The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

University of Cape Town

Validity of an assessment centre as a measure of competency ratings for development in a large financial institution.

By

Barbara Bashabe

(BSHBAR003)

A dissertation submitted in partial fulfillment of the requirements for the award of the Degree of Master of Commerce in Organisational Psychology

Faculty of Commerce

University of Cape Town

2006

DECLARATION:

This work has not been previously submitted in whole, or in part, for the award of any degree. It is my own work. Each significant contribution to, and quotation in, this dissertation from the work, or works of other people has been attributed, and has been cited and referenced.

Signature: Signed by candidate Date: 23/MAY/2007

ACKNOWLEDGEMENT

I would like to thank God, my heavenly father, who has blessed me abundantly and brought me thus far in life to accomplish this piece of work.

To my family, your prayers, love and support, I will always be grateful.

This research paper would also not have been successfully completed without the time, advice and support from my supervisor, Associate Professor Burger van Lill. In addition, my heartfelt gratitude goes out to the financial organisation that provided the data and more specifically to the assessment manager, whose dedication to the advancement of knowledge is applaudable.

Last but not least, I would like to thank my friends who have guided me and offered tremendous support throughout this project.

University of Cape Town

TABLE OF CONTENTS

University of Cape Town

University of Cape Town

LIST OF FIGURES AND TABLES

TABLES

University of Cape Town

University of Cape Town

FIGURES

University of Cape Town

ABSTRACT

Assessment centres are used to make decisions for employee advancement and development within organisations. These centres comprise of a number of instruments which are used to establish competencies needed to perform effectively. The validity of this measure is yet to be fully ascertained as previous studies have shown that although assessment centres have predictive validity, they lack construct validity. The research study therefore set out to establish whether assessment centres used within financial institutions have predictive validity but lack construct validity. For purposes of this study, two sets of data each comprising of responses from 120 and 91 employees were used. This data was retrieved from a financial institution within the Western Cape.

Contrary to previous studies, there was low predictive validity within this study. Some of the factors that could have contributed to the low correlation between performance and assessment ratings include a rather small sample that was used, having different criterion being measured in the evaluations than in the assessment centres, range restriction, complexity of assessing behaviour, and having a number of dimensions from which to derive an Overall Assessment Rating (OAR). However, similar to previous studies, this study also found low construct validity amongst the ratings. In addition, a further analysis indicated that both performance and personality measures are valid methods in predicting performance.

University of Cape Town

1. INTRODUCTION

A number of methods have been developed to assess performance and potential growth of individuals within an organisation. However, assessment centres have gained popularity in their ability to predict successful on-the-job performance. Unlike other methods, assessment centres are able to simulate various work contexts and therefore can be effectively used to determine whether an individual has the necessary skills needed to perform the required task as per job requirement or not (Thornton & Byham, 1982). As a result, assessment centres are widely used by organisations to establish whether an individual possesses the necessary competencies for effective job performance. It is important to note that assessment centres are used to establish competencies needed to perform within other jobs other than the individual's current position. This is because, information regarding an employee's actual performance on the current job can be easily established through observations and performance appraisals (Thornton & Byham, 1982).

An assessment centre is a procedure used to measure an individual's potential to perform effectively using a number of measuring instruments (Thornton & Byham, 1982). The original assessment centres were referred to as industrial centres and were used to consider line management for promotion (Thornton & Byham, 1982). Today, assessment centres are widely used for selection and development purposes within organisations. However, studies have shown that assessment centres have predictive validity (Gaugler, Rosenthal, Thornton & Bentson, 1987) but lack construct validity (Sackett & Dreher, 1982). This means that assessment centres can successfully predict job success but do not measure the underlying construct. Therefore the way in which assessment centres work, is yet to be ascertained (Lance, Newbolt, Gatewood, Foster, French, & Smith, 2000; Klimoski & Brickner, 1987).

As a result, this study was set out to establish what the underlying constructs within an assessment centre in the Western Cape are and their relation to job

University of Cape Town

success. Individual assessments used for development purposes within this centre had to be analysed for purposes of this study in order to establish whether they offer a valid measure of competency ratings. The data that was used was retrieved from a financial institution within the Western Cape and it comprised of two data-sets that were used to determine the predictive and construct validity of assessment centres.

In addition, due to findings within previous studies, three hypotheses were derived for purposes of this study. The hypotheses stated that there will be a high predictive validity between assessment centre ratings and performance management ratings, there will be a low correlation of construct ratings across a single dimension and there will be a high correlation between ratings of different traits within a single exercise.

Predictive validity was determined by assessing the relationship between assessment center ratings and the subsequent performance ratings of individuals. With regards to construct validity, Campbell and Fiske's (1959) Multitrait-multimethod approach was used for analysis. In addition, an analysis of variance (ANOVA) was carried out in order to establish the extent of variance between the traits and methods used in order to make meaningful interpretations of the correlations. It should be noted that due to the limited sample size, statistical techniques such as factor analysis could not be used to discover the underlying patterns of relationships amongst the variables. Nonetheless, this study yielded some interesting findings some of which have been found within previous studies.

There are six chapters within this research paper. The first chapter comprises of a brief introduction to the study. The second chapter consists of previous studies that have been conducted within a similar research domain. The third chapter consists of the method used within the present study. The fourth chapter comprises of the results. The fifth chapter consists of a discussion of the results in relation to previous studies and the final chapter comprises of a concluding remark to the study.

University of Cape Town

It should be noted that generalisations based on the findings within this study should be done with caution as the sample was not large enough and representative of the population.

University of Cape Town

2. LITERATURE REVIEW

Overview

According to Edenborough (2005), there is no one measure that offers exhaustive information regarding the most appropriate path for individual development. However, assessment centres have gained popularity as effective measures of determining present and future individual job performance because these centres comprise of a combination of various assessment tools (Klimoski & Brickner, 1987; Tziner, Ronen & Hacohen, 1993). Therefore, detailed and comprehensive information regarding an individual's performance can be obtained. As a result, various studies have set out to establish whether assessment centres actually measure what they purport to measure and whether these centres are good predictors of future job performance.

Assessment Measures

Assessment is the process used to gather a wide array of information in order to reach a decision regarding an individual (Wheeler & Haertel, 1993; Foxcroft & Roodt, 2003). There are various methods used to gather individual information for development purposes. These include assessment centres, structured interviews and tests (Spencer, McClelland & Spencer, 1992). However, previous studies conducted by Smith and Boyle (as cited in Spencer et al., 1992) have shown that assessment centres have the highest criterion validity of .65 with regards to determining successful performance.

Assessment Centres

An assessment centre is an event comprising of established work simulations in which individuals' behaviour is observed by a group of trained observers, for specified behavioural dimensions applicable to a specific job (Arthur, Day, McNelly & Edens, 2003; Foxcroft & Roodt, 2003; Edenborough, 2005). Simulation exercises enable an individual to engage in a job situation and to display skills needed to perform effectively (Thornton & Byham, 1982). Assessment centres were first used by the military during the Second World

University of Cape Town

War for selection of officers (Edenborough). However, AT & T was the first company to use assessment centres for industrial purposes (Olshfski & Cunningham, 1986). Today, these centres are used by organisations to determine whether individuals possess required competencies to perform a specific task through the replication of behaviour (Edenborough).

There are various tools used within an assessment centre to determine suitability for a job. A leaderless group exercise is one example and within this method, participants are given a business problem to solve and the relevant emergent skills such as leadership skills and initiative are evaluated (Thornton & Byham, 1982). Another example is a background Interview. This comprises of an informal or formal discussion between two or more people. The purpose of the interview is to obtain information regarding an individual's previous work experience, education, biographical data as well as the individual's work preferences (Thornton & Byham, 1982). Tests are also used within an assessment centre to ascertain an individual's underlying personality structure (Thornton & Byham, 1982). Examples of tests used include personality and projective tests. On the other hand, written analysis exercises are used to measure an individual's cognitive and intellectual abilities (Thornton & Byham, 1982). In-baskets are also used and these comprise of various documents that are presented to an individual in order to establish how they would deal with the administrative work in a typical managerial job (Thornton & Byham, 1982). Within a business case study, participants are presented with a problem and are asked to recommend a solution (Thornton & Byham, 1982). These are some of the techniques that can be used within an assessment centre to establish an individual's competencies.

An assessment centre can also be used to identify potential and training needs amongst employees (Byham, 1971). These assessment centres used for development are sometimes referred to as development centres. One notable characteristic of development centres is that there is detailed feedback offered to an individual regarding their performance, as the focus is on growth and improvement of an individual's skills (Thornton & Byham, 1982; Lievens, 2002; Edenborough, 2005). When information retrieved from an

University of Cape Town

assessment center is communicated back to the individual, it provides a basis for self-insight and development planning. Likewise, management can use this information to further develop an individual and enhance their performance.

Boam and Sparrow (1992) maintain that employee development within an organisation focuses on getting the right skills for a particular job and making the best of what the employee has to offer. As a result, assessment centres used for development focus on establishing competencies needed for effective future performance. Results from assessment centres are usually used to formulate training and development plans for employees in order to improve job performance (Thornton & Byham, 1982). It is important to note that assessment centres establish job competencies and general abilities that are difficult to measure with the use of individual standardised psychological measures (Kuptsch, Kleinmann & Koller, 1998). This is because individual assessment methods usually measure competencies that have not been adequately defined. On the contrary, the use of a number of assessment measures within the assessment centre method enables the acquisition of comprehensive information regarding an individual's competencies. Therefore, more accurate results regarding an individual's abilities and skills can be obtained.

University of Cape Town

Competencies

Boyatzis (1982) popularised the current usage of the term competency in his book 'The competent Manager' (Woodruffe, 1992; Edenborough, 2005). A competency can be defined as an underlying characteristic or ability that enables an individual to perform a job effectively (Boyatzis, 1982; Hornby & Thomas, 1989; Woodruffe, 1992; Wheeler & Haertel, 1993; Edenborough, 2005). Competencies include skills, values, attitudes, individual attributes, abilities, accumulated knowledge or any other characteristic, which an individual can use, within a specific context, to accomplish a task. According to Boyatzis' model of effective job performance, competencies are necessary but not sufficient for effective job performance. The job demands (job requirements that necessitate specific capabilities) together with the organisational environment, and individual competencies determine specific

desirable actions or behaviour. Consequently, consistency in effective performance will occur if there is a 'fit' between these three components (Boyatzis, 1982).

In other words, for one to effectively define a competency, various aspects have to be taken into consideration. These include actions, places where the competencies are performed, sequence of behaviour, results and intended effects (Boyatzis, 1982). In addition, competencies to be measured have to be clearly defined and established based on prior competencies as demonstrated by superior performers (Edenborough, 2005; Spencer et al., 1992). As a result, in order for competency measures to be valid, the situational context in which they are being applied has to be considered. It is important to note that some jobs such as science, research and engineering do not provide assessment measures of performance (Edenborough). Consequently, in these types of jobs it is more relevant to assess whether or not the person in the job is following certain procedures or processes that are thought to be important to the organisation.

Validity

Validity is the extent to which an assessment tool measures what it is supposed to measure (Wheeler & Haertel, 1993). Assessment centres should be valid in order to be effective in measuring constructs. Foxcroft and Roodt (2003) maintain that assessment centres have gained in popularity because assessees can relate to the exercises and simulations as well as perceive them as relevant and appropriate within the work context. However despite evidence of face, content and predictive validity, assessment centres' construct validity is yet to be ascertained (Haaland & Christiansen, 2002).

Face Validity

This refers to credibility and acceptance of an assessment measure (Saunders, 2000; Foxcroft and Roodt, 2003). It is important to note that lack of face validity can have a negative effect on other forms of validity because face validity makes an assessment measure highly acceptable to participants. On

University of Cape Town

the contrary, low face validity can bias the participants' reactions to the assessment measures consequently, yielding inaccurate responses.

Criterion Validity

This form of validity is concerned with relating the assessment measurement to an external criterion of performance (Edenborough, 2005; Olshfski & Cunningham, 1986). It determines what a person's score should be on some criterion measure in order for one to successfully perform on the job (Saunders, 2000; Foxcroft & Roodt, 2003).

Predictive Validity

This is a subset of criterion validity and it is established by comparing assessment ratings to an individual's performance over time (Edenborough, 2005). Saunders (2000) maintains that predictive validity is difficult to establish, as it requires controlled research over time. Assessment centres have a relatively good predictive validity as these centres effectively determine successful job performance. However, it is important to establish how predictive a test needs to be for it to be effective, as low levels of predictive validity can still be useful in cases where the numbers to be selected are very small in relation to those being tested (Edenborough).

Concurrent Validity

This is most appropriate in situations where predictive validity cannot be easily established (Edenborough, 2005). Given the difficulties associated with establishing predictive validity, organisations use concurrent validity, as it can be determined immediately. It refers to the extent to which scores external to the assessment measure are used to confirm assessment ratings (Foxcroft & Roodt, 2003). Consequently, high performers are differentiated from low performers using an external measurement score and the results are compared to the assessment ratings. It is important to note that unlike predictive validity, concurrent validity is not affected by minor changes in criteria over time.

University of Cape Town

Construct Validity

This is the extent to which an assessment tool measures a particular construct or characteristic it is meant to measure (Foxcroft & Roodt, 2003; Edenborough, 2005). It is concerned with establishing whether the assessment tool is measuring what it is intended to measure. For example, if a test is meant to measure competencies, evidence should be found that the test it is actually measuring individual competencies. Establishing construct validity involves extensive studies and comparison with other assessment tools measuring a similar construct. In order to establish construct validity of assessment centres, discriminant and convergent validity have to be determined.

Discriminant Validity

This refers to the extent to various traits or abilities that are being measured within a method minimally correlate (Bycio, Alvares & Hahn, 1987; Shore, Thornton & Shore, 1990; Foxcroft & Roodt, 2003). It establishes 'heterotrait-monomethod' and 'heterotrait-heteromethod' correlations within the exercises.

Convergent Validity

On the other hand, convergent validity refers to the extent to which ratings of the same construct or ability are significantly correlated across dimensions (Bycio, et al., 1987; Shore et al., 1990; Foxcroft & Roodt, 2003). This is sometimes referred to as 'Monotrait-heteromethod' correlation.

Content Validity

This reflects relevant material in the assessment measure with relation to required aspects of a role or job (Edenborough, 2005; Saunders, 2000). Assessment measures should be evaluated for content validity in order for them to be effective. The content validity of assessment centres can be derived from their job related content (Byham, 1971).

University of Cape Town

Validity of Assessment Centres

The criterion validity of assessment centres is not disputed as numerous studies have shown that assessment centres are good at predicting future performance on the job (Schleicher, Day, Meyes & Riggio, 2002). Researchers have however failed to find evidence for construct validity. Some studies (Sackett & Dreher, 1982; 1984) have found evidence for discriminant validity but not convergent validity. Consequently, they argue against the use of assessment centres on construct validity grounds because there is no evidence to support consistency of behaviour across exercises. Sackett and Dreher maintain that the absence of construct validity could be as a result of situational exercise-specific determinants of assessee's behaviour.

A study by Tziner, Ronen and Hacohen (1993) that compared assessor ratings to supervisor ratings collected over a period of four years showed that assessment centres have long-term criterion validity. It was also established that managers as assessors are better predictors of overall job performance than psychologists. Psychologists and managers could be better suited to predict performance, however different schemes may cause either group of assessors to be attentive to different behaviour. Consequently, they suggested that in order to establish high construct and predictive validity, assessors should be chosen based on the criterion measure for which the assessment centre is to be validated. Lievens (2002) also verified that the choice of assessors played a big role in determining assessment centre validity. Lance, Newbolt, Gatewood, Foster, French and Smith (2000) also found that assessment centres demonstrated criterion related validity and not construct validity as general performance on the assessment measures was consistently related to the overall assessment rating (OAR). However, Bycio et al. (1987) maintain that construct validity would improve if competencies were measured within exercises rather than across dimensions.

