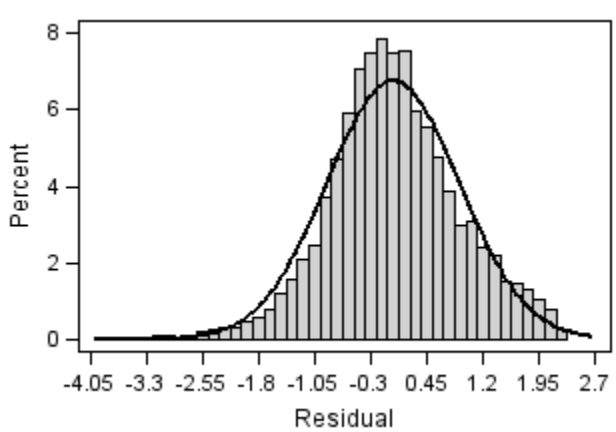
Table 25. HLM post hoc analysis type means

Type	Estimate
Colleague	4.71
Manager	4.72
Self	4.72
Subordinate	4.92

Looking at Tables 24 and 25, there is a significant difference in ratings between colleagues and subordinates (t_{14E3}= -.22; *p*<.001) with colleagues providing lower ratings as indicated by lower means when compared with subordinates (colleague μ = 4.71; subordinate μ=4.92). This result was also found between managers and subordinates (t_{14E3}= -9.94; *p*<.001) with managers providing lower ratings (manager μ = 4.72; subordinate μ=4.92). A significant difference was also found between self-ratings and participant ratings (t_{14E3}= -10.44; *p*<.001) with self-ratings being lower than subordinates' (self μ = 4.72; subordinate μ=4.92).

Furthermore all assumptions were met in order to conduct the HLM analysis, which was the normal distribution of residuals with no outliers clearly visible (Refer to Figure 8).

Figure 8. Distribution of residuals (b)



4.4 Summary of key findings

In answering the first research question, the degree of consistency varied between the different groups of raters. Self-ratings were significantly different from ratings by colleagues, managers and subordinates. However, self-ratings were found to be lower than ratings by all other rater groups. On average subordinate ratings were higher than all other rater groups and managers and colleagues were shown on average to show no significant differences.

The HLM analysis revealed that that differences observed were due to individual differences, and not to the type of rater group the respondent belonged to.

Referring to the second research question, strong correlations were observed between all subscales across all rater groups. Furthermore, significant differences were found between rater groups across all subscales. However, results showed no clear pattern or grouping to isolate particular areas of congruency, or the lack thereof, within the 360-degree feedback questionnaire and its subscales.

5. Discussion

5.1 Key findings

This study is aimed at investigating the degree of consistency of ratings between the different rater groups of a particular 360-degree feedback questionnaire. These groups include the following:

- Participants referred to as “self” in analyses
- Subordinates and peers referred to as colleagues in analyses
- Supervisors referred to as managers in analyses.

The study is also aimed at identifying the areas of consistency and inconsistency in this context. Taking this aim into account, this chapter will discuss the key findings from the analyses conducted in relation to the current study’s research questions. This will be followed by limitations of the study as well as recommendations for future research.

In answering the first research question, the degree of consistency varied between the different groups of raters. In line with the expectation of finding differences between raters, self-ratings were significantly different from ratings by colleagues, managers and subordinates. Surprisingly, however, self-ratings were found to be lower than ratings by all other rater groups. When gathering the data that was used in this study, participants were allowed to select the respondents themselves. This self-selection of raters, presented a problem of bias that influenced the ratings (Carson, 2006). A review of past research showed that raters might intentionally inflate ratings for political reasons, reluctance to rate peers harshly, or due to the pervading culture and context (Fleenor et al., 2010). Furthermore, favourable rater-ratee relationships have been shown to display higher ratings compared to self-rating (Antonioni & Park, 2001, Fleenor et al., 2010). This finding corresponds with the results of this study.

Results of the current study showed subordinates’ ratings to be higher than self-ratings on average. An explanation for this rating inflation may be ascribed to fear of consequences of providing negative or honest feedback that may be less favourable than the self-ratings of participants (Toegel & Conger, 2003). In a study that looked at

45 managers who participated in a management development programme in an American public-sector financial institution, self-ratings were found to be lower than ratings by colleagues. This may be ascribed to participants having had low self-perception (Fletcher & Baldry, 2000). Over-estimation or under-estimation has been shown in past research to be linked to blind spots or low levels of self-awareness (Calhoun et al., 2010). This corresponds with findings in this study, and provides a possible explanation for them. The current research finding was surprising in that the majority of research within the 360-degree feedback context showed self-ratings to be more lenient and higher than ratings by others (Atkins & Wood, 2002, Heijden & Nijhof, 2004, Patiar & Mia, 2008, Yammarino & Atwater, 1993). This finding was supported by results of a study of department managers of 4 and 5 star luxury hotels in Australia, where self-ratings were affected by leniency being higher than other raters on average (Patiar & Mia, 2008).