Arthur et al., (2003) suggest that current estimates of assessment centre criterion validity could be underestimated as these are commonly derived from combining ratings from various dimensions in order to get an overall assessment rating. A metanalysis study by Gaugler, Rosenthal, Thornton and

University of Cape Town

Bentson (1987), found that assessment centres had a criterion validity of r = .37. They also found that OARs are related to measures of performance and success on the job.

Reviews by Huck and Bray (1976) as well as Klimoski and Strickland (1977), revealed that assessment centre ratings have a positive correlation with subsequent job performance. However, the positive correlation could have been as a result of subsequent developmental or administrative decisions that were made, based on the assessment centre results and hence were not an accurate reflection of effective job performance. In other words, individuals could get rewarded (for example, promotion, better salary and increased benefits) as a result of the good ratings from the assessment centre and not based on good performance. This could result in what Klimoski and Brickner (1987) referred to as criterion contamination, which means that the rating on one measure influences the rating on another. However, Howard (as cited in Tziner et al., 1993) recommended that in order to deal with criterion contamination, validity of assessment centres has to be established after a while, as decisions to positively reward an individual would not be based on previous assessment centre ratings but rather on the individual's performance. In addition, Olshfski and Cunningham (1986), maintain that in order to evaluate the predictive validity of assessment centres, results must be kept away from anyone involved in making decisions within the organisation.

On the other hand, however, Tziner et al. (1993) argued that studies conducted over a long period of time were ambiguous, as some showed no relationship between time and predictive validity while others found a positive relationship. In addition what is considered as successful performance may change over time due to changes in the external environment and hence studies of assessment centre validity carried out after awhile might not be relevant, as ratings will no longer be applicable.

Klimoski and Strickland (1977) gave yet another reason that could account for criterion contamination. They suggested that assessors' notion as to what

University of Cape Town

constitutes successful behaviour could be contaminated by the organisation's view of successful performance, likeability and an individual's appearance and hence lead to a bias in judgment in which these 'favourable' competencies are preferred rather than looking out for competencies that lead to effectiveness. Consequently performance decisions would be based on the assessment centre ratings, which could be contaminated.

Tziner et al. (1993) suggest that another way of establishing criterion validity is by using various methods to measure a construct. However, the use of various methods brings up the issue of construct validity. According to Bycio, Alvares and Hahn (1987), assessment centre measures are not cross-situational as specific behaviour is only manifested within a given context. Hence the various methods might not adequately measure similar constructs as the contexts within which the methods are used would be different. Neidig and Neidig (1984) also argued that non-existent convergent validity in assessment centres was as a result of the different contextual situations that required the manifestation of a variation in behaviour. Consequently, the lack of convergent validity in assessment centres could be attributed to differences in performance rather than measurement error, as behaviour is situationally determined (Sackett & Dreher, 1982; Neidig & Neidig). In addition, Sackett and Dreher maintain that there are wide variations between exercises for different kinds of behaviour to be manifested.

Despite evidence of predictive validity, Klimoski and Brickner (1987), question the effectiveness of assessment centres, as construct validity cannot be obtained. In other words, how can assessment centres be effective at measuring constructs if it cannot be established whether these centres are measuring the traits that are purported to be measured? Lievens (2002) found that the reason as to why construct validity cannot be easily established in cross-situational exercises was as a result of poor design of the method, assessor unreliability and inconsistent assessee performances. Improvements on the assessment methods such as having fewer dimensions to be rated and behavioural checklists yielded higher construct validity in some studies. Lievens also found that evidence for convergent validity was established when

University of Cape Town

there was consistency of behaviour amongst the assessees within the various exercises and when Industrial Psychologists and managers were used as assessors. In addition, it was also noted that in order to establish construct validity, there has to be a high inter-rater reliability as assessors usually rotate amongst exercises and hence do not rate the same individual (Lievens).

Olshfski and Cunningham (1986) also argue that one major constraint to assessment centre validity is the definition of concepts in precise, operational terms. For example, the definition of a manager could involve a number of concepts such as kind, yet firm; open to differences, yet not waffling; and flexible yet decisive (Olshfski & Cunningham). As a result of the broad definitions, attributes to look out for might not be so specific. This poses a major constraint for those who design assessment centres, as there is no precise theoretical anchor to operationalise the required competencies and yet measures require a more precise definition. Yet another constraint for management is the ease at which decisions can be based on in appropriate criteria derived from assessment centre ratings which could not be a true reflection of the job requirements. This could be as a result of assessment centres making use of definitions of competencies from job analyses in which the present occupant's contribution cannot be easily distinguished from the position itself. Markus, Thomas and Allpress (2005) also draw attention to the issue of competency definition. They maintain that accurate assessment of competencies is hindered by the broad terms used to define competencies related to contextual performance.

According to Hoeft and Schuler (2001), assessment centres exhibit good content validity and satisfactory criterion validity. However, poor outcomes are as a result of less than satisfactory construct validity. Although assessment centres have good content validity, very few studies have been conducted to address this aspect of assessment centres. The best way of establishing content validity is by assessing the linkage between job requirements and the constructs being measured (Hoeft and Schuler). In addition, in order to exhibit content validity all relevant knowledge aspects, skills, abilities and other characteristics have to be taken into account. However, Sackett and Dreher

University of Cape Town

(1982; 1984) found that content validity was more relevant if assessment centres were being used to establish current job performance (concurrent validity) rather than future training needs (predictive validity). They went on to say that when assessment centres are used as predictors of future performance, then evidence of predictive validity would have to be established.

In a study conducted by Robie, Osburn, Morris and Etchegaray (2000), it was established that the use of the same assessor on the measurement of constructs within an exercise and across various dimensions increased discriminant validity. However, this may be expensive for assessment centres that use one assessor to rate each candidate. However, Robie et al, also suggest that videotaping assessment centre performance could eliminate the costs involved with using so many assessors.

Contrary to previous studies, Shore et al. (1990) established construct validity amongst dimension ratings. However the validity was determined using final dimension ratings as determinants of underlying constructs rather than within exercise dimension ratings. They maintain that final dimension ratings are relevant and consequently should be used to determine construct validity, as these are used for decision making.

Binning and Barrett (as cited in Arthur, Woehr & Maldegen, 2000) note that the absence of construct validity amidst sufficient criterion and content validity is contradictory, as unitarian conceptualisation of validity suggests that content and criterion validity are just different ways of verifying construct validity of an assessment measure. However Sackett (as cited in Hoeft and Schuler, 2001) argues that the presence of content validity does not provide ultimate proof of criterion and construct validity.

The in-basket is sometimes used as a single measure for development purposes (Spangenberg & Theron, 2003). In a study establishing validity of the In-basket method as opposed to an assessment centre, Greyling, Visser & Fourie (2003) found no evidence for discriminant validity and very little

University of Cape Town

evidence for convergent validity within this assessment technique. Their study also yielded poor predictive validity and hence suggested that the in-basket should never be used as a single method for measuring competencies. They suggested however, that careful development and application of the in-basket, particularly with regard to the scoring methods, might improve its validity (Spangenberg & Theron).

Reliability of Performance Ratings

Murphy and Davidshofer (1991) state that assessment measures are not entirely consistent and therefore this impacts on both the construct and predictive validity of a measuring instrument. This means that in order for a device to be valid, it has to be reliable. On the other hand however, reliability is not a sufficient condition for validity. The consistency of any measuring device is referred to as reliability (Kerlinger & Lee, 2000; Foxcroft & Roodt, 2003).

Performance evaluation is one of the most widely used techniques for establishing an individual's performance on the job (Viswesvaran, Ones & Schmidt, 1996). As a result it is important to establish the consistency of performance ratings because important decisions are usually based on these scores. Performance evaluations are subjective scores derived from peer, supervisor, subordinate, self, customer assessments of an individual (Viswesvaran, Ones & Schmidt). In their meta-analysis study, Viswesvaran, Ones and Schmidt found that supervisory ratings were more reliable than peer ratings.

It is also important to note that Viswesvaran, Ones and Schmidt (1996) also compared inter-rater and intra-rater performance ratings, and found that supervisory ratings had higher inter-rater reliability than peer ratings. In addition, they found that the intra-rater reliability scores were higher than the inter-rater reliability scores. The findings of this study showed that the choice of assessors can significantly impact on the performance ratings, and this could in turn affect the consistency of the scores.

University of Cape Town

South African Context

According to Jansen and De Jong (as cited in Greyling, Visser and Fourie, 2003), assessment centres have increased in popularity within South Africa because of their objectivity and emphasis on job related traits as opposed to individual skills, hence minimising bias (Saunders, 2002). Assessment centres focus on job related traits rather than language abilities and consequently minimise the influence that culture would have on assessment. In addition, studies done in South Africa have shown that although assessment centres lack construct validity, these centres are good predictors of job performance (Spangenberg & Theron, 2003). However, it is important to note that culture has a significant impact on attributes and behaviour that are measured by assessment centres (Briscoe, 1997). Consequently, the content validity of assessment centres within a multi cultural context like South Africa is highly questioned.

Saunders (2002) maintains that the assessment method in South Africa is culturally incompatible, to a certain extent, with some of the cultures as represented in the assessee group. However, Saunders goes on to argue that assessment centres predominantly observe behaviour and as a result minimise the influence that culture would have on assessment as compared to other assessment methods. This is because behaviour techniques do not rely heavily on language abilities and thus culture bias is minimised. In addition, the exercises focus on job related activities. It should be noted though, that some methods used in assessment centres require language proficiency such as interviews and tests.

Relevance of Rating Scale Content

Another issue that has to be addressed whilst analysing the correlation coefficients of the assessment centre variables, is the relevance of the rating scale content. This is also known as the 'content issue' in rating scales (Kavanagh, MacKinney and Wolins, 1971). Although there is little evidence to show that personality traits account for little variance in job performance (Kavanagh et al.), it is important to establish whether an employee's

University of Cape Town

personality affects their performance on the job. The National Industrial Conference Board (as cited in, Kavanagh et al.) maintains that although assessment measures continue to evaluate personality traits, overt or performance-oriented behaviours are more relevant to job performance than the personality-oriented behaviours (covert behaviour). Likewise, Ghiselli and Brown (1955) go on to state that the most relevant content for rating scales is performance traits. Consequently, the relative representativeness of traits has to be established from the analysis of correlations in order to ascertain whether performance traits contribute more towards on-the-job performance than personality traits.

Conclusion

There are various tools that can be used to assess individuals for development purposes. Research has shown that assessment centres are the most widely used measures as these centres enable the assessment of individuals, using a wide array of techniques in a simulated work context. In addition various studies have established that assessment centres have criterion and content validity and consequently can predict successful job performance. However, the presence of construct validity is highly disputed. It is important to note that validity can never be regarded as established once and for all or universally for any particular assessment measure as different situational contexts determine the manifestation of appropriate behaviour (Gaugler et al., 1987). Consequently, in a multicultural context like South Africa, a balance has to be reached on the design and use of assessment centres in order to ensure that the centres actually do measure what is set out to be measured. In addition, it is imperative to establish whether performance traits are more relevant to job performance than personality traits.

The research question within this study aimed at establishing whether individual assessments used for development within a large financial institution offer a valid measure of competency ratings. The main purpose of this study therefore was to determine the predictive and construct validity of assessment centres. As mentioned above, in order for a measuring device to have validity, it has to be reliable. Therefore, the internal consistency of the

University of Cape Town

competency scores was also ascertained as well as the reliability of the performance ratings. There were three hypotheses within this study and the first one claimed that there would be a high predictive validity between assessment centre ratings and performance management ratings. The second hypothesis stated that the construct ratings across a single dimension would minimally correlate (Convergent validity). Lastly, there would be a high correlation between ratings of different traits within a single exercise (Discriminant validity). Therefore the study expected to find that development centres can effectively predictive performance of an individual but lack construct validity.

University of Cape Town

3. METHOD

Participants

Competency ratings of employees of a financial institution within the Western Cape were used within this study. There were two data-sets that were used for this study because a single data set comprising of individuals' raw competency scores as well their subsequent performance ratings could not be retrieved. The first set of data that was used to establish predictive validity comprised of 120 employees who had participated in development centres for the last three years. They had been assessed at first, second and functional management levels. The second set of data comprised of 91 employees. Within the first data-set, twenty competencies were assessed and five performance ratings were taken as illustrated in table 3.3.

The second data-set that was used for establishing construct validity comprised of 91 employees and the competencies with which these employees were assessed against were identified as those important to function successfully as a manager at a 'People Leader' level. There are ten competencies that were assessed, using seven methods. It should be noted that within this data-set, not all employees were assessed on similar competencies as well. An illustration detailing the competencies and methods used within the assessment process is shown in table 3.4.

University of Cape Town

Race and Gender distribution

The racial classification comprised of two groupings and these were Black and White. The Black category consisted of all employees who were classified as Indian, African and Coloured. The percentage of Black to Whites was much higher within the first data-set, while the second data set had an equal distribution of race. On the other hand, however the gender distribution within the first data-set was almost equal, but there were more males than females. The distribution of race and gender within these two groups has been illustrated in figures 3.1 and 3.2. In addition, the percentage distribution of

race and gender has also been shown in tables 3.1 and 3.2.

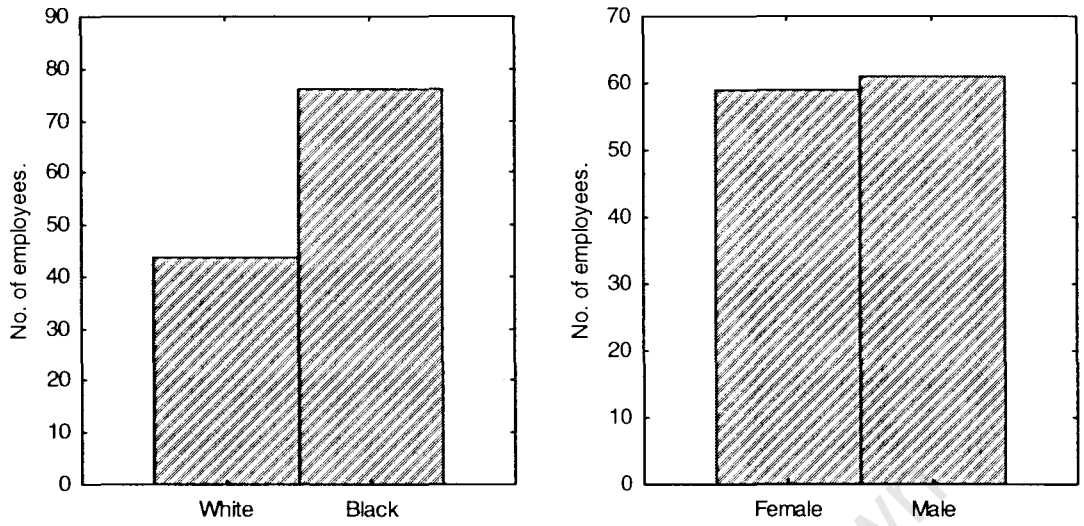


Figure 3.1. Race and gender distribution within the first data-set

Table 3.1.