Looking at other groups of raters, manager and colleague groups showed no significant differences in the current study. This was supported by findings on a study that was carried out on a sample of 2 213 employees of a Midwestern American insurance company, where manager and peer ratings were shown to be the most similar to each other with strong correlations between them (Beehr et al., 2001). However, results indicated that manager ratings were lower than subordinates' which conformed to expectations. Subordinates were also shown to have higher ratings than colleagues. These results were supported by a study carried out on 197 shift bosses from a big Romanian company, where colleague and subordinate ratings were higher than self-ratings and manager ratings (Liviu et al., 2009). However, contrary to the current research findings, supervisors were shown to have lower ratings compared with self-ratings (Liviu et al., 2009). This was attributed to a more conservative approach in evaluation that was adopted by managers and supervisors (Patiar & Mia, 2008).

Within the current study, ratings for all groups showed that, in terms of leadership performance, self-ratings provided the most nuanced picture of the self, followed by ratings of managers, colleagues and subordinates. This finding corresponds with the review of past research where accuracy decreases from managers to peers to

subordinates (Carson, 2006). Assuming that the nuanced picture provided by self-ratings is the most accurate, this is correct. However, this may not always be the case. Further evidence of this assertion was found in the study conducted by Beehr et al. (2001) where self-ratings provided a unique perspective, which might have been better, but was not always valid. A study conducted on a sample of 252 middle managers within the public utility industry at the University of Michigan Business School in America, showed that self-ratings were almost similar to supervisor or manager ratings (Hooijberg & Choi, 2000). This result corresponds with findings of this research where self-rating and manager ratings were not the same, but very similar in providing a nuanced portrait of individual participants. Furthermore, past research provides an explanation for this result by pointing out that in general individuals do not act similarly among rater groups that provide differing perspectives (Fleenor et al., 2010). In line with this, in the context of this study, managers and colleagues may have had similar frames of references that may account for the similarity in ratings of participants.

The results of the current study also revealed that differences observed were due to individual differences, and not to the type of rater group the respondent belonged to. Support of this finding was found in a study conducted on a resident programme evaluation of the American Board of Internal Medicine, which found that there was very little difference between evaluator categories, but that the difference was more pronounced between individual residents (Weigelt et al., 2004). This also corresponded with findings of a study that was carried out among 2 350 managers who participated in a management skills profile, and who showed that differences were associated more strongly with individual differences than the level of raters (Sytsma & Hezlett, 1998). Possible reasons for this may be found in large standard deviations, rater effects or the inadequate training of raters (Weigelt et al., 2004).

Referring to the second research question, strong correlations were observed between all subscales across all rater groups. Furthermore, significant differences were found between rater groups across all subscales. However, results showed no clear pattern or grouping to isolate particular areas of congruency, or the lack thereof, within the 360-degree feedback questionnaire and its subscales. Therefore, it is more accurate to

assume that congruency, or the lack thereof, can be interpreted in terms of whole constructs, and not their components, such as leadership performance. This could be ascribed to the problem of a lack of a common definition and consensus on the components of effective or transformational leadership (Kerin, 2010). Alternatively, this lack of discrimination between dimensions can be explained by results from a study on 2 213 employees of a Midwestern American insurance company (Beehr et al., 2001). Here, the halo effects and biases may have masked differences across dimensions of leadership performance (Beehr et al., 2001). Therefore, ratings of this particular 360-degree feedback questionnaire can only yield insight on leadership as a whole, failing to determine congruency or patterns of difference for the components of leadership.

In summary; it should be noted that the evidence presented in this study provides further support for the assertion that investigating congruency between groups of raters in a 360-degree feedback context is complex. Furthermore, owing to the fact that ratings were sourced from many organisations and from different types of participants, with varying numbers and types of respondents, important differences and similarities may have been masked in the current study. This may be the result of a lack of context specificity or other descriptive information to aid the interpretation of results. This is because the assumption of transportability of 360-degree feedback questionnaires to other groups and organisations without testing the validity or examining the context may confound results (Craig & Clark, 2009). Results can be confounded by either concealing critical differences or highlighting differences which may not be valid (Craig & Clark, 2009). This is a particularly salient factor in the current study where one standardised instrument has been utilised by individuals from different organisations and contexts. Furthermore, considering that the data obtained for this study was sourced across different organisations and contexts, the positive and negative biases from each context could balance out, and conceal any differences that may have been present in the original data.

5.2 Limitations of the study

As with all research, there are limitations to this study, and in order not to over-generalise the findings presented above, the conclusions that are drawn should be considered in the light of the limitations of the study.

Firstly, the archival data that was sourced and used for analysis in this study contained no demographic information or any descriptive contextual information. The researcher was not able to make conclusions regarding demographical information, nor was she able to link data to specific organisations, locations or contexts. This has limited the generalisability of the present study. The reason being that context can affect results in various ways (Fleenor et al., 2010). Demographic data has been shown in past research to influence ratings and equivalence within the 360-degree feedback context (Gillespie, 2005). The pervading organisational culture may also be highly political, preventing honest feedback or have a different conceptualisation of what good performance is compared with other organisations (Fleenor et al., 2010). Furthermore, not being able to control these possible extraneous variables, sets up the possibility of potential rival hypotheses which may have impacted on the findings of the current study. The above factors have resulted in the researcher approaching the findings of this study with caution.