Race and gender distribution within the first data-set

	Count	Percent
White	44	37
Black	76	63

	Count	Percent
Female	59	49
Male	61	51

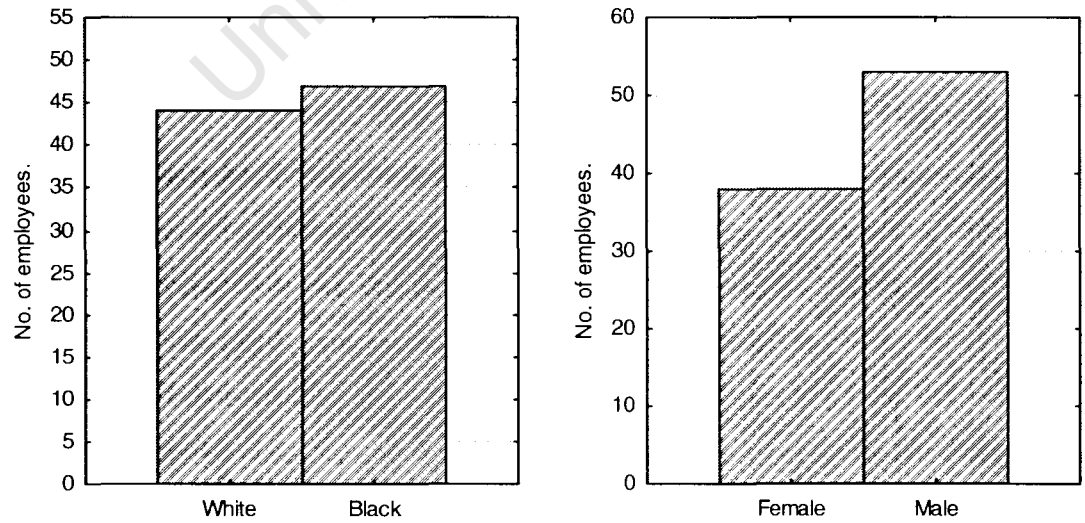


Figure 3.2. Race and gender distribution within the second data-set

University of Cape Town

Table 3.2.

Race and gender distribution within the second data-set

	Count	Percent
White	44	48
Black	47	52

	Count	Percent
Female	38	42
Male	53	58

Procedure

As a result of time constraints regarding the research project, previously collected information had to be used. This data was retrieved from a financial institution within the Western Cape. There were two sets of documented data that were used within this study. The first set of data consisted of compiled data that had received some form of selection and summarising. The assessment centre data comprised of an overall assessment rating for the individuals. In addition, only the final performance ratings were obtained. The names of the participants had been removed and hence the information was anonymous. This competencies and performance ratings used within this data-set are illustrated in table 3.3.

University of Cape Town

Table 3.3.

Competencies and Performance Ratings for the first data-set.

Competencies	Performance Ratings
1. Thinking Skills	1. Performance Rating taken on 01 July 2001
2. Business Awareness	2. Performance Rating taken on 01 July 2002
3. Aligning Performance for Success	3. Performance Rating taken on 01 July 2003
4. Commitment	4. Performance Rating taken on 31 Dec 2003
5. Team Orientation	5. Performance Rating taken on 31 Dec 2004
6. Client Focus	
7. Ownership Assuming	
8. Decision Making	
9. Initiating Action	
10. Planning and Organising.	
11. Continuous Learning	
12. Building Relationships	
13. Facilitating Change	
14. Energy	
15. Enterprise Innovation	
16. Building Successful teams	
17. Developing Others	
18. Adaptability	
19. Risk Taking	
20. Stress Tolerance	

University of Cape Town

It is important to note that the competencies which were assessed for each individual differed. In other words, not all participants were assessed on all the competencies. Therefore four sub-groupings as shown in the results section, were made for easier analysis within the first data-set.

The second set of data comprised of raw data that the researcher retrieved from the organisation. There was no form of processing that had been previously done on this data, as it comprised of raw scores. There was a wide range of information and the selection process used to choose the data to use within this study was based on the managerial level that had the most data. Consequently, the most recent assessment centre ratings for the 'People Leader' managers were retrieved. The data competencies within this data-set are illustrated below.

Table 3.4.

Competencies and Assessment Methods used

	Assessment Method						
Competency	Business Case Presentation	Career Interview	Written Exercise	OPQ	Role Play	Group Exercise	In-Basket
Customer Focus	●		●	●	●		●
Building Relationships		●		●			
Commitment		●		●	●	●	
Aligning Performance for Success				●	●		●
Facilitating Change	●	●		●	●	●	
Thinking Skills	●		●	●	●	●	●
Decision Making	●		●	●	●	●	●
Initiating Action	●		●	●			●
Planning & Organising				●			●
Business Awareness	●			●			

University of Cape Town

Methodology

The methodology used within this study can be described as cross sectional as data within different contexts is being investigated in order to establish a relationship (Collis & Hussey, 2003). The main concept within this study is the validation of development centres and this has been accomplished by analysing individuals' competency scores obtained within the development centre and by comparing these scores with subsequent performance ratings. More specifically, data derived from development centres has been compared to the performance ratings of individuals in order to establish predictive validity. In addition, the development centre ratings derived from various methods are being statistically analysed to establish construct validity.

The consistency of the rating scale content as well as the reliability of the performance scores have also been ascertained by comparing the scores within the two assessment techniques. Furthermore, the relevance of content ratings has been established by carrying out an analysis of variance on the correlation matrices between the variables.

University of Cape Town

Measuring Instruments

Performance Ratings

The first data-set that was used to establish predictive validity comprised of 5 performance ratings. The manager and employee agree upon a performance contract, which is reviewed in the middle of the year and a final evaluation is made at the end of the year. The performance agreements differ from one employee to another depending on their position but they typically include key result areas, key performance indicators, measures and targets and a rating per each key result area on a 5 point scale where below 2 is regarded as under performing, 2.5 as acceptable performance and above 3.5 as excellent performance. The performance agreement also includes individual deliveries (+-60%), a team component (+-30%) and a "living the values" section (10% measured using a 360^0 questionnaire). It is important to note that some of the performance data was based on specialist roles or different levels of

management than assessed against as not everyone that participated necessarily went into a managerial or a higher-level management role.

Assessment Centres

The evaluators used within the development centres were qualified Psychologists, Psychometrists and trained role-players. A 5-point scale was used and it comprised of the following:

Table 3.5.

Assessment Centre Scales

Scale	Meaning
1	Development Required
2	Coaching Required
3	Competent
4	Exceeds Requirements
5	Outstanding

University of Cape Town

Occupational Personality Questionnaire (OPQ)

This instrument has been used within the assessment centre to establish and individual's behaviour at work, by measuring or assessing personality. It comprises of a 32 – point scale, and was developed by obtaining feedback from organisations, input from employees as well as by using personality theories (Barrett, Kline, Paltiel & Eysenck, 1996). The OPQ was first developed in the United Kingdom and is presently being used in 40 countries and is available in 27 languages (SHL, 1999). However, although it was not particularly adapted for South Africa, it is nonetheless used.

Workplace behaviour is assessed along three dimensions and these include Relationships with People, Thinking style and Feelings and Emotions. Another dimension that focuses on establishing the degree of competitiveness and achievement orientation cuts across the three dimensions (Barrett, Kline, Paltiel & Eysenck, 1996). Table 3.6 illustrates the scales used within the OPQ.

Table 3.6.

Scales used within the OPQ

OPQ Dimensions		
Relationship with People	Thinking style	Feelings and Emotions
Persuasive	Data rational	Relaxed
Controlling	Evaluative	Worrying
Outspoken	Behavioural	Tough minded
Independent minded	Conventional	Optimistic
Outgoing	Conceptual	Trusting
Affiliative	Innovative	Emotionally controlled
Socially confident	Variety seeking	Vigorous
Modest	Adaptable	Competitive
Democratic	Forward thinking	Achieving
Caring	Detail conscious	Decisive
	Conscientious	Consistency
	Rule following	

Respondents are required to assess themselves on 104 questions by indicating which statement, from a set of four statements, is the most/least descriptive of their behaviour (Saville & Holdsworth, 1999).

University of Cape Town

Reliability of the OPQ

Using a standardization sample size of 807 respondents, the reliability of the OPQ was established (Saville & Holdsworth, 1999). Two thirds of the sample size comprised of employees within seven different organisations, the rest of the respondents were undergraduate students. The table 3.7 shows the internal consistency estimates within the OPQ. The alpha co-efficients were found to range between 0.67 and 0.88. Democratic and Evaluative were the only scales that fell below 0.70.

Validity of the OPQ

Saville & Holdsworth (2004) conducted a study within the insurance sector to establish the predictive validity of the OPQ. Participants consisted of middle and senior managers, and no significant correlations were found. Table 3.7 also shows the correlations between the OPQ and the performance ratings.

Table 3.7

Internal Consistency of scores and predictive validity of the OPQ

OPQ Dimensions								
Relationship with People	r	α	Thinking styles	r	α	Feelings & Emotions	r	α
Persuasive	0.11	0.81	Data rational	-0.04	0.88	Relaxed	0.03	0.85
Controlling	0.12	0.87	Evaluative	0.18	0.67	Worrying	-0.03	0.88
Outspoken	0.04	0.76	Behavioural	0.10	0.82	Tough-minded	0.05	0.82
Independent-minded	-0.05	0.72	Conventional	-0.14	0.74	Optimistic	-0.04	0.80
Outgoing	-0.02	0.85	Conceptual	0.08	0.79	Trusting	-0.03	0.81
Affiliative	0.01	0.82	Innovative	0.08	0.88	Emotionally-controlled	-0.03	0.85
Socially confident	0.04	0.83	Variety seeking	0.04	0.72	Vigorous	-0.17	0.75
Modest	-0.01	0.81	Adaptable	0.09	0.82	Competitive	0.12	0.86
Democratic	0.06	0.68	Forward thinking	0.09	0.75	Achieving	0.12	0.79
Caring	-0.12	0.78	Detail conscious	-0.14	0.80	Decisive	0.01	0.80
			Conscientious	-0.16	0.82			
			Rule following	-0.37	0.84			

Note. N = 807

University of Cape Town

Quantitative Data

The data that was collected was numerical and it was statistically analysed to derive conclusions. The two sets of data were analysed separately in order to establish predictive and construct validity.

Reliability and Validity of Study

Reliability is concerned with the extent to which one's study is replicable (Collis & Hussey, 2003). The methods used to retrieve and analyse data within this study, have been clearly stated to enable the replicability of the study. Consequently, this study is open to scrutiny and inspection and comparable results should be obtained should another similar research that is compatible with the theory being used, be conducted (Welman & Kruger, 2002). In addition, validity or the extent to which this study establishes the validity of development centres has been upheld by ensuring that the procedures used are an accurate measure of development centre validity. Previous research studies (Kavanagh et al., 1971) have shown that the methods used establish development centre validity within this study as well as the relevance of rating scale content, are accurate measures for assessing validity.

University of Cape Town

Ethical considerations

Ethical clearance was obtained from the commerce faculty board in order to carry out this research. The issues that were addressed pertained to how anonymity and confidentiality of the participants as well as the organisation would be ensured. In addition the ethical issues pertaining to scientific research were addressed

4. RESULTS

Predictive validity

Performance Ratings

As a result of having a considerable amount of missing data within the performance scores, one performance rating was used and this was obtained by getting an average rating for each employee from all the available performance ratings scores. Illustrated below is the mean and standard deviation amongst the performance ratings.

Table 4.1

Mean and standard deviation within the Performance Ratings

	Mean	Std Dev.	N
PR 01-Jul-01	2.94	0.56	89
PR 01-Jul-02	3.10	0.57	103
PR 01-Jul-03	3.21	0.62	114
PR 31-Dec-03	3.26	0.58	90
PR 31-Dec-04	3.26	0.51	114
Average Mean	3.15		

Note. Case-wise deletion of missing data was used. Total number of employees within this data-set was 120. N = number of participants in the performance evaluation

University of Cape Town

Reliability of Performance Ratings

In addition to using correlation coefficients to establish validity, they can also be used to ascertain the reliability of a measuring device. Consequently, the reliability coefficient within the performance scores comprised of the correlation between the different scores of individuals within the performance scores. The main aim of estimating reliability in the performance scores is to establish how much variability in the ratings is due to errors in measurement and how much is due to variability in true scores (Murphy & Davidshofer,

1991). Using the test-retest method, it will be possible to determine whether an individual's scores on two different measures correlate. Differences between both performance measures would be attributed to measurement error. Therefore if the difference is large then one would conclude that the measurement errors are a major source of variability (Murphy & Davidshofer, 1991). On the other hand, small differences would indicate a highly reliable method.

The table below illustrates the reliability coefficients between the five performance scores. There are four significant scores and the correlation coefficients range from 0.39 to 0.44. These are much lower than desirable as reliability coefficients usually range between 0.80 and 0.90 (Anastasi & Urbina, 1997). This means that there is wide variation between these scores and this variance could be attributed to measurement error. The most important aspect to note within this table is that the correlation coefficients are only significant within the ratings that were taken in a shorter time interval of each other. In other words, each performance rating was significantly correlated with the next rating that was taken and these ratings were not significantly related to the other performance scores taken at a much later or earlier time.

University of Cape Town

Table 4.2.

Correlation coefficients for the various performance ratings

	01-Jul-01	01-Jul-02	01-Jul-03	31-Dec-03
01-Jul-02	.44**			
01-Jul-03	.23	.52**		
31-Dec-03	.13	.13	.44**	
31-Dec-04	.02	.15	.31	.39*

*$p < 0.05$

**$p < 0.02$

Development centre Ratings

An overall assessment rating that had been obtained by getting an average competency score of an individual within the various methods, was used within this study to establish predictive validity. It should be noted that four sub-groupings were made within this development centre data as not all employees had been assessed on similar competencies. These sub-groupings comprised of individuals who had been assessed on the same traits. However, one sub-group consisted of only 3 employees and therefore was left out. The sub-groupings enabled comparisons to be made within the data. The correlation coefficients between the development centre ratings and performance ratings were computed using case-wise deletion of missing data. Case-wise deletion of missing data keeps out all cases that have missing data in at least one of the selected variables (Statsoft, 2003). On the other hand, however, the mean and standard deviation within this data-set was computed with a mean substitution of missing data. Table 4.3 shows the mean and standard deviation within this data-set.

University of Cape Town

Table 4.3.

Mean and standard deviation for the first data-set with mean substitution of missing data

Competency	Mean	Std.Dev.
Thinking Skills	2.42	1.07
Business Awareness	1.83	1.07
Aligning Performance for Success	2.37	0.45
Commitment	2.42	0.66
Team Orientation	2.79	0.33
Client Focus	2.60	0.68
Ownership Assuming	2.71	0.37
Decision Making	2.58	0.81
Initiating Action	2.43	0.85
Planning & Organising	2.69	0.76
Continuous Learning	2.07	0.94
Building Relationships	2.57	0.61
Facilitating Change	2.35	0.20
Energy	2.86	0.56
Enterprise Innovation	2.35	0.37
Building a Successful Team	2.24	0.52
Developing Others	2.22	0.49

Note. n = 120

University of Cape Town

Internal consistency of the Development centre scores

An analysis of the internal consistency was carried out on 17 competencies (3 competencies with the least values were eliminated) in order to establish

whether the items within the different measures of competence were reliable in a sense that the different measures assess a coherent construct.

Table 4.4.

Internal consistency amongst the competency scores

Competency	Item -Total Correlation	Squared Multiple R	Alpha (α) if deleted
Thinking Skills	0.49	0.42	0.68
Business Awareness	-0.04	0.31	0.76
Aligning Performance for Success	0.35	0.38	0.70
Gaining Commitment	0.22	0.28	0.71
Team Orientation	0.18	0.24	0.71
Client Focus	0.46	0.46	0.69
Ownership Assuming	0.31	0.30	0.71
Decision Making	0.65	0.52	0.66
Initiating Action	0.52	0.46	0.67
Planning & Organising	0.50	0.40	0.68
Continuous Learning	0.10	0.22	0.73
Building Partnerships	0.32	0.35	0.70
Facilitating Change	0.22	0.42	0.71
Energy	0.41	0.28	0.69
Enterprise Innovation	0.17	0.28	0.71
Building a Successful Team	0.29	0.49	0.70
Developing Others	0.38	0.58	0.70

Note. n = 120; Cronbach Alpha = 0.71; and Average inter-item correlation = 0.15

The table above illustrates the internal consistency amongst the items. This was derived by substituting means for all the missing data. It shows the correlations between one competency and the total competency score, the squared multiple correlation between the respective competency and all other

University of Cape Town

competencies and lastly it illustrates the coefficient alpha or the internal consistency of the scale if α is deleted. It can be concluded that all the scores are internally consistent as there is minimal variance within the inter-item correlation. The Cronbach Alpha of 0.71 is quite high indicating that the scores have a relatively high reliability.

Analyses

In order to establish predictive validity, two analytical measures as explained below were used.

Correlation matrices

These were used to establish the relationship between the performance ratings and the development centre ratings.

Multiple regressions

In order to determine the strength of the linear relationship between the assessment centre ratings and performance ratings, the co-efficient of determination, also known as R^2, was used. The p-value in the multiple regression results was also analysed because it indicates the probability of support for the alternative hypothesis vs. the null-hypothesis (Keller & Warrack, 2003). The smaller the p-value, the more statistical evidence exists to support the alternative hypothesis.

University of Cape Town

Group 1

The first analysis for establishing predictive validity was carried out on 42 employees who had been assessed on 10 competencies. Figure 4.1 illustrates the race and gender distribution within this group.

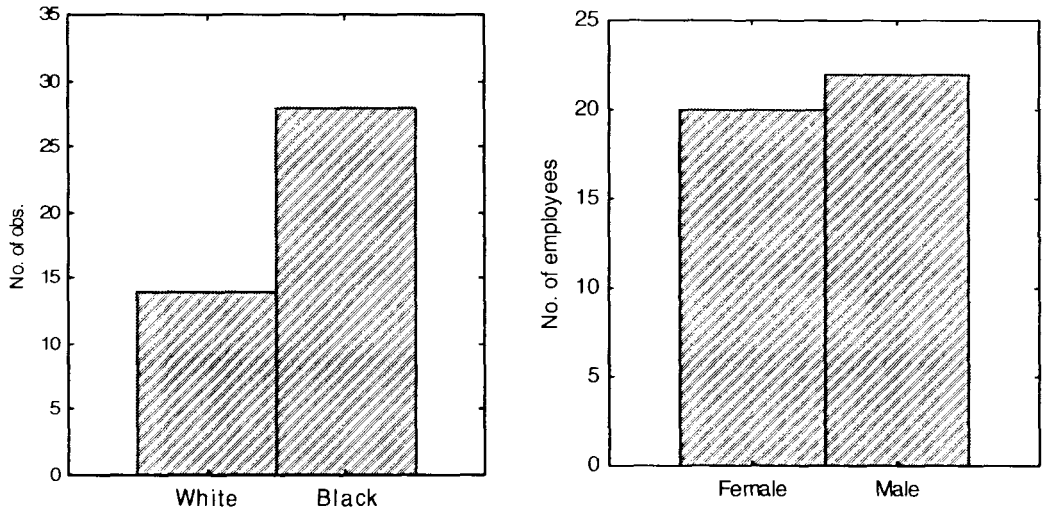


Figure 4.1. Race and gender distribution within Group 1

Table 4.5 shows the correlation matrix between the competencies and the performance rating for the employees.

Table 4.5.

Correlation Matrix for Group 1

Competency	Correlation between Performance Rating and Competency
Thinking Skills	-0.03 (p=0.86)
Business Awareness	0.03 (p=0.86)
Aligning Performance for Success	0.13 (p=0.40)
Commitment	-0.01 (p=0.96)
Team Orientation	0.01 (p=0.92)
Client Focus	0.09 (p=0.55)
Ownership Assuming	0.08 (p=0.63)
Decision Making	-0.12 (p=0.45)
Initiating Action	0.20 (p=0.21)
Planning & Organising	0.43* (p=.0005)

Note. $n = 42$ for each computed correlation

*$p < 0.05$

University of Cape Town

The only significant correlation between the two variables is 0.43 for the Planning and Organising competency (p=0.005). However, this is a relatively low correlation with the performance rating. This also means that there is a low linear relationship between the competency rating and the performance rating. A multiple regression analysis (co-efficient of determination) was also carried out on the same group to establish how much variation within the performance ratings was as a result of the variation in the competency ratings, this is illustrated in table 4.6.

Table 4.6.

Regression Summary for Group 1

Competency	t(31)	*p*
Thinking Skills	-0.28	0.78
Business Awareness	0.11	0.91
Aligning Performance for Success	-0.30	0.76
Commitment	0.76	0.45
Team Orientation	0.83	0.41
Client Focus	1.78	0.08
Ownership Assuming	-0.74	0.47
Decision Making	-2.83	0.01*
Initiating Action	0.87	0.39
Planning & Organising	3.15	0.004*

Note. $R = 0.66$; $R^2 = 0.43$; *Adjusted* $R^2 = 0.25$; $F_{10,31} = 2.36$;

*$p < 0.05$

Table 4.6 indicates that the co-efficient of determination is 0.43. This means that 43% of the variation in the performance ratings is explained by the variation in the competency ratings. A further analysis indicates that Planning and Organising as well as Decision Making contribute significantly to the performance rating. In other words, 31% of the variation in the performance

University of Cape Town

ratings is explained by the variation within these two competencies. Table 4.7 shows the regression summary for these two competencies. The F value (2.36) also indicates that the model is significant ($p<0.05$).

Table 4.7.

Regression Summary for Planning and Organising and Decision Making

	t(39)	*p*
Decision Making	-2.69	0.011*
Planning & Organising	4.10	0.0002*

Note. $R = 0.56$; $R^2 = 0.31$; *Adjusted* $R^2 = 0.25$; $F_{10,31} = 2.36$

*$p < 0.05$

Group 2

The second group comprised of 17 employees that were assessed on 10 competencies. Figure 4.2 shows the race and gender distribution within this group.

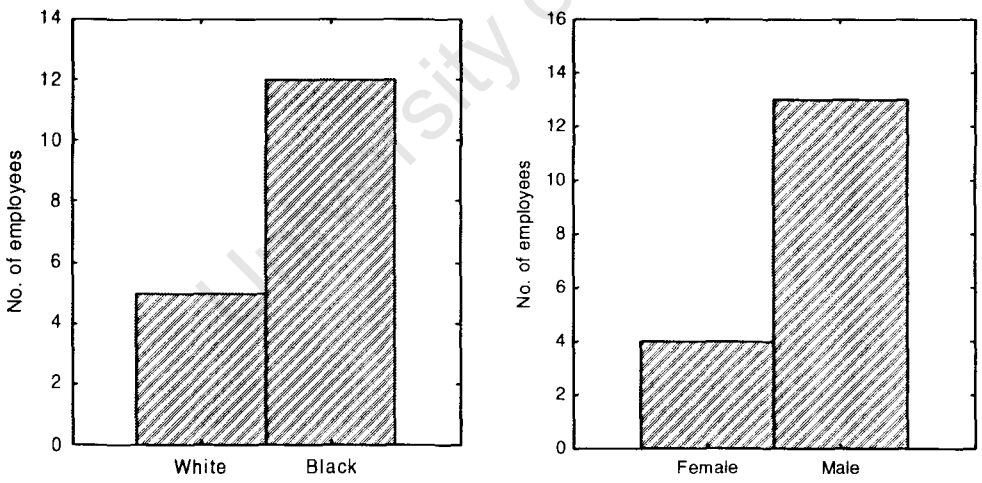


Figure 4.2. Race and gender distribution within Group 2

A correlation analysis indicated that Business Awareness and Decision Making competencies had a significant relationship with the performance rating. An illustration of this is given in table 4.8.

University of Cape Town

Table 4.8.

Correlation Matrix for Group 2

Competency	Correlation between Performance Rating and Competency
Thinking Skills	0.18 (p=0.50)
Business Awareness	0.5411* (p=0.025)
Aligning Performance for Success	0.17 (p=0.51)
Commitment	0.17 (p=0.51)
Client Focus	0.09 (p=0.17)
Ownership Assuming	-0.15 (p=0.57)
Decision Making	0.5218* (p=0.032)
Initiating Action	0.03 (p=0.90)
Facilitating Change	0.41 (p=0.10)
Enterprise Innovation	0.30 (p=0.24)

Note. $n = 7$ for each computed correlation

*$p < 0.05$

University of Cape Town

The correlation co-efficient for Business Awareness (r = 0.54) and Decision Making (r = 0.52) is relatively high which means that there is a moderate linear relationship between these two competencies and the performance rating. A regression analysis indicated that there were no significant relationships between the variables and the performance rating (table 4.9). However the co-efficient of determination of 0.6 (table 4.9) showed that 60% of the variance within the performance rating could be attributed to the variance in the competency ratings. The overall F value of 9.12 also indicates that the model is significant.

Table 4.9.

Regression Summary for Group 2

	t(6)	*p*
Thinking Skills	-0.43	0.68
Business Awareness	0.08	0.94
Aligning Performance for Success	0.94	0.38
Commitment	-1.20	0.28
Client Focus	0.72	0.50
Ownership Assuming	-1.54	0.17
Decision Making	0.77	0.47
Initiating Action	0.52	0.62
Facilitating Change	-0.12	0.91
Enterprise Innovation	0.49	0.64

Note. $R = 0.78$; $R^2 = 0.60$; *Adjusted* $R^2 = -0.05$; $F_{10,6} = 9.12$

On the contrary however, an individual analysis of all the competencies showed that Business Awareness and Decision Making had a significant relationship with the performance rating as shown in tables 4.10 and 4.11. In addition, 30% of the variation within the performance ratings could be explained by the variation within these competency ratings. The overall F (6.2 and 5.6 respectively) also indicates that both models are significant ($p<0.05$).

Table 4.10.

Regression Summary for Business Awareness

	t(15)	*p*
Business Awareness	2.5	0.025*

Note. $R = 0.54$; $R^2 = 0.29$; *Adjusted* $R^2 = 0.25$; $F_{1,15} = 6.2$

*$p < 0.05$

University of Cape Town

Table 4.11.

Regression Summary for Decision Making

	t(15)	p-level
Decision Making	2.34	0.03*

Note. $R = 0.52$; $R^2 = 0.27$; Adjusted $R^2 = 0.22$; $F_{1,15} = 5.6$

*$p < 0.05$

Group 3

The last analysis for establishing predictive validity was carried out on a group of 58 employees. The race and gender distribution within this group is illustrated in figure 4.3. There were 12 competencies that were assessed within this group and an average performance rating for each individual was used. The correlation matrix between the competency ratings and the performance ratings in table 4.12 below shows that there are no significant relationships. In addition, the correlation co-efficients between the competency ratings and performance ratings are all very low. It is also important to note that Business Awareness, Commitment, Initiating Action and Energy are negatively correlated with the performance rating.

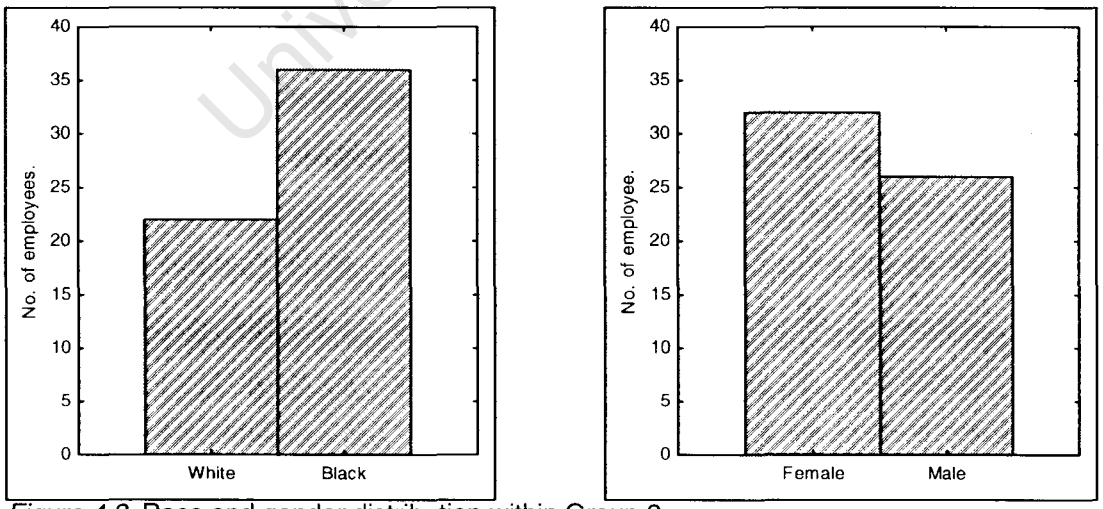


Figure 4.3. Race and gender distribution within Group 3

University of Cape Town

Table 4.12.

Correlation Matrix for Group 3

Competency	Correlation between Performance Rating and Competency
Thinking Skills	0.03 (p=0.90)
Business Awareness	-0.88 (p=0.51)
Commitment	-0.18 (p=0.18)
Client Focus	0.02 (p=0.87)
Decision Making	0.01 (p=0.94)
Initiating Action	-0.03 (p=0.83)
Planning & Organising	0.23 (p=0.08)
Continuous Learning	0.11 (p=0.42)
Building Relationships	0.16 (p=0.24)
Energy	-0.12 (p=0.34)
Building a Successful Team	0.81 (p=0.55)
Developing Others	0.24 (p=0.07)

Note. $n = 58$ for each computed correlation

A regression analysis also indicates that only 24% (table 4.13) of the variance within the performance ratings could be explained by the variance within the competency ratings. In addition, the Planning and Organising competency has a significant relationship with the performance rating. The F value (1.18) also indicates that the model is significant ($p<0.05$).

University of Cape Town

Table 4.13.

Multiple Regression Summary for Group 3

Competency	t(45)	p
Thinking Skills	-0.02	0.99
Business Awareness	0.99	0.33
Commitment	-0.83	0.41
Client Focus	-0.38	0.70
Decision Making	-0.79	0.43
Initiating Action	-0.37	0.71
Planning & Organising	2.14	0.04*
Continuous Learning	1.30	0.20
Building Relationships	0.58	0.57
Energy	-1.30	0.20
Building a Successful Team	-0.29	0.77
Developing Others	1.52	0.13

Note. $R = 0.49$; $R^2 = 0.24$.; *Adjusted* $R^2 = 0.04$; $F_{12,45} = 1.18$;

*$p < 0.05$

University of Cape Town

As a result of having smaller groups, an analysis on the whole data-set was carried out in order to establish whether there would be any findings that were different from the ones above. However, in order to compute the data, all missing variables were substituted with a mean. In addition, the competencies with the most missing data were removed and these were adaptability, risk taking and stress tolerance. The correlation analysis did not yield any significant values. However, the multiple regression summary had three competencies with significant correlations and these were Business Awareness (r = 0.02), Initiating Action (r = 0.04) and Planning and Organising (r = 0.001). Table 4.14 below illustrates the multiple regression summary with a substitution of means for the missing data.

Table 4.14.