Related to this particular limitation, the questionnaire that was used was a standardised generic tool used by all participants regardless of the organisation or context. The use of generic instruments may skew results, and present the problem of generalisability to all organisations (Carson, 2006). This is because no two organisations are the same, each presenting unique factors which may impact on results from 360-degree feedback questionnaires (Craig & Clark, 2009). These may be particularly salient factors as participants were drawn from various organisations, of which no information was available to establish context of results for better interpretation.

Secondly, the data was sourced from participants and respondents who completed the questionnaire over the Internet. This presented the potential problem of method error, as the researcher was not able to control conditions or provide appropriate instructions

during the completion of the questionnaire. Method effects have been shown to be the largest source of error in performance ratings (Sytsma & Hezlett, 1998). Therefore any misinterpretation or compromising conditions could not be noted or controlled during analysis. An example of this was the complete lack of control in the selection of raters. This presented the possibility that participants might have been able to choose only friends or individuals with whom they were sharing good relationships. These conditions presented possible biases in ratings from participants (Hooijberg & Choi, 2000). These, and other related factors, could have resulted in unintended biases or extraneous variables within the current study.

Another limitation of the current study is that although the researcher conducted analyses to investigate construct validity of the 360-degree feedback questionnaire (Leadership 3), modifications were not made to the original questionnaire that was followed by further factor analyses. Modification of the questionnaire was beyond the scope of this study. The researcher aimed to highlight any possible improvements of the questionnaire in order to aid interpretation of results as well as provide suggestions to the test developer for future administrations. This has implications for the interpretation of results of the current study. Furthermore, even though suggestions were made to improve the questionnaire, the confirmatory factor analysis revealed that the existing structure was adequate which further supported the researcher's approach to conduct analyses on data from the original questionnaire.

Lastly, by averaging the ratings for analysis purposes, important variations may have been concealed (Conger & Toegel, 2002). This provides a possible reason for not being able to identify any specific areas or dimensions of congruency or incongruency within the dimensions of leadership from the particular 360-degree feedback questionnaire that was used in this study.

5.3 Recommendations for future research

Although the current study yielded insight into the consistency between different groups of raters of a particular 360-degree feedback questionnaire the researcher was not able

to identify specific areas of consistency and inconsistency. Therefore the researcher feels that further valuable insight may be gained from future research endeavours.

One suggestion is to possibly modify the current study's questionnaire based on recommendations from the researcher, and to analyse the data within this context. Alternatively, this proposed modified questionnaire may be administered to a new sample followed by analysis. The researcher conducted validation exercises within the current study in order to establish construct validity. Establishing construct validity has been explained in preceding chapters to be one of the most important criteria in establishing the validity of an questionnaire. However, future endeavours may want to explore other types of validity such as face validity to further improve the structure of the current questionnaire. In future studies or administration of the current questionnaire, modification may entail re-structuring the questionnaire in terms of grouping items under the specified sub-scales differently. Alternatively, different versions of the same questionnaire may be constructed for each group of raters in the 360-degree feedback process.

Secondly, to increase generalisabiltiy of findings and control for extraneous variables for future administrations of the current questionnaire, demographic and contextual information for each participant and organisation can be gathered. Following from this, using this contextual information may allow for narrower comparisons in order to examine dimensions of performance more effectively. This approach could not be replicated in the current study owing to the fact that the archival data used did not contain any demographic or contextual information.

Whilst the current study used data from a generic questionnaire administered to participants from different organisations, future research should use questionnaires that have been modified for specific organisations. This approach may also ensure that any internal and external cultural factors are taken into account and controlled during analysis. Furthermore, inclusion of a qualitative component will provide further insights and contextual information. This qualitative data may be able to enable future researchers to identify the specific components of leadership performance that are consistent between rater groups which quantitative data may not be able to do.

Lastly, given the difficulty in isolating the theoretical dimensions that constitute the concept of leadership performance, future research could explore the different components of leadership performance utilising a different approach. As mentioned above a qualitative component which was not available in the current data, may be gathered and explored to isolate the theoretical dimensions. This may provide a stronger theoretical foundation for the questionnaire of this study, and may possibly yield clearer information regarding the areas of consistency in the 360-degree feedback context.

6. Conclusion

In today's volatile, changing and competitive world of work, organisations are continuously looking for ways to increase and measure performance accurately and fairly (Banu & Umamaheswari, 2009). This need has spurred on the popularity of using 360-degree feedback questionnaires to measure performance in organisations. Research in this area has mainly focused on the psychometric properties of 360-degree feedback instruments and processes, with an increase in research examining the ratings between and within groups of respondents (Hassan & Rohrbaugh, 2009). In line with this trend the current study has investigated relationships between groups of respondents, as well as between the ratings of respondents and participants to contribute to knowledge regarding the manner in which different groups in organisations perceive each other, and dimensions that contribute to performance.