Multiple Regression Summary with mean substitution of missing Data

Competency	t(102)	p-level
Thinking Skills	-0.26	0.79
Business Awareness	2.42	0.02*
Aligning Performance for Success	0.04	0.97
Commitment	-0.71	0.48
Team Orientation	-0.44	0.66
Client Focus	0.62	0.54
Ownership Assuming	-0.39	0.69
Decision Making	-0.68	0.50
Initiating Action	-2.06	0.04*
Planning & Organising	3.45	0.001*
Continuous Learning	1.17	0.25
Building Relationships	-0.87	0.39
Facilitating Change	1.11	0.27
Energy	-0.88	0.38
Enterprise Innovation	0.54	0.59
Building a Successful Team	-0.81	0.42
Developing Others	1.73	0.09

Note. $R^2 = 0.2$

*$p < 0.05$

University of Cape Town

Construct Validity

The second data-set was used to establish construct validity. Case-wise deletion of missing data yielded 'no valid cases' within the second data-set and consequently no analyses could be computed as the missing data was randomly distributed across cases. As a result, correlations were computed using pair-wise deletion of missing data. This method derives correlations for all cases that have valid data on the two variables being analysed. It is important to note that the number of employees for each analysis differs as pair-wise deletion of missing data was used to compute the values.

Construct validation enables one to establish whether an assessment measure actually measures a specific construct (Murphy & Davidshofer, 1991). In order to establish construct validity, two statistical techniques have been used. These are Multitrait-multimethod approach and Analysis of Variance (ANOVA). The underlying relationship amongst the variables could not be analysed using factor analysis due to the limited sample size.

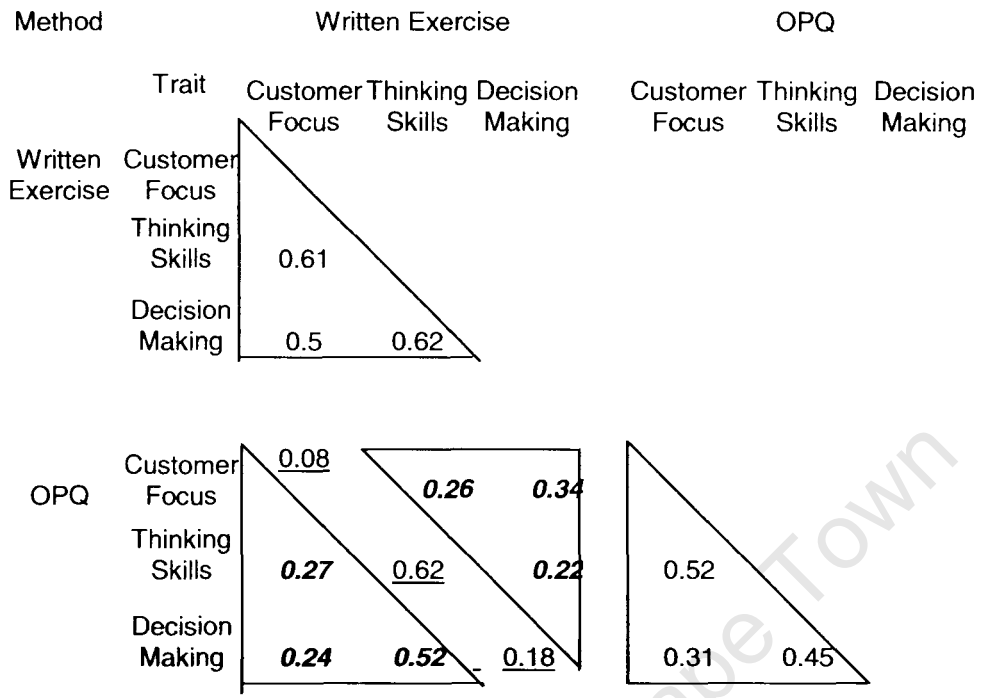
Multitrait-multimethod approach

Within this approach, the correlations amongst various assessment methods and traits being measured, take the form of a Multitrait-multimethod matrix as illustrated in table 4.15 (Murphy & Davidshofer, 1991). Campbell and Fiske's (1959) Multitrait-multimethod approach has become one of the widely used methods for establishing construct validity (Kleinmann & Koller, 1997). This approach provides a wide array of information as it enables the comparison of correlations amongst the different methods and traits (Murphy & Davidshofer; Lawler, 1967). Table 4.15 provides an example illustrating Multitrait-multimethod correlation matrix. The correlation co-efficients within this table have been based on some of the data within this study.

University of Cape Town

Table 4.15.

Example of a Multitrait-multimethod matrix



Note. The underlined coefficients depict convergent validity. The bold and italicised coefficients illustrate Discriminant Validity using the Heterotrait-heteromethod approach. Likewise, the values in the top most triangle illustrate Discriminant validity using the Heterotrait-monomethod approach.

University of Cape Town

There are three separate tables that have been used to illustrate the inter-correlation matrices, within this study. This is as a result of having various competencies being assessed using different methods and therefore one table could not be used to illustrate the Multitrait-multimethod correlation. Table 1 (see appendix A) is a 39 x 39 matrix and shows all the correlations between the variables. Within this table, the shaded cells illustrate the discriminant validity coefficients using the Heterotrait-heteromethod approach. The convergent validity coefficients have been underlined and the rest of the coefficients illustrate discriminant validity using the Heterotrait-monomethod approach.

For easier analysis however, the Heteromethod-monotrait (convergent validity) and Heterotrait-monomethod (discriminant validity) matrices have

been illustrated in 2 separate tables as shown in Table 4.16 and Table 4.17, respectively. Using Campbell & Fiske's guideline for evaluating Multitrait-multimethod matrices, the three ways in which these correlations were analysed to establish convergent and discriminant validity are described below.

Convergent Validity

Heteromethod-monotrait approach.

This approach has been used to establish what is referred to as convergent validity. Using this mode of analysis, the ratings for a single construct within various methods should be high and should converge to yield similar results. An illustration has been given in table 4.15 above and the convergent validity co-efficients are the underlined values. According to Campbell and Fiske (1959), convergent validity exists when correlations amongst various methods measuring a similar construct are significantly different from zero. It is important to note that this approach only considers correlations of a single trait and various assessment measures hence the term monotrait-multimethod (Murphy & Davidshofer, 1991). Table 4.16 below shows the Heteromethod-monotrait (convergent validity) coefficients within this study.

University of Cape Town

Table 4.16.

Inter-correlations between various methods for a single trait (Heteromethod–monotrait approach)

Methods	In-basket	Business Case Presentation	Written Exercise	OPQ	Role Play	Career Interview	Group Exercise
			Customer Focus				
Business Case Presentation	0.21 (n=32)						
Written Exercise	--	0.33 (n=4)					
OPQ	**0.29* (n=51)**	0.17 (n=38)	0.08 (n=24)				
Role Play	0.19 (n=23)	0.20 (n=25)	**0.56* (n=22)**	**0.43* (n=48)**			
			Commitment				
Group Exercise						0.12 (n=23)	
Role Play						0.29 (n=21)	0.17 (n=43)
OPQ					**0.43* (n=48)**	0.23 (n=25)	**0.26* (n=63)**
			Aligning Performance for success				
OPQ	**0.30* (n=58)**						
Role Play	0.02 (n=14)			**0.39* (n=36)**			
			Facilitating Change				
Group Exercise						0.29 (n=34)	
Business Case Presentation						--	--
OPQ		**0.58** (n=25)**				--	**0.32* (n=39)**
Role Play		**0.57** (n=31)**		**0.61** (n=29)**		0.45 (n=7)	0.19 (n=7)

University of Cape Town

Table 4.16 (continued)

Methods	In-basket	Business Case Presentation	Written Exercise	OPQ	Role Play	Career Interview	Group Exercise
			Thinking Skills				
Group Exercise	**0.43** (n=47)**						
Business Case Presentation	**0.49** (n=39)**						**0.41** (n=34)**
Written Exercise	0.87 (n=3)	**0.62** (n=24)**					0.58 (n=4)
OPQ	0.19 (n=38)	0.19 (n=54)	**0.62** (n=21)**				0.26 (n=32)
Role Play	0.06 (n=26)	0.10 (n=25)	--	0.17 (n=16)			0.04 (n=25)
			Decision Making				
Group Exercise	**0.53** (n=44)**						
Business Case Presentation	**0.42** (n=43)**						**0.39* (n=29)**
Written Exercise	0.50 (n=3)	0.69 (n=5)					**0.54* (n=20)**
OPQ	0.27 (n=54)	**0.49** (n=44)**	0.18 (n=21)				0.17 (n=57)
Role Play	**0.47** (n=29)**	0.17 (n=29)	**0.83** (n=20)**	0.26 (n=46)			**0.40** (n=44)**
			Initiating Action				
Business case Presentation	0.30 (n=39)						
Written Exercise	--	**0.68** (n=19)**					
OPQ	0.21 (n=54)	0.22 (n=58)	0.38 (n=17)				
			Planning & Organising				
OPQ	**0.39** (n=51)**						
			Building Relationships				
OPQ						0.15 (n=45)	

University of Cape Town

Table 4.16 (continued)

Methods	In-basket	Business Case Presentation	Written Exercise	OPQ	Role Play	Career Interview	Group Exercise
			Business Awareness				
Group Exercise	**0.80** (n=15)**						
Business case Presentation	**0.63** (n=25)**						**0.64** (n=19)**
OPQ	0.30 (n=20)	**0.46* (n=23)**					0.20 (n=18)

Note. The significant values are in bold font and unavailable data is indicated by --.

*$p < 0.05$

**$p < 0.01$

Table 4.16 above shows that the inter-correlations for Customer Focus, Building Relationships, Commitment, Aligning Performance for Success, and Planning and Organising were moderately low with the highest correlation being 0.56 for Customer Focus (Role Play and Written Exercise) and the lowest 0.26 for Commitment (OPQ and Group Exercise). The other competencies i.e. Thinking Skills, Decision Making, Initiating Action, Facilitating Change and Business Awareness had some moderately high correlations, with the highest correlation being 0.8 for Decision Making (Role Play and Written Exercise) and Business Awareness (Group Exercise and In-basket) and the lowest 0.32 for Facilitating Change (OPQ and Group Exercise). These correlations are also illustrated in the 39 x 39 matrix within the appendix as the underlined values.

University of Cape Town

Discriminant validity

Heterotrait-monomethod approach.

This approach was used to establish discriminant validity by analysing correlations amongst different traits, within a single assessment method (Murphy & Davidshofer, 1991). The correlations between measures of different constructs should be small or lower than correlations between measures of the same trait. According to Murphy and Davidshofer, constructs are chosen to be clearly different, therefore measures of the constructs should not correlate highly. However, sometimes these correlations are larger than the correlations between various methods and various traits (heterotrait-heteromethod) and this indicates what is referred to as method bias (Murphy & Davidshofer). In other words, a small proportion of the correlation derived from this approach is attributed to the common method of measurement i.e. different traits are measured using the same method.

The discriminant validity co-efficients obtained using the Heterotrait-monomethod approach have been illustrated in table 4.17 below. Within this table, the statistically significant correlations within Role Play and OPQ methods were quite low, with the highest value being 0.70 and the lowest 0.07.

On the other hand however, there were some relatively high correlations within the Business Case Presentation exercise, Written Exercise, Role Play, Group Exercise and In-Basket. The highest correlation within these methods was 0.8 (Decision making and thinking skills as measured by the In-basket). These values are also illustrated in table 1 (see appendix A) within the unshaded sections.

University of Cape Town

Table 4.17.

Inter-correlations between traits within a single method (Heterotrait-monomethod approach)

Competencies	Customer focus	Building Relationships	Commitment	Aligning Performance for success	Facilitating Change	Thinking skills	Decision making	Initiating Action	Planning & Organising
				OPQ					
Building Relationships	**0.52** (n=62)**								
Commitment	**0.35** (n=77)**	**0.44** (n=60)**							
Aligning Performance for success	**0.29** (n=79)**	0.23 (n=62)	**0.37** (n=78)**						
Facilitating Change	0.12 (n=77)	0.23 (n=61)	**0.27* (n=75)**	**0.32** (n=77)**					
Thinking skills	**0.52** (n=62)**	0.23 (n=45)	0.19 (n=63)	0.16 (n=64)	0.24				
Decision making	**0.31** (n=75)**	**0.34** (n=57)**	**0.37** (n=73)**	**0.35** (n=75)**	**0.26* (n=73)**	**0.45** (n=60)**			
Initiating Action	**0.34** (n=75)**	**0.50** (n=59)**	**0.38** (n=73)**	**0.37** (n=75)**	**0.35** (n= 74)**	0.22 (n=59)	**0.40** (n=72)**		
Planning & Organising	**0.26* (n=58)**	**0.36* (n=56)**	**0.40** (n=56)**	**0.42** (n=57)**	0.25 (n=57)	0.21 (n=44)	**0.31* (n=55)**	**0.49** (n=56)**	
Business Awareness	0.07 (n=24)	0.19 (n=7)	0.14 (n=25)	0.22 (n=25)	0.25 (n=23)	0.40 (n=24)	0.32 (n=24)	0.20 (n=23)	0.11 (n=7)

University of Cape Town

Table 4.17 (continued)

Competencies	Customer focus	Building Relationships	Commitment	Aligning Performance for success	Facilitating Change	Thinking skills	Decision making	Initiating Action	Planning & Organising
				Business Case Presentation					
Facilitating Change	0.47 (n=15)								
Thinking skills	0.10 (n=24)				**0.63** (n=34)**				
Decision making	0.19 (n=19)				**0.51* (n=17)**	**0.66** (n=52)**			
Initiating Action	**0.31* (n=43)**				**0.50** (n=35)**	**0.37** (n=65)**	**0.41** (n=51)**		
Business Awareness	0.49 (n=10)				**0.47* (n=27)**	0.21 (n=28)	0.49 (n=10)	0.33 (n=28)	
				Written Exercise					
Thinking skills	**0.62** (n=21)**								
Decision making	**0.50* (n=21)**					**0.62** (n=22)**			
Initiating Action	**0.52* (n=18)**					**0.55* (n=19)**	0.38 (n=18)		

University of Cape Town

Table 4.17 (continued)

Competencies	Customer focus	Building Relationships	Commitment	Aligning Performance for success	Facilitating Change	Thinking skills	Decision making	Initiating Action	Planning & Organising
				Role Play					
Commitment	**0.29* (n=57)**								
Aligning Performance for success	**0.49** (n=41)**		**0.52** (n=43)**						
Facilitating Change	**0.57** (n=40)**		**0.50** (n=39)**	**0.70** (n=37)**					
Thinking skills	0.07 (n=33)		**0.44** (n=34)**	0.22 (n=20)	0.26 (n=15)				
Decision making	**0.42** (n=55)**		**0.48** (n=56)**	**0.33* (n=42)**	**0.47** (n=38)**	0.26 (n=35)			
				Group Exercise					
Facilitating Change			**0.67** (n=36)**						
Thinking skills			**0.54** (n=45)**		**0.58** (n=39)**				
Decision making			**0.44** (n=58)**		**0.58** (n=36)**	**0.72** (n=44)**			

University of Cape Town

Table 4.17 (continued)

Competencies	Customer focus	Building Relationships	Commitment	Aligning Performance for success	Facilitating Change	Thinking skills	Decision making	Initiating Action	Planning & Organising
				In-Basket					
Aligning Performance for Success	**0.69** (n=50)**								
Thinking skills	**0.49** (n=48)**			**0.65** (n=55)**					
Decision making	**0.59** (n=51)**			**0.69** (n=58)**		**0.80** (n=56)**			
Initiating Action	**0.47* (n=47)**			**0.73** (n=55)**		**0.60** (n=52)**	**0.72** (n=55)**		
Planning & Organising	**0.50** (n=47)**			**0.79** (n=54)**		**0.55** (n=51)**	**0.67** (n=54)**	**0.67** (n=52)**	

Note. The significant values have been highlighted in bold font

* $p < 0.05$

** $p < 0.01$

University of Cape Town

Heterotrait-heteromethod approach

This approach was used to establish discriminant validity by analysing the correlations between various methods and traits. These correlations should be small and lower than the correlations within different traits as measured by a single method (heterotrait-monomethod) (Murphy & Davidshofer, 1991).

The values in the shaded cells within Table 1 (see appendix) illustrate the correlations between various methods and traits. These values should be lower than the correlations within the Heterotrait-monomethod approach mentioned above. The highest correlation within this table was 0.84 for the Customer Focus and Thinking Skills competencies as measured by the Role play and Written Exercise respectively. The other high scores were 0.82 for the Facilitating Change trait as measured by the Role plays and Building Relationships as measured by the Career Interview. There was yet another high score of 0.83 for the Building Relationships (as measured by the OPQ) and Facilitating Change (as measured using Role Plays). However most of the correlations within this table were relatively low.

Extent of Variance between Methods and Traits

Analysis of Variance (ANOVA)

Although the Multitrait-multimethod approach enables conclusions about construct validity to be made, it is difficult to make relative judgments using this approach as data derived from this approach is usually difficult to interpret (Kavanagh, Mackinney & Wolins, 1971; Murphy & Davidshofer, 1991; Schmitt, Coyle & Saari, 1977). Kleinmann and Koller (1997) state that construct validity can be better estimated and improved upon by applying statistical analyses on the Multitrait-multimethod matrices. Consequently, an analysis of variance has been used within this study to make meaningful interpretations of the correlations.

University of Cape Town

Considering Employees as random and traits and methods as fixed, the three-way classification model below has been hypothesised to describe the correlation data (Kavanagh et al., 1971). The interpretation of this model is based on the degree of variation within the Multitrait-multimethod matrix.

$$Y_{IJK} + \mu + \alpha_i + \beta_j + \gamma_\kappa + (\alpha\beta)_{ij} + (\alpha\gamma)_{i\kappa} + (\beta\gamma)_{j\kappa} + \epsilon ij\kappa$$

Where: Y_{IJK} = Rating of employee$_I$ on a trait$_J$ by method$_K$

α_i = effect of employee $_i$

β_j = effect of trait $_j$

γ_κ = effect of method $_\kappa$

$\epsilon ij\kappa \sim$ ND $(0, \sigma_\epsilon^2)$ i.e. error term $\epsilon ij\kappa$ is normally distributed with a mean of 0 and variance σ_ϵ^2

With this model, it will be possible to establish:

i. Employee variance which indicates convergent validity;

ii. Employee by trait variance which indicates discriminant validity

iii. Employee by method variance which indicates the amount of Source bias (halo effect) (Kavanagh et al., 1971).

The computations for the sums of squares of the above effects from the matrix, the degrees of freedom (df) and the expected mean squares (MS) (Kavanagh et al., 1971) have been calculated as shown in the table 4.18.

University of Cape Town

Table 4.18.

Computation of Sums of Squares, Degrees of Freedom and Expected Means.

Source	df	SS	MS
Employee (E)	N-1	Nnm ($\check{r}_0$)	$\sigma_\varepsilon^2 + nm\,\sigma_m^2$
E x Trait (T)	(N-1) (n-1)	Nnm ($\check{r}_{bt}$ - $\check{r}_0$)	$\sigma_\varepsilon^2 + m\,\sigma_{m\ x\ T}^2$
E x Method (M)	(N-1) (m-1)	Nnm ($\check{r}_{bm}$ - $\check{r}_0$)	$\sigma_\varepsilon^2 + n\,\sigma_{m\ x\ s}^2$
Error	(N-1) (n-1) (m-1)	Nnm (1 – $\check{r}_{bt}$ – $\check{r}_{bs}$ + $\check{r}_0$)	σ_ε^2

Note. N = total number of employees; n = number of traits; m = number of methods; $\check{r}_0$ = average correlation of all variables in the multitrait-multimethod matrix; $\check{r}_{bt}$ = average correlation between traits; $\check{r}_{bm}$ = average correlation between methods; σ_ε^2 = variance between employees; σ_m^2 = variance between methods.

The correlations within this study were calculated using pair-wise deletion of missing data therefore, the total number of employees (N) could not be easily established between each group. As a result, in order obtain computations with matching pairs of Ns, N was obtained by getting the total number of employees in each trait under analysis. An average was then obtained by dividing this number with the number of traits under analysis.

It should be noted that the analysis of variance will only be computed for only four sets of data listed below as these have the most correlation values. In addition, in the case of inverse correlations between traits, absolute values were used to compute the Mean of Squares (MS).

- Between the Business Presentation and OPQ method
- Between Role plays and OPQ method
- Between In-basket and OPQ method
- Between Performance assessment measures and personality measures. There is only one personality measure, which is the OPQ. The performance measures include all the methods with which behaviour is assessed i.e.

University of Cape Town

Business Case Presentation, Career Interview, Written Exercise, Role Play, Group Exercise and In-basket.

•

Analysis of Variance between the Business Presentation method and OPQ

The analysis of variance of correlations amongst employees in table 4.19 below gave an F value of 3.13. This F value at 5% significance level ($p > 0.05$), is greater than 1.32 and consequently, it is not significant. Therefore, it can be concluded that there is no differentiation amongst the employees that is attributable to the methods used. This also means that there is no evidence for convergent and discriminant validity. Likewise, the employee x trait interaction, as well as the employee x method interactions are not significant ($p > 0.05$).

Table 4.19.

Analysis of Variance (ANOVA) of Correlations between Business Case Presentation and OPQ

Source	df	MS	F
Employee	58	5.22	3.13
Employee x Trait	290	0.28	0.17
Employee x Method	58	0.04	0.02
Error	290	1.67	

Note. Number of employees = 59; number of traits = 6; number of methods = 2;

$\check{r}_{bm} = 0.43$; $\check{r}_{bt} = 0.31$; $\check{r}_{0} = 0.43$

Analysis of variance between performance measures and personality measures.

The F value of 4.13 in table 4.20 below was lower than 1.32 ($p>0.05$). Therefore the F value was not significant. It can be concluded therefore that there is no

University of Cape Town

differentiation amongst the employees. Likewise, an F value of 0.29 was lower than 1 and consequently not significant. As the result, the h_0 was also rejected. Since this value is the weakest, it can be also be concluded that there is less discriminant validity than convergent validity.

Table 4.20.

Analysis of Variance (ANOVA) of Correlations between Performance Measures and Personality Measures

Source	df	MS	F
Employee	52	4.95	4.13
Employee x Trait	468	0.35	0.29
Employee x Method	52	1.47	1.23
Error	468	1.19	

Note. Number of employees = 53; number of traits = 10; number of methods = 2;

$\check{r}_{bm} = 0.31$; $\check{r}_{bt} = 0.39$; $\check{r}_{0} = 0.24$

Analysis of Variance between Role Plays/In-baskets and OPQ.

Similarly the analyses in table 4.21 and 4.22 did not yield any significant F values. Therefore, there was no differentiation between the employees based on the method used. In addition, the F values for the effect between employee x trait was the smallest within both analyses which indicated that there was less discriminant validity than convergent validity.

University of Cape Town

Table 4.21.

Analysis of Variance (ANOVA) of Correlations between Role Plays and OPQ

Source	df	MS	F
Employee	62	5.58	3.17
Employee x Trait	310	0.26	0.15
Employee x Method	62	0.89	0.50
Error	310	1.76	

Note. Number of employees = 63; number of traits = 6; number of methods = 2;

$\check{r}_{bm}$ = 0.38; $\check{r}_{bt}$ = 0.35; $\check{r}_{0}$ = 0.46

Within the above analyses, it was concluded that there was no evidence for method bias as the F values within the employee x method effect were small and not significant.

Table 4.22.

Analysis of Variance (ANOVA) of Correlations between In-Baskets and OPQ

Source	df	MS	F
Employee	63	7.02	3.55
Employee x Trait	315	0.21	0.10
Employee x Method	63	3.67	1.86
Error	315	1.97	

Note. Number of employees = 64; number of traits = 6; number of methods = 2;

$\check{r}_{bm}$ = 0.27; $\check{r}_{bt}$ = 0.49; $\check{r}_{0}$ = 0.57

Relevance of Rating Scale Content

Using the above analyses it was possible to establish the relevance of the rating scale content. This was attained by analysing the variance between the performance measures and personality measures in order to establish whether one

University of Cape Town

method was better than the other at assessing performance. In addition, in order to ascertain the relevance of each individual assessment technique in effectively assessing a competency, a correlation analysis was carried out by comparing the various individual scores for each method with the final score.

The analysis of variance between Performance measures and Personality measures in table 4.20 above shows that there is little or no variation between the methods as an F-value of 1.23 is lower than an alpha of 1.32 (5% significance level) which means that both methods yield similar results. Consequently, it cannot be concluded that either method is better in assessing employee performance.

Correlation of Dimension ratings within the Overall Assessment Rating (OAR) for each competency.

This analysis was done in order to establish whether some methods contributed more to the OAR than others. Pair-wise deletion of missing data was used to compute the data within this section.

Table 4.23.

Correlation of ratings for Customer Focus

Customer Focus	
Measured by	Correlation of Final Score on competency
In-basket	0.57** (n=51)
Business Case Presentation	0.56** (n=46)
Written Exercise	0.52** (n=26)
OPQ	0.42** (n=80)
Role Play	0.58** (n=59)

*$p < 0.05$

**$p < 0.01$

University of Cape Town

Table 4.23 shows that there was a moderately high correlation between the individual scores and the final score. However, Role Play and In-Baskets had the highest correlation with the final score. All the values within this table were highly significant ($p < 0.01$).

Table 4.24.

Correlation of ratings for Building Relationships

Building Relationships	
Measured by	Correlation of Final Score on competency
Career Interview	0.33* (n=45)
OPQ	0.57** (n=62)

*$p < 0.05$

**$p < 0.01$

Within the above table, the ratings for OPQ had a moderately high correlation (0.57) with the final score.

Table 4.25.

Correlation of ratings for Commitment

Commitment	
Measured by	Correlation of Final Score on competency
Career Interview	0.33 (n=27)
Group Exercise	0.46** (n=68)
OPQ	0.49** (n=79)
Role Play	0.60** (n=59)

*$p < 0.05$

**$p < 0.01$

University of Cape Town

Table 4.25 shows that Role Play had scores with the highest correlation with the final score (0.60). In addition, it should be noted that the Career Interview had no significant relationship with the final score.

Table 4.26.

Correlation of ratings for Aligning Performance for Success

Aligning Performance for Success	
Measured by	Correlation of Final Score on competency
In-basket	0.73**(n=57)
OPQ	0.54** (n=78)
Role Play	0.63** (n=43)

*$p < 0.05$

**$p < 0.01$

Table 4.26 above shows that the Role Play (0.63) and In-Basket (0.73) had ratings that had a highly correlated and significant relationship with the final score.

Table 4.27.

Correlation of ratings for Facilitating Change

Facilitating Change	
Measured by	Correlation of Final Score on competency
Career Interview	0.25 (n=41)
Group Exercise	0.47**(n=40)
Business Case Presentation	0.63**(n=36)
OPQ	0.61**(n=77)
Role Play	0.69**(n=40)

*$p < 0.05$

**$p < 0.01$

University of Cape Town

Within table 4.27, the scores for Career Interview were not significantly related to the final score. On the other hand, OPQ (0.61), Role Play (0.69) and Business Case Presentation (0.63) had scores that were moderately correlated with the final score.

Table 4.28.

Correlation of ratings for Thinking Skills

Thinking Skills	
Measured by	Correlation of Final Score on competency
In-basket	0.60**(n=54)
Group Exercise	0.51**(n=49)
Business Case Presentation	0.60**(n=69)
Written Exercise	0.82**(n=23)
OPQ	0.45**(n=63)
Role Play	0.37*(n=35)

*$p < 0.05$

**$p < 0.01$

Table 4.28 above shows that the scores for Written Exercise had a very high correlation with the final rating (0.82).

University of Cape Town

Table 4.29.

Correlation of ratings for Decision Making

Decision Making	
Measured by	Correlation of Final Score on competency
In-basket	0.63**(n=57)
Group Exercise	0.72**(n=60)
Business Case Presentation	0.54**(n=53)
Written Exercise	0.86**(n=22)
OPQ	0.43**(n=74)
Role Play	0.70**(n=57)

*$p < 0.05$

**$p < 0.01$

The most significant and highly correlated values within the previous table were Written Exercise (0.86), Group Exercise (0.72), Role Play (0.70) and In-Basket (0.63).

Table 4.30.

Correlation of ratings for Initiating Action

Initiating Action	
Measured by	Correlation of Final Score on competency
In-basket	0.46**(n=56)
Business Case Presentation	0.53**(n=68)
Written Exercise	0.91**(n=18)
OPQ	0.56**(n=75)

*$p < 0.05$

**$p < 0.01$

University of Cape Town

Table 4.30 also shows that the Written Exercise had a very high correlation with the final score (0.91). The rest of the scores were moderately correlated with the final score.

Table 4.31.

Correlation of ratings for Planning and Organising

Planning and Organising	
Measured by	Correlation of Final Score on competency
In-basket	0.72**(n=53)
OPQ	0.57**(n=57)

*$p < 0.05$

**$p < 0.01$

There were two methods used to assess the Planning and Organising competency and the In-Basket had the highest correlation with the final score.

Table 4.32.

Correlation of ratings for Business Awareness

Business Awareness	
Measured by	Correlation of Final Score on competency
Business Case Presentation	0.69**(n=30)
OPQ	0.53**(n=25)

*p=<0.05

**p=<0.01

The Business Case Presentation had the highest correlation with the final score (0.69). Nonetheless, OPQ had a moderately high and significant correlation the final score (0.53).

University of Cape Town

5. DISCUSSION

Criterion validity

The aim of this study was to establish the predictive and construct validity of a development centre. With reference to previous studies conducted to establish the validity of assessment centres, it was hypothesised that development centres have high predictive validity. In order to establish the predictive validity of the development centre, competency scores of employees were compared to subsequent performance ratings. Previous research has found that assessment centres have high predictive validity as these centres effectively determine successful job performance (Schleicher et al., 2002). In other words, the higher the competency scores, the higher the on-the-job performance of individuals. However, some studies (Gaugler et al., 1987; Klimoski & Strickland, 1977) found that assessment centres are more predictive of promotion suitability of individual than job performance.

The present study found a low correlation between most of the competency ratings and performance ratings. The highest significant correlation values were r = 0.54 for Business Awareness and r = 0.52 for Decision Making within Group 3. These competencies also accounted for 30% of the variance (R^2 = 0.3) within the performance rating. Furthermore, the correlation between Planning and Organising and the performance rating within Group 1 was moderately high (r = 0.43). The regression analysis within Group 1 also showed that the Planning and organising as well as Decision Making (R^2 = 0.31) accounted for 31% of the variation in the performance ratings. On the other hand however, the rest of the competencies had low correlations with the performance rating, which were also not significant.

In a study conducted by Klimoski and Strickland (1977) it was established that assessment centre ratings had a positive correlation with subsequent job performance. However, this positive correlation could have been as a result of other factors that might have influenced the individuals' performance ratings (Strickland, 1977). For example subsequent developmental decisions could have

University of Cape Town

been biased by the assessment centre results and as a result the performance ratings were not an accurate reflection of effective job performance. That is to say, individuals could have been rewarded for good ratings obtained from the assessment centre and not necessarily for their good performance and this in turn could have impacted on the performance score given to the individual. Klimoski and Brickner (1987) called this criterion contamination, which means that the rating on one measure influences the rating on another.