The current study found that the degree of consistency varied between the different groups of raters. In line with the expectation of finding differences between raters, self-ratings were significantly different from those of colleagues, managers and subordinates. However, surprisingly, self-ratings were found to be lower than all other rater groups. Even though the majority of research findings in this area found that self-ratings were higher than others' ratings, this study highlights the potential consequences of self-selecting raters. Favourable rater-ratee relationships have been shown to display higher ratings compared with self-rating (Antonioni & Park, 2001; Fleenor et al., 2010). Raters may intentionally inflate ratings for political reasons, reluctance to rate peers harshly, or as a result of the pervading culture and context (Fleenor et al., 2010). This points to the need for the appropriate choice of raters for completing the particular 360-degree feedback questionnaire that was used to gather data for this study.

Within the current study, ratings for all groups showed that, in terms of leadership performance, self-ratings provided the most nuanced picture of the self followed by ratings of managers, colleagues and subordinates. This finding makes sense as self-ratings provide a unique perspective on performance and the individual (Beehr et al., 2001). However, self-ratings may not always be accurate or valid (Beehr et al., 2001).

Therefore, although this gives the most nuanced picture, self-awareness needs to be high in order to yield ratings that are accurate and insightful in terms of leadership performance. Apart from managers and colleagues, results showed significant differences between all rater groups. This is explained by the fact that in general individuals do not act similarly among rater groups that provide differing perspectives (Fleenor et al., 2010). In line with this, and in the context of this study, managers and colleagues may have had similar frames of references that may account for the similarity in ratings of participants.

It was also discovered that differences between ratings were not a product of rater group membership, but rather of individual differences. This needs to be taken into account when interpreting ratings, and perhaps each individual and his or her ratings should be looked at in isolation. This is because averaging ratings tend to mask important differences, or alternatively highlight inconsequential ones (Atwater et al., 1998). Furthermore, the use of the qualitative component, which was not provided in the original data set, will also yield richer insight into leadership performance.

Strong correlations were observed between all subscales across all rater groups and significant differences were found between rater groups across all subscales. However, results showed no clear pattern or grouping to isolate particular areas of congruency or lack thereof within the particular 360-degree feedback questionnaire and its subscales. This could mean that this particular questionnaire can only assess leadership performance as a single construct without being able to distinguish differences between ratings for the subscales it comprises. Alternatively, a common understanding of performance should be provided to participants and respondents in terms of subscales to ensure patterns of differences across subscales. Therefore, it is recommended that future research efforts could explore the different components of leadership performance in greater detail and with greater rigour. This may provide a stronger theoretical foundation for the questionnaire of this study, and may possibly yield clearer information regarding the areas of consistency in the 360-degree feedback context.

Overall, it is hoped that the present study will stimulate researchers to explore the dynamics of the current 360-degree feedback questionnaire as well as the rich area of

360-degree feedback. A deeper understanding into the areas of consistency and inconsistency with an empirical approach is required to improve the use and understanding of 360-degree feedback methods (Hasssan & Rohrbaugh, 2009). Taking into account all the findings of the current study, in relation to the particular questionnaire that was used in this study, interpretation of future ratings from the administration of this particular 360-degree feedback questionnaire is hoped to be more accurate, and provide greater depth into leadership performance of individuals. This can translate into better evaluation and the design of leadership development programmes within organisations that can be based on findings from the administration of this questionnaire.

Reference List

Alimo-Metcalfe, B. (1998). 360 Degree Feedback and Leadership Development. *International Journal of Selection and Assessment, 6,* 35-44.

Antonioni, D., & Park, H. (2001). The Relationship Between Rater Affect and Three Sources of 360-Degree Feedback Ratings. *Journal of Management, 27,* 479-495.

Arnold, C. L. (1992). An Intorduction to Hierarchical Linear Models. *Measuremtn and Evaluation in Counselling and Development, 25,* 58-90.

Atkins, P. W. B., & Wood, R. E. (2002). Self- Versus Others' Ratings as Predictors of Assessment Center Ratings: Validation Evidence for 360-Degree Feedback Programs. *Personnel Psychology, 55,* 871-904.

Atwater, L. E., Brett, J. F., & Charles, A. C. (2007). Multisource Feedback: Lessons Learned and Implications for Practice. *Human Resource Management, 46*, 285-307.

Atwater, L. E., Ostroff, C., Yammarino, F. J., & Fleenor, J. W. (1998). Self-Other Agreement: Does it Really Matter?. *Personnel Psychology, 51,* 577-598.

Atwater, L., Waldman, D., Ostroff, C., Robie, C., & Johnson, K. M. (2005). Self-Other Agreement: Comparing its Relationship with Performance in the U.S. and Europe. *International Journal of Selection and Assessment, 13,* 25-40.

Banu, C. V., & Umamaheswari, P. (2009). A Study on 360 Degree Performance Appraisal Systems in Reliance Life Insurance, Udumalpet. *The IUP Journal of Management Research, 7,* 65-78.

Bartholomew, D. J., Steele, F., Moustaki, I., & Galbrath, J. I. (2008). *Analysis of Multivariate Social Science Data (2nd ed.).* New York: CRC Press Taylor & Francis Group.

Becton, J. B., & Schraeder, M. (2004). Participant Input into Rater Selection: Potential Effects in the Quality and Acceptance of Ratings in the Context of 360-Degree Feedback. *Public Personnel Management, 33,* 23-32.

Beehr, T. A., Ivanitskaya, L., Hansen, C. P., Erofeev, D., & Gudanowski, D. M. (2001). Evaluation of 360 Degree Feedback Ratings: Relationships with Each Other and With Performance and Selection Predictors. Journal of Organizational Behavior, 22, 775-788.

Calantone, R., & Vickery, S. K. (2009). Special Topic Forum on Using Archival and Secondary Data Sources in Supply Chain Management Research. *Journal of Supply Chain Management, Jul 2009,* 1-3.

Calhoun, A. W., Rider, E. A., Peterson, E., & Meyer, E. C. (2010). Multi-rater Feedback with Gap Analysis: An Innovative Means to Assess Communication Skill and Self-insight. *Patient Education and Counseling, 80,* 321-326.

Carretero-Dios, H., & Pérez, C. (2007). Standards for the Development and Review of Instrumental Studies: Considerations About Test Selection in Psychological Research. *International Journal of Clinical and Health Psychology, 7,* 863-882.

Carson, M. (2006). Saying It Like It Isn't: The Pros and Cons of 360-Degree Feedback. *Business Horizons, 49,* 395-402.

Conger, J., & Toegel, G. (2002). Action Learning and Multi-rater Feedback as Leadership Development Interventions: Popular but Poorly Deployed. *Journal of Change Management, 3,* 332-348.

Cook, J. (2010). *Thornhill Associates (Pty) Ltd.* Johannesburg: Thornhill Associates (Pty) Ltd.

Craig, S. B., & Clark, A. P. (2009). Book Reviews of Lepsinger, R. & Lucio, A. The Art and Science of 360 Degree Feedback. (2nd Ed.) San Francisco: Pfeiffer, 2009. *Personnel Psychology, 62,* 629-657.

Cramer, D. (2003). *Advanced Quantitative Data Analysis.* Maidenhead: Open University Press.

DeNisi, A. S., & Kluger, A. N. (2000). Feedback Effectiveness: Can 360-Degree Appraisals Be Improved?. *The Academy of Management Executive, 14,* 129-139.

Diefendorff, J. M., Silverman, S. B., & Greguras, G. J. (2005). Measurement Equivalence and Multisource Ratings for Non-Managerial Positions: Recommendations for Research and Practice. *Journal of Business and Psychology, 19,* 399-425.

Eichinger, R. W., & Lombardo, M. M. (2004). Perspectives: Patterns of Rater Accuracy in 360-Degree Feedback. *Human Resource Planning,* 23-25.

Fleenor, J. W., Smither, J. W., Atwater, L. E., Braddy, P. W., & Sturm, R. E. (2010). Self-other Rating agreement in Leadership: A Review. *The Leadership Quarterly, 2010,* 1-30.

Fletcher, C., Baldry, C., & Cunningham-Snell, N. (1998). The Psychometric Properties of 360 Degree Feedback: An Empirical Study and a Cautionary Tale. *International Journal of Selection and Assessment, 6,* 19-34.

Fletcher, C., & Baldry, C. (2000). A Study of Individual Differences and Self-Awareness in the Context of Multi-Source Feedback. *Journal of Occupational and Organizational Psychology, 73,* 303-319.

Forte, J. (2008). Give Feedback, Get Performance. The American Salesman, 53, 26-30.

Foster, C. A., & Law, M. R. F. (2006). How Many Perspectives Provide a Compass? Differentiating 360-Degree and Multi-Source Feedback. *International Journal of Selection and Assessment, 14,* 288-291.

Ghorpade, J. (2000). Managing Five Paradoxes of 360-Degree Feedback. *The Academy of Management Executive, 14,* 140-150.

Gillespie, T. L. (2005). Internationalizing 360-Degree Feedback: Are Subordinate Ratings Comparable?. *Journal of Business and Psychology, 19,* 361-382.

Hassan, S., & Rohrbaugh, J. (2009). Incongruity in 360-Degree Feedback Ratings and Competing Managerial Values: Evidence from a Public Agency Setting. *International Public Management Journal, 12,* 421-449.

Hooijberg, R., & Choi, J. (2000). Which Leadership Toles Matter to Whom? An Examination of Rater Effects on Perceptions of Effectiveness. *Leadership Quarterly, 11,* 341-364.

Howell, D. C. (2004). *Fundamental Statistics for the Behavioral Sciences* (5th Ed.). Belmont: Thomson Wadsworth.

Huck, S. W. (2004). *Reading Statistics and Research (4th ed.).* USA: Pearson Education Inc.

Joreskog, K., & Sorbom, D. (1993). *LISREL8:Structural Equation Modeling with the SIMPLIS Command Language*. Hilldale N.J.: Lawrence Erlbaum Associates.