It should be noted that the reliability of performance ratings used within this study were relatively low as the correlations ranged between 0.39 and 0.52. These low correlations could have been attributed to the error variance, which corresponds to the "random fluctuations of performance" from one measurement to another (Anastasi & Urbina, 1997, p. 92). This means that the performance scores are more susceptible to random changes within the work environment. There are various intervening conditions that could be attributed to the wide variance within performance ratings of this study. These include different positions or jobs within which these scores were taken. In other words, the employees could have held different positions, which could have caused variance between the scores. As a result, the criterion measures used within the different performance evaluations could have been different.

In addition, these individuals could have been scored by different evaluators, which could have also contributed to the measurement variance within the performance ratings. The test-retest reliability method used also shows the extent to which these scores could be generalised from one situation to another. The higher the reliability, the more generalisable the results (Anastasi & Urbina, 1997). The moderate reliability indicates that these scores are relatively generalisable. Anastasi and Urbina also maintain that the reliability of measurement devices is usually restricted to short-term time intervals as retest correlations progressively decrease over time. Evidently, within this study, the performance ratings were taken within a long-term time interval (within 6 to 12 months of each other) and

University of Cape Town

therefore, the generalisability of the results is limited. The correlations between these performance ratings were also not significant.

In yet another study conducted by Huck and Bray (1976), it was established that criterion contamination did not influence the supervisors' ratings for their employees. This study found that there was no difference in the ratings between those in which the supervisors knew the assessment centre ratings of their employees and those that did not know that ratings. Thornton and Byham (1982) go on to mention that there might be criterion contamination initially as decisions could be based on the assessment centre results. However, in the long-term, they argue that the assessment centre results do not influence further decisions. Therefore criterion contamination might not have influenced these results as the performance ratings used were taken a number of months after the assessment centre process.

Jones, Herriot, Long and Drakley (1991) point out that the low correlation between competency ratings and performance ratings could be attributed to assessors' inability to effectively combine information from individual assessment instruments. This could possibly be as a result of having a number of dimensions from which an Overall Assessment Rating (OAR) has to be derived and therefore can lead to the loss of information from each component instrument. In addition, assessors might attach more weighting to less valid predictors, which in turn would affect the correlation between the OAR and performance scores (Lance et al., 2000).

In a study conducted by Bycio et al. (1987) it was found that there was low predictive validity. They pointed out that this could have been as a result of the exercises' misrepresentation of job requirements. Exercise specific ratings should correlate with performance ratings to the extent that the exercises accurately represent the required competencies.

Another reason for the low predictive validity within this study could be attributed to the rather small sample that was used. Gaugler et al. (1987) state that sampling

University of Cape Town

error accounts for a proportion of the variability within validity co-efficients. It was inevitable to use sub-groupings with smaller sample sizes within this study as opposed to one large sample, as not all participants had been assessed on similar traits. Consequently, combining the various groups could have led to a large number of missing data. Nevertheless, an analysis was still done on the whole data set by substituting missing means. This was done in order to establish if different results from those within the sub-groupings could have been obtained. As stated above, this integrated analysis still yielded a low co-efficient of determination (R^2 = 0.2) between the competency ratings and performance ratings.

Yet another reason for low predictive validity within assessment centres could be attributed to the moderate/high levels of range restriction of assessment results (Gaugler et al., 1987; Thornton & Byham, 1982). Not every individual that is assessed advances within their job and consequently the behavioural traits that are assessed within the assessment centre may not be the same as those measured by the performance ratings. Subsequently, the sample gets restricted as a relatively small subgroup gets evaluated. Notably, within this study, some of the performance data was based on specialist roles or different levels of management than assessed against within the assessment center, as not everyone that participated necessarily went into a higher level management role. This was a possible explanation for the low predictive validity within as not all participants within the assessment centre went on to advance to the positions for which they were assessed. Consequently, the performance ratings were based on different criterion than those with which the employees had been assessed on within the assessment centre.

Low predictive validity within assessment centres could also be attributed to the complexity of having to assess individuals' effectiveness on a higher level job where competencies are usually different from those required within the current job (Thornton & Byham, 1982). Similarly, Lance et al. (2000) state that an assessor's misjudgement of critical behaviour could lead to low predictive validity as the behaviours being assessed could be different from those needed to perform

University of Cape Town

successfully. Therefore, this would lead to the performance ratings not being similar to the assessment centre ratings. Lance et al. go on to point out that assessors may consider assessment centre exercises as work sample tests in which success is determined by the final performance on the test. On the other hand, successful performance on assessment measures should be established by observing the ongoing desirable behaviours within the various exercises and not only by assessing the resultant performance on the entire exercise. In other words, the dimension ratings should also be taken into consideration. Consequently, assessors need to focus on all the various traits that lead to success within the entire exercise if they are to effectively assess similar competencies as those assessed on the job.

Furthermore, Thornton and Byham (1982) mention that assessment centres usually lack representative validation samples as participants are usually those recommended by their supervisors as having potential to perform successfully in a higher position. Consequently, the sample is usually biased and not representative and therefore could lead to a high correlation between the performance ratings and assessment centre ratings. On the other hand, however, when self-nomination is used to select participants, the overall assessment centre ratings are usually lower. Thornton and Byham go on to state that the few self-nominees who get high ratings within the assessment centre go on to perform more successfully when they advance to higher positions than those who get chosen by their supervisors.

It should be noted that one other factor that should be considered whilst establishing validity of assessment centres is the reliability of the rating scales. The reliability of the criteria with which the assessment centre scores are validated against can significantly vary and this in turn yields invalid results. Therefore, this can result in low correlations between the performance ratings and the assessment centre ratings. Thornton and Byham (1982) state that low reliability of performance ratings could be as a result of halo effects and leniency of supervisors. Using the test re-test reliability analysis, the reliability co-efficients ranged between 0.39 and 0.52, which was moderately low. However, it should be noted that these low

University of Cape Town

correlations could have been as a result of different criterion being evaluated as a result of development or promotion of an employee. On the other hand however, the internal consistency analysis of the competency scores yielded a considerably high cronbach alpha of 0.71.

Bycio et al., (1987) note that the Multitrait-multimethod correlation co-efficients of development centres should be positive as this indicates that the assessment centre ratings and the job requirements are related. Consequently, effective development decisions can be made from the results of the analysis as this shows that the criteria being measured is relevant. Most of the correlation co-efficients within the Multitrait-multimethod matrix were positively correlated and it can therefore be concluded that the constructs being measured were relevant and had a positive relationship with the job.

Construct validity

Previous studies had shown that assessment centres have low construct validity and as a result, the second and third hypotheses within this study stated that assessment centres have low convergent validity and high discriminant validity. Most of the convergent validity correlations (Heteromethod-monotrait approach) were below 0.7 and therefore moderately low, with the lowest being 0.26 for Commitment. However, there were two high correlations of 0.8 for Decision Making and Business Awareness competency. On the contrary, the Heterotrait-monomethod yielded some relatively high correlations with the lowest correlation being 0.31 and the highest 0.79. Most of the correlations ranged between 0.4 and 0.7. The Heterotrait-heteromethod approach (see shaded correlations in table 1 within the Appendix) also yielded some relatively high correlations with the highest correlation being 1.0 for the Thinking Skills and Facilitating Change competencies as measured by Written Exercise and Group Exercise respectively (see page 102 and 105). These correlations indicated that there is no evidence for discriminant validity. Therefore this study found that there was relatively low construct validity.

University of Cape Town

Lievens (2002) found that low construct validity could be attributed to the assessment centre design such as inexperienced assessors, having a number of dimensions within each exercise as well as not using behavioural checklists. In addition, inter-rater reliability was found to impact on construct validity as well. In other words, the variance within exercise ratings along a dimension could be attributed to the variance within the assessors. Lievens states that inter-rater variance within a development centre can be determined by having all the assessors rating all the individuals. In consistency within these ratings would mean that there is variance amongst the assessors' ratings. Consequently, a well-designed assessment centre would yield high construct validity. This could explain some of the high construct validity ratings within the development centre. The assessors used within the development centre were well trained and comprised of Psychologists , Psychometrists and trained Role Players.

Neidig and Neidig (1984) also mention that minimal or a lack of construct validity should not preclude the job-relatedness of an assessment centre as there are various factors that could affect the validity of the constructs (Neidig & Neidig, 1984). Furthermore, Lance et al. (2000), state that the situational specificity of exercises used in assessment centres is a cause of variance within performance and this inadvertently has an effect on construct validity. This means that assessees demonstrate different behaviour within various exercises, which leads to inconsistency within the observed traits. As a result, assessors rate them differently on similar traits within various exercises as different behaviour is exhibited. Bycio et al. (1987) conducted a study in which the cross-situational consistency and discriminant validity of assessment centres was to be established. Their results showed that assessment centres measure situation specific competencies. Therefore an improvement of the validity could be attained by measuring more behaviour related traits and reducing the "cognitive demand" placed on the raters so that they can observe all behaviours (Bycio et al., 1987, p. 473). Lievens also mentions that in order to establish construct validity, the level of consistency of individual performance should be verified. Therefore, it could be

University of Cape Town

concluded that the relatively high construct validity within this study could be attributed to the assessor competence.

An analysis of variance indicates the amount of variance within the methods used (Zedeck, 1986). Within an assessment, assessors should not look out for different behavioural traits along a dimension as this would lead to low convergent validity. The analyses of variance within this study indicated that there was no evidence for method bias as the F values within the employee x method effect were small and not significant. It should be noted that having similar ratings within an exercise can also be attributed to halo effect (Robie, Osburn, Morris, Etchegaray & Adams, 2000;). This in turn limits the cross situational consistency of ratings across exercises, thus leading to low convergent validity and high correlations within an exercise amongst the different trait ratings (discriminant validity) (Zedeck).

In addition, as a result of the significant results that were not significant, it cannot be established whether there was a greater effect due to the method (method bias) than to the consistency of assessee behaviour and the assessor ratings. Nedig and Neidig (1984) point out that the main objective of assessment centres should be to observe required on-the-job behaviour using an array of measures. Consequently, variance within the ratings should be expected due to the different methods used and this variance should not be attributed to measurement error. In addition, in order for a method to have job relatedness, accurate representation of the job requirements is necessary.

It should also be noted that the analysis of variance indicated that there was no difference between the personality measures and performance measures. In a study conducted by Rothstein and Johnston (1996) to ascertain the relative validity of personality measures and assessment centres, it was also found that although personality measures and assessment centres differed in the way they measured behaviour, they were both equally valid measures in predicting performance.

University of Cape Town

In a further study, Silverman et al. (1986) wanted to ascertain whether consistency amongst ratings within various exercises could be attained. This was done by having the assessors rate each dimension across exercises before coming up with an overall rating for the entire dimension. Results showed that this did not improve the cross-situational consistency of the ratings. Based on this study it was unclear whether the inconsistency was as a result of differing behaviour exhibited by assessees, inaccurate rating of behaviour by assessors, or variations in the behaviour that assessees have to display as a result of dissimilar exercises.

Furthermore, the findings within this study show that some methods contribute more to the OAR than others. Consequently, it can be suggested that fewer methods (those with the highest correlation) should be used in the assessment process. For example, the Career Interview should not have been used to assess the Building Relationships and Facilitating Change competencies as it did not have a significant relationship with the final rating (OAR). On the other hand, the methods that could have been used to assess Thinking Skills are Written Exercise ($r = 0.82$), In-Basket ($r = 0.60$), Business Case Presentation ($r = 0.60$) and Group Exercise ($r = 0.51$). With regards to the Decision Making Competency, Written Exercise ($r = 0.86$), Group Exercise ($r = 0.72$), Role Play ($r = 0.70$) and In-Basket ($r = 0.63$) should have been used as these were highly correlated with the overall assessment rating. Written Exercise also had the highest correlation when used to measure Initiating Action skills ($r = 0.91$).

It should be noted that Written Exercise scores had the highest correlation ($r = 0.52$ for assessing Customer Focus; $r = 0.82$ for assessing Thinking Skills; $r = 0.86$ for assessin Decision Making; and $r = 0.91$ for assessing Initiating Action) with the final scores within this study. This means that they make a substantial contribution to the OAR. With regards to the Career Interview, Thornton and Byham (1982) stated that the interviewing technique within assessment centres had been proven to be valid. On the contrary, the analyses within this study showed that there was a low correlation between the Career Interview and the OAR. Therefore, considering that the Interviewing technique is a valid measure, these findings could be attributed to

University of Cape Town

the stereotypes held by the interviewer and unstandardised information that leads to inter-rater disagreement (Thornton & Byham). Another reason for the low correlation could be attributed to the fact that interviews provide descriptions of past behaviour while other assessment methods observe behaviour related to performance (Thornton & Byham).

In another study conducted by Borman (1982) it was found that the structured interview was less valid than the simulation exercises. McDaniel, Whetzel, Schmidt and Maurer (1994) found that the validity of the interview technique was dependent on its content, the way it was conducted and the nature of the criterion. It was also established that situational interviews were more valid than job-related and psychological interviews. In addition, structured interviews were better at assessing the required traits than unstructured interviews. However, further studies should be conducted to establish the validity of the interview technique as a method used within the assessment process.

Limitations and Recommendations

Limitations

One limitation to this study was in relation to the sample size. In order to make generalisations from a study, the sample size should be large enough and representative of the entire population (Babbie & Mouton, 2001). The sample used within this study was not large enough as data could only be retrieved from one financial institution within the Western Cape. Lance et al. (2000) point out that as a result of the time and financial cost involved in the assessment process, most studies have relatively small samples. The limited use of assessment centres impacted on the sample size within this study, as assessment centre data could not be easily obtained. Consequently, generalisations of the findings within this study were done with caution.

In addition, Kleinmann and Koller (1997) state that correlations derived from the Multitrait-multimethod approach are usually based on observable variables that

University of Cape Town

could potentially contain measurement errors. Therefore conclusions based on the correlations within the Multitrait-multimethod matrix could be disputed if the reliability of the scales is unclear. However it should be noted that within this study, the internal consistency of the scales was computed and therefore it was established that the scales used were reliable.

Similarly, the constructs used within this study are latent and therefore it is important to establish the definition of the constructs under investigation so as to ensure that the assessment methods used actually measure what has been set out to be measured. Welman and Kruger (2002) state that the correlation co-efficient could be affected by extraneous variables such as the meaning of the constructs. Within this study, the operational definition of constructs was not obtained therefore the content validity of the constructs could not determined.

Yet another limitation within this study pertained to the availability of similar studies with which plausible comparisons could be made. Gaugler et al. (1987) mention that there are no standardised assessment centre procedures and therefore variability of the assessment methods may cause certain variables to moderate assessment centre validity. For example, the level of assessor training and inter-rater reliability differ amongst studies and this could impact on the comparisons being made.

University of Cape Town

Recommendations

Although most studies found that assessment centres have predictive validity, lack of predictive validity could be attributed to various factors and can be improved upon. One way of ensuring that the assessment process has predictive validity is by reducing the number of dimensions or behaviours to observe for each trait (Jones et al., 1991). Consequently, the instruments used should measure broader attributes or constructs as this would enable assessors to easily identify relevant traits within exercises. In addition, Jones et al. found that validities of assessment centres were considerably lower than validities of individual assessment measures as individual assessments are usually constructed to measure specific ability. On

the other hand, assessment centres usually measure a wide range of competencies. Therefore, the dimensions used within an assessment centre should relate to specific criterion outcomes and the weighting for each trait should depend on an instruments' predictive validity. This means that competencies should be weighted relatively, according to the predictive validity of the instrument used.

This study could not establish the validity of assessment centres within a multicultural context like South Africa because the data-sets were not large enough to enable these analyses. However, although Thornton and Byham (1982) stated that assessment centres are equally valid predictors of job performance within all designated groups, it would still be desirable to establish whether these findings are generalisable to South Africa. In order to verify the fairness of assessment instruments, it is important to establish the content validity of a measuring instrument. In addition, predictive and construct validity should be ascertained separately for the various groups being assessed.

Lastly, Epstein (1983) points out that behaviour is influenced by the context and consequently is situation specific. It should therefore not be expected to find consistent behaviour being exhibited within different contexts as each situational is unique and therefore disguises the consistency of behaviour. As a result, in order to obtain findings with consistent behaviour, further studies should cancel out situational effects.

University of Cape Town

6. CONCLUSION

This study set out to establish the validity of assessment centres within a financial institution. Data from a financial institution within the Western Cape was retrieved and the predictive and construct validity of the assessment centre was investigated.

In relation to previous studies, this study did find some evidence for predictive and construct validity. More specifically this research study found that there was a relatively low correlation between the assessment centre ratings and performance management ratings (predictive validity), there was a relatively low correlation of construct ratings across a single dimension (convergent validity) and a relatively high correlation between ratings of different traits within a single exercise. The low predictive validity could have been attributed to a number of factors such as the low reliability of the performance scores, range restriction and a rather small sample used within this study. On the other hand, the relatively low construct validity within this study could have been attributed to the assessment centre design and the situational specificity of assessee behaviour.

It is recommended that further research should be carried out to establish the validity of assessment centres within a multicultural context like South Africa. In addition, the situational effect should be considered and cancelled when analysing the validity of assessment centres so that the impact of cross situational variance on validity is minimized.

University of Cape Town

References

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). New Jersey: Prentice Hall International.

Arthur, W., Day, E., McNelly, T., & Edens, P. (2003). A meta-analysis of the criterion-related validity of assessment center dimensions. *Personnel Psychology, 56,* 125-154.

Arthur, W., Woehr, D., & Maldegen, R. (2000). Convergent and discriminant validity of assessment center dimensions: A conceptual and empirical re-examination of the assessment center construct-related validity paradox. *Journal of Management, 26*, 813-835

Babbie, E., & Mouton, J. (2001). *The practice of social research.* Cape Town: Oxford University Press.

Barrett, P., Kline, P., Paltiel L., & Eysenck, H. (1996). An evaluation of the psychometric properties of the concept 5.2 Occupational Personality Questionnaire. *Journal of Occupational and Organizational Psychology, 69*, 1-19

Boam, R., & Sparrow, P. (Eds.). (1992). *Designing and achieving competency. A competency-based approach to developing people and organizations.* London: McGraw-Hill

Borman, W. (1982). Validity of behavioral assessment for predicting military recruiter performance. *Journal of Applied Psychology, 67*, 3-9

Boyatzis, R. (1982). *The competent manager: A model for effective performance.* New York: Wiley

Briscoe, D. (1997). Assessment centers: Cross-cultural and cross national issues. *Journal of social behavior and personality, 12,* 261-271

Bycio, P., Alvares, K., & Hahn, J. (1987). Situational specificity in assessment centre ratings: A confirmatory factor analysis. *Journal of applied psychology, 72*, 463-474

Byham, W. (1971). The assessment center as an aid in management development. *Training and Development Journal, 25*, 10-22

Campbell, D., & Fiske, D. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105

University of Cape Town

Collis, J., & Hussey, R. (2003). *Business research: a practical guide for undergraduate and postgraduate students* (2nd ed.). New York: Palgrave Macmillan

Edenborough, R. (2005). *Assessment methods in recruitment, selection and performance: A manager's guide to psychometric testing, interviews and assessment centers*. London: Kogan Page.

Epstein, S. (1983). The stability of confusion: A reply to Mischel and Peake. *Psychological Review, 90*, 179-184

Foxcroft, C., & Roodt, G. (Eds.). (2003). *An introduction to psychological assessment in the South African context*. Cape Town: Oxford University Press.

Gaugler, B., Rosenthal, D., Thornton, G., & Bentson, C. (1987). Meta analysis of assessment center validity. *Journal of Applied Psychology, 72*, 493-511

Ghiselli, E., & Brown, C. (1955). *Personnel and industrial psychology*. New York: McGraw-Hill.

Goffin, R., Rothstein, M., & Johnston, N. (1996). Personality testing and the assessment centre: Incremental validity for managerial selection. *Journal of Applied Psychology, 81*, 746-756.

Greyling, L., Visser D., & Fourie, L. (2003). Construct validity of competency dimensions in a team leader assessment centre. *South African Journal of Industrial Psychology*, *29*, 10-19

Haaland, S., & Christiansen, N. (2002). Implications of trait-activation theory for evaluating the construct validity of assessment center ratings. *Personnel Psychology, 55*, 137-163

Hoeft S., & Schuler, H. (2001). The conceptual basis of assessment centre ratings. *International Journal of Selection and Assessment, 9*, 114-123

Hornby, D., & Thomas, R. (1989). Towards a better standard of management. *Personnel Management, 21,* 52-5

Huck, J., & Bray, D. (1976). Management assessment center evaluations and subsequent job performance of Black and White females. *Personnel Psychology, 29*, 13-30

Jones, A., Herriot, P., Long, B., & Drakley, R. (1991). Attempting to improve the validity of a well-established assessment centre. *Journal of Occupational Psychology, 64*, 1-21.

University of Cape Town

Kavanagh, M., MacKinney, A., & Wolins, L. (1971). Issues in managerial performance: Multitrait-multimethod analyses of ratings. *Psychological Bulletin*, 75, 34-49.

Keller, G., & Warrack, B. (2003). *Statistics for management and economics* (6th ed.). Pacific Grove: Thomson Learning Inc.

Kerlinger, F., & Lee, H. (2000). *Foundations of behavioral research* (4th ed.). Texas: Harcourt College Publishers.

Kleinmann, M., & Koller, O. (1997). Construct validity of assessment centres: appropriate use of confirmatory factor analysis and suitable construction principles. *Journal of Social Behaviour and Personality, 12*, 65-84.

Klimoski, R., & Brickner, M. (1987). Why do assessment centers work? The puzzle of assessment center validity. *Personnel Psychology, 40*, 243-262

Klimoski, R., & Strickland, W. (1977). Assessment centres: Valid or merely prescient. *Personnel Psychology, 30*, 353-361

Kuptsch, C., Kleinmann, M., & Koller, O. (1998). The chameleon effect in assessment centers: The influence of cross-situational behavioral consistency on the convergent validity of assessment centers. *Journal of Social Behavior and Personality, 13*, 103-117

Lance, C., Newbolt, W., Gatewood, R., Foster, M., French, N., & Smith, D. (2000). Assessment center exercise factors represent cross-situational specificity, not method bias. *Human Performance, 13*, 323-353

Lawler, E. (1967). The multitrait-multirater approach to measuring managerial job performance. *Journal of Applied Psychology, 51*, 369-381.

Lievens, F. (2002). Trying to understand the different pieces of the construct validity puzzle of assessment centres: An examination of assessor and assessee effects. *Journal of Applied Psychology, 87*, 675-686

Markus, H., Thomas, H., & Allpress, N. (2005). Confounded by competencies? An evaluation of the evolution and use of competency models. *New Zealand Journal of Psychology, 34*, 117-126

McDaniel, M., Whetzel, D., Schmidt, F., & Maurer, S. (1994). The validity of employment interviews. *Journal of Applied Psychology, 79*, 599-616.

Murphy, K., & Davidshofer, C. (1991). *Psychological testing: Principles and applications.* New Jersey: Prentice-Hall, Inc.

University of Cape Town

Neidig, R., & Neidig, P. (1984). Multiple assessment center exercises and job relatedness. *Journal of Applied Psychology, 69*, 182-186

Olshfski, D., & Cunningham, R. (1986). Establishing assessment centre validity: An examination of methodological and theoretical issues. *Public Personnel Management, 15*, 85-97

Robie, C., Osburn, H., Morris, M., Etchegaray, J., & Adams, K. (2000). Effects of the rating process on the construct validity of assessment center dimension evaluations. *Human Performance, 13*, 355-370

Sackett, P., & Dreher, G. (1982). Constructs and assessment center dimensions: Some troubling empirical findings. *Journal of Applied Psychology, 67*, 401-410

Sackett, P., & Dreher, G. (1984). Situation specificity of behaviour and assessment center validation strategies: A rejoinder to Neidig and Neidig. *Journal of Applied Psychology, 69*, 187-190

Saunders, E. (2000). *Assessing human competence. Practical guidelines for South African managers.* Randburg: Knowres Pub.

Saunders, E. (2002). *Assessing human competence. Practical guidelines for the South African manager.* Randburg: Knowres Pub.

Sauders, M., Lewis, P., & Thornhill A. (2003). *Research methods for business students* (3rd ed.). Harlow: Prentice Hall

Schleicher, D., Day, D., Mayes, B., & Riggio, R. (2002). A new frame for Frame of Reference training: Enhancing the construct validity of assessment centers. *Journal of applied Psychology, 87*, 735-746

Schmitt, N., Coyle, B., & Saari, B. (1977). A review and critique of analyses of multitrait-multilethod matrices. *Multivariate Behavioral Research, 12*, 447-478.

Saville & Holdworth Ltd. (1999). *OPQ32: Manual and user's guide*. Surrey: SHL Group Plc

Saville & Holdworth Ltd. (2004). *The Occupational Personality Questionnaire in South Africa : psychometric research perspectives*. Surrey: SHL Group Plc

Shore, T., Thornton, G., & Shore, L. (1990). Construct validity of two categories of assessment center dimension ratings. *Personnel Psychology, 43*, 101-116

University of Cape Town

Silverman, W., Dalessio, A., Woods, S., & Johnson, R. (1986). Influence of assessment center methods on assessors' ratings. *Personnel Psychology, 50,* 71-90

Spangenberg, H., & Theron, C. (2003). Validation of the high performance leadership competencies as measured by an assessment measure in-basket. *South African Journal of Industrial Psychology, 29*, 29-38

Spencer, M., McClelland, C. & Spencer, M. (1992). *Competency assessment methods. A history and state of the art.* New York: Hay/McBer

Statsoft (2003). *Electronic statistics textbook*. Retrieved 2 October 2006 from http://www.statsoft.com/textbook/stathome.html

Thornton, G., & Byham, W. (1982). *Assessment centres and managerial performance*. New York: Academic Press

Tziner, A., Ronen, S., & Hacohen, D. (1993). A four-year validation study of an assessment center in a financial corporation. *Journal of Organizational Behavior, 14,* 225-237

Viswesvaran, C., Ones, D., & Schmidt, F. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology, 81*, 557-574

Welman, J., & Kruger, S. (2002). *Research methodology for the business administrative sciences*. (2nd ed.). Cape Town: Oxford University Press.

Wheeler, P., & Haertel, G. (1993). *Resource handbook on performance assessment and measurement*. Berkeley: The Owl Press

Woodruffe, C. (1992). What is meant by a competency? In *Designing and achieving competency. A competency-based approach to developing people and organizations*. London: McGraw-Hill

Zedeck, S. (1986). A process analysis of the assessment center method. In Staw, B. & Cummings L (eds.). *Research in organizational behavior*. Greenwich: JAI.

University of Cape Town

APPENDIX

Table 1.

Inter-correlations between Methods and Traits (Heterotrait-Monomethod; Heteromethod-Monotrait; and Heterotrait-Heteromethod approaches).

Variables	Customer focus_In-basket	Customer focus_BizCase Presentation	Customer focus_Written Exercise	Customer focus_OPQ	Customer focus_Role Play	Building Relationships_Career Interview	Building Relationships_ OPQ
Customer focus_In-basket	1.00						
Customer focus_BizCase Presentation	0.21	1.00					
Customer focus_Written Exercise	--	-0.33	1.00				
Customer focus_OPQ	0.29*	0.17	0.08	1.00			
Customer focus_Role Play	0.19	0.20	.56*	0.43*	1.00		
Building Relationships_Career Interview	0.22	0.21	--	0.02	0.01	1.00	
Building Relationships_ OPQ	0.26	0.07	0.18	0.52*	0.40*	0.14	1.00
Commitment_Career Interview	0.35	-0.03	-0.33	0.29	0.37	0.23	0.16
Commitment_Group Exercise	-0.07	0.02	0.20	0.02	0.27	0.18	0.03

University of Cape Town

Variables	Customer focus_In-basket	Customer focus_BizCase Presentation	Customer focus_Written Exercise	Customer focus_OPQ	Customer focus_Role Play	Building Relationships_Career Interview	Building Relationships_ OPQ
Commitment_OPQ	0.02	0.06	0.30	0.35*	0.54*	0.05	0.44*
Commitment_Role Play	-0.17	0.44*	0.25	0.43*	0.29*	0.19	0.30
Aligning Performance for success_In-basket	0.69*	0.24	--	0.18	-0.17	0.32*	0.24
Aligning Performance for success_OPQ	.065	-0.12	0.18	0.29*	0.17	-0.04	0.23
Aligning Performance for success_Role Play	-0.08	0.42	0.61*	0.22	0.49*	--	0.42
Facilitating Change_Career Interview	--	-0.19	0.82	0.09	0.22	0.16	0.24
Facilitating Change_Group Exercise	0.23	0.25	0.33	0.19	0.18	0.23	0.10
Facilitating Change_Bizcase Presentation	0.50	0.46	0.49*	0.47*	0.42*	--	0.75*
Facilitating Change_OPQ	0.10	-0.10	0.48*	0.12	0.16	0.37*	0.23
Facilitating Change_Role Play	0.50	0.39	0.58*	0.39*	0.56*	0.82*	0.83*

UNIVERSITY OF CAPE TOWN

Variables	Customer focus_In-basket	Customer focus_BizCase Presentation	Customer focus_Written Exercise	Customer focus_OPQ	Customer focus_Role Play	Building Relationships_Career Interview	Building Relationships_ OPQ
Thinking skills_In-basket	0.50*	0.19	-0.11	0.24	0.20	0.12	0.25
Thinking skills_Group Exercise	0.33*	0.04	0.40	0.12*	0.45*	0.12	0.04
Thinking skills_Bizcase Presentation	0.40*	-0.10	0.57*	0.26*	0.48*	-0.07	0.35*
Thinking skills_Written Exercise	--	-0.61	0.62*	0.56*	0.84*	-0.87	0.80
Thinking skills_OPQ	0.22	0.22	0.27	0.52*	0.50*	-0.12	0.23
Thinking skills_Role Play	0.05	0.52*	0.50	0.21	-0.07	-0.08	0.19
Decision making_In-basket	0.59*	0.13	-0.20	0.23	0.23	0.09	0.29*
Decision making_Group Exercise	0.46*	0.23	0.58*	0.11	0.47*	0.40*	-0.10
Decision making_Bizcase Presentation	0.27	0.19	0.60	0.29*	0.56*	-0.12	0.32*
Decision making_Written Exercise	--	-0.58	0.50*	0.34	0.60*	-0.50	0.47

University of Cape Town

Variables	Customer focus_In-basket	Customer focus_BizCase Presentation	Customer focus_Written Exercise	Customer focus_OPQ	Customer focus_Role Play	Building Relationships_Career Interview	Building Relationships_ OPQ
Decision making_OPQ	0.04	0.16	0.24	0.31*	0.38*	-0.17	0.34*
Decision making_Role Play	0.02	-0.07	0.35	0.57*	0.43*	-0.04	0.34
Initiating Action_In-basket	0.47*	0.14	0.48	0.10	0.01	0.27	0.19
Initiating Action_OPQ	0.20	0.17	0.24	0.34*	0.14	0.32*	0.50*
Initiating Action_Bizcase Presentation	0.39*	0.31*	0.23	0.24	0.38*	-0.01	0.38*
Initiating Action_Written Exercise	--	--	0.52*	0.54*	0.44	--	--
Planning