Kanouse, D. (1998). Why Multi-Rater Feedback Systems Fail. *HRFocus, January 1998,* 3.

Kerin, R. (2010). The Process of Change in Leadership Development Using 360-Degree Feedback to study the Roles of Reflection, Planning, and Support. *The George Washington University, 2010,* 1-163.

Kluwin, T. N., & Morris, C.S. (2006). Lost in a Giant Database: The Potentials and Pitfalls of Secondary Analysis for Deaf Education. *American Annals of the Deaf, 151,* 121-128.

Leedy, P. D. (1993). *Practical research: Planning and design (5th ed.).* USA: Macmillan Publishing Company

Levine, M. (2010). Taking the Burn Out of the 360° Hot Seat. *T+D, August 2010,* 40-45.

Liviu, I., Emil, C., Salantă, I., & Butilcă, D. (2009). The Use of 360-Degree Feedback Method. *Annals of the University of Oradea, Economic Science series, 18,* 300-306.

Maylett, T., & Riboldi, J. (2007). Using 360° Feedback to Predict Performance. *T + D, Sep 2007,* 48-52.

Miles, J. & Banyard, P. (2007). *Understanding and Using Statistics in Psychology: A practical Introduction.* London: SAGE Publications Ltd

Monroe, J. (1995). Measurement For Beginners: Review of a Handbook of Test Construction: Introduction to psychometric Design by Paul Kline, Methuen, 1986. *New Ideas in psycho., 13,* 187-189.

Moses, J., Hollenbeck, G. P., & Sorcher, M. C. (1993). Perspective: Does "360-Degree Feedback" Result in Going Around in Circles?. *Human Resource Management, 32,* 283-297.

Neale, J., & Liebert, R. (1986). *Science and behaviour: An introduction to methods of research.* New York: Prentice Hall International.

O'Leary-Kelly, S. W., & Vokurka, R. J. (1998). Empirical Assessment of Construct Validity. *Journal of Operations Management, 16,* 387-405.

Patiar, A., & Mia, L. (2008). The Effect of Subordinates' Gender on the Difference Between Self-ratings, and Superiors' Ratings, of Subordinates' Performance in Hotels. *International Journal of Hospitality Management, 27,* 53-64.

Peiperl, M. A. (2001). Getting 360° Feedback Right. *Harvard Business Review, January 2001,* 142-147.

Pollman, V. A. (1990). Some Faulty Assumptions that Support Using Multi-Rater Feedback for Performance Appraisal. Sourced from London, M., Wohlers, A. J., &

Gallagher, P. (1990). 360-Degree Feedback Surveys: A Source of Feedback to Guide Management Development. *Journal of Management Development, 9,* 17-31.

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioural research: Methods and data analysis.* USA: McGraw-Hill Inc.

Schreiber, J., Stage, F. K., King, J., Nora, A., & Barlow, E. A. (2006). Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review. *Journal of Educational Research, 99,* 323-337.

Singh, N. R. (2010). A Conceptual Vision on 360 Degree Assessment for Entrepreneurial Appraisal. *Advances in Management, 3,* 25-31.

Sireci, S. G. (2007). On Validity Theory and Test Validation. *Educational Research, 36,* 477-481.

Smith, A. F. R., & Fortunato, V. J. (2008). Factors Influencing Employee Intentions to Provide Honest Upward Feedback Ratings. *J Bus Psychol, 22,* 191-207.

Spangenberrg, H. H., & Theron, C. (2004). Development of a Questionnaire for Assessing Work Unit Performance. *SA Journal of Industrial Psychology, 30,* 19-28.

Stangor, C. (2010). *Introduction to Psychology.* Irvington: Flat World Knowledge.

Sytsma, M.& Hezlett, S. A. (1998). Trait, Rater and Level Effects in 360-Degree Performance Ratings. *Personnel Psychology, 51,* 557-576.

Toegel, G., & Conger, J. A. (2003). 360-Degree Assessment: Time for Reinvention. *Academy of Management Learning & Education, 2,* 297-311.

Tosti, D. T., & Addison, R. M. (2009). 360-Degree Feedback: Going Around in Circles?. *Performance Improvement, 48,* 36-39.

Van der Heijden, B. I. J. M., & Nijhof, A. H. J. (2004). The Value of Subjectivity: Problems and Prospects for 360-Degree Appraisal Systems. *International Journal of Human Resource Management, 15,* 493-511.

Wech, B. A., & Heck, A. L. (2004). An Introduction to hierarchical Linear Modeling for Marketing Researchers. *Marketing Bulletin, 15,* 1-8.

Whitley, Jr., B. E. (2002). *Principles of Research in Behavioral Science* (2nd Ed.). New York: McGraw-Hill.

Weigelt, J. A., Brasel, K. J., Bragg, D., & Simpson, D. (2004). The 360-Degree Evaluation: Increased Work with Little Return?. *Current Surgery, 61,* 616-626.

Wimer, S. (2002). The Dark Side of 360-Degree Feedback: The Popular HR Intervention Has an Ugly Side. *TD, Sep 2007,* 37-42.

Woo, S. E., Sims, C. S., Rupp, D. E., & Gibbons, A. M. (2008). Development Engagement Within and Following Developmental Centers: Considering Feedback Favorability and Self-Assessor Agreement. *Personnel Psychology, 61,* 727-759.

Yammarino, F. J., & Atwater, L. E. (1993). Understanding Self-Perception Accuracy: Implications for Human Resource Management. *Human Resource Management, 32,* 231-247.

Yogun, E. (2009). The Traditional Problem of Modern Technique: Acceptance of 360 Degree System. E-*Journal of New World Sciences Academy Social Sciences, 4*, 198-209.

Zimmerman, R. D., Mount, M. K., & Golf III, M. (2008). Multisource Feedback and Leaders' Goal Performance: Moderating Effects of Rating Purpose, Rater Perspective, and Performance Dimension. *International Journal of Selection and Assessment, 16,* 121-133.

APPENDIX 1

THORNHILL ASSOCIATES (PTY) LTD
Reg 1995/00 1962/07
PO Box 1943, Gallo Manor, 2052, South Africa

22 April 2010

TO WHOM IT MAY CONCERN

Ms Sarika Kara will be conducting research on data from the Thornhill Associates database for degree purposes. The database consists of scores representing ratings of participants by respondents for the purpose of feedback. The data that will be supplied will not include any names of those giving or receiving feedback, but will simply be the scores for statistical treatment.

We do give her permission to use the data and will be collaborating with her in making it available to her.

All participants in Thornhill 360-degree feedback questionnaires agree to terms and conditions which include permission for the anonymous use of the data in academic research.

We look forward to this useful collaboration between ourselves and the University.

Managing Director
Counselling Psychologist

www.thornhill.co.za
+27 11 608 3089 / +27 82 895 8023
Managing Director: MA (Couns. Psy.) (Unisa); MA (Cape Town)

APPENDIX 2

LEADERSHIP3 360° FEEDBACK QUESTIONNAIRE ITEMS
© J Cook. 3 June 2005

L3.0

HEAD: CREATING DIRECTION

1. Identifies the "big picture" strategic context of problems and opportunities
2. Communicates a clear, powerful and attractive vision for the future of the organisation
3. Contributes a very thorough understanding of how business works
4. Analyses complex situations insightfully
5. Comes up with effective solutions to business problems
6. Is open and receptive to the new ideas of others and different ways of doing things
7. Introduces his/her own creative and innovative thinking
8. Challenges existing ways of doing things
9. Takes tough decisions when necessary, even when they are unpopular
10. Encourages realism in decisions by finding and facing the truth, even when it is unpleasant

HEART 1: ENABLING OTHERS

11. Listens attentively and accurately
12. Treats all people with respect, regardless of race, gender, or other diversity
13. Actively builds on the benefits of diversity to draw the best out of all people
14. Allows others to work in their own way, recognising their strengths, limitations and dreams
15. Contributes supportively and expertly to the success of peers and superiors
16. Demonstrates a set of values that subordinates respect and buy into
17. Supports others' decisions
18. Celebrates others' successes and achievements
19. Expresses confidence in the ability of others
20. Inspires support and enthusiasm in others for a shared vision and mission

HEART 2: ENABLING SELF

21. Recognises his/her own emotions and how they impact on others
22. Is aware of his/her strengths and weaknesses and is not afraid to disclose them
23. Obtains and attends to feedback from other people without defensiveness
24. Comes across as confidently optimistic, even when others are discouraged
25. Maintains a high level of personal energy and determination
26. Maintains composure and self-control under difficult circumstances
27. Steps forward to take the initiative unprompted, regardless of his/her formal position
28. Is consistent in what he/she says and does in all situations
29. Is consistently dependable in all situations
30. Keeps going with resilience in the face of difficulties or opposition.

HEART 3: ENABLING GROUPS AND THE ORGANISATION

31. Sums up decisions and puts what the team is thinking into words or action steps
32. Spots potential conflict, brings disagreements into the open, and deals with them promptly
33. Comes across in an assured, forceful, impressive, decisive and unhesitating way
34. Encourages constructive debate and open discussion
35. Is effective at winning people over to his/her point of view
36. Reads the organisation's power dynamics so well he/she can deal with politics effectively
37. Builds consensus and support for initiatives, especially among key decision makers
38. Forms strong networks with people across the organisation
39. Attracts the most talented staff to work with him/her
40. Gains unusual loyalty from staff to the extent that they go the extra mile for him/her.

HANDS: EXECUTING

41. Insists on high standards of performance and service to clients / customers / colleagues
42. Can be counted on personally to deliver on time and up to standard
43. Clearly states what is required, from whom, by when, how and why
44. Meets with direct reports often to confirm direction and monitor progress
45. Gives accurate, insightful and fair feedback on performance
46. Has the courage to implement tough consequences for poor performance
47. Sets up and/or improves systems, processes and structures to achieve optimal results
48. Takes active steps immediately to ensure the team obtains the resources needed to do the job
49. Ensures that staff receive training or coaching to equip them for specific requirements.
50. Empowers others to take initiative, make decisions and work effectively on their own.

LEADERSHIP IMPACT ITEMS

51. When I am with this person I feel understood
52. When I am with this person I feel respected
53. When I am with this person I feel stretched to go beyond my comfort zone
54. When I am with this person I feel that I want to do things right all the time
55. When I am with this person I feel confident in my own abilities
56. When I am with this person I feel appreciated and recognised
57. When I am with this person I feel energised
58. When I am with this person I feel the urgency of what we need to do
59. When I am with this person I feel that work is fun
60. When I am with this person I feel that we have a future in this organisation
61. When I am with this person I feel motivated to give my very best all the time
62. When I am with this person I feel proud to be part of a winning team

Comments

APPENDIX 3

Participants:			Respondents:		
Number of respondents	**Number of participants with that many total respondents**	**Number of participants with that many respondents who have responded**	**Respondent type**	**Total respondents**	**Have responded**
0	355	375	Manager	2768	2505
1	50	52	Customer	1	1
2	38	48	Direct report	36	35
3	60	81	Subordinate	8008	7562
4	199	208	Regional Manager	4	4
5	253	276	Senior	23	23
6	330	332	Internal client	2	2
7	300	317	Team Member	53	40
8	287	293	Self	2646	2610
9	279	302	Other	3430	3162
10	267	249	Research Exec	31	26
11	179	170	Broker Consultant	4	2
12	121	116	Colleague	7289	6683
13	91	82	Key Broker	12	7
14	73	69			
15	64	41	**Total**	24307	22662
16	30	28			
17	39	26			
18	19	14			
19	19	8			
20	16	11			
>20	64	35			
Total participants:	**3133**	**3133**			

UMTHOMBO WOPHUHLISO LOLUNTU
School of Human & Community Development

Psychology
School of Human & Community Development
University of the Witwatersrand
Private Bag 3, WITS, 2050
Tel: (011) 717 4500 Fax: (011) 717 4559

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

APPENDIX 4

DATE: 11/04/2010

Good Day

My name is Sarika Kara and I would like to invite your organisation to participate in a research study that I am currently conducting for the purposes of obtaining my Masters degree in Industrial Psychology at the University of the Witwatersrand.

My research focuses on leadership performance in the context of 360-degree feedback. Past research shows that 360-degree ratings can provide insight into specific performance related constructs, such as leadership performance. Furthermore, in the context of the 360-degree feedback process, leadership has been shown to be positively related to accurate self-perceptions with the result of increased performance. Therefore, my research will also investigate this theoretical link. Comparisons between participants and respondents in this unique context will provide a unique way of investigating leadership performance of individuals.

I would like to request your permission to access your existing data base obtained from the administration of your leadership 360-degree feedback questionnaire. There are no further requirements regarding participation in this research.

This research is not intended to investigate any individuals or organisations, but rather to establish general trends among managers and professionals. As such, responses will only be examined in relation to all other responses. In addition, the results obtained through my statistical analyses will be kept confidential. Feedback and general results will be made available to you in the form of an executive summary.

Thank you for taking the time to read this letter. Should you have any queries, please do not hesitate to contact either myself, or my supervisor, Fiona Donald.

Yours Sincerely,

Miss Sarika Kara
Industrial Psychology Masters Student
082 732 9169

Mrs Fiona Donald
Supervisor /Lecturer
011 717 4507

128

APPENDIX 5

Thu, May 27, 2010 9:30:19 AM
RE: L3 research - ethics clearance
From: t
To: Fiona Donald <Fiona.Donald@wits.ac.za>
Cc: Sarika Kara <sarikakara@yahoo.com>

Hi Fiona,

1. Happy for the results to be published anywhere.

2. Also in principle happy for names to be used - I did suggest to Sarika that I would like to see the text before finally approving the use of names, although point 3 probably deals with this.

3. Only consider me as fourth author if I make a material contribution. But probably not, as it may appear to compromise the independence of the study (not that we are validating an instrument, mind you - so let's wait and see).

Thanks.

Envy you the thought of a sabbatical!

Regards,

APPENDIX 6

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

HUMAN RESEARCH ETHICS COMMITTEE (SCHOOL OF HUMAN & COMMUNITY DEVELOPMENT

CLEARANCE CERTIFICATE	**PROTOCOL NUMBER: MORG/10/003 IH**
PROJECT TITLE:	Consistence between raters in a 360^0 leadership feedback instrument
INVESTIGATORS	Sarika Kara
DEPARTMENT	Psychology
DATE CONSIDERED	26/05/10
DECISION OF COMMITTEE*	Approved

This ethical clearance is valid for 2 years and may be renewed upon application

DATE: 28 April 2011

CHAIRPERSON
(Professor G. Eagle)

cc Supervisor:

Fiona Donald
Psychology

DECLARATION OF INVESTIGATOR (S)

To be completed in duplicate and **one copy** returned to the Secretary, Room 100015, 10th floor, Senate House, University.

I/we fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure be contemplated from the research procedure, as approved, I/we undertake to submit a revised protocol to the Committee.

This ethical clearance will expire on 31 December 2012

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

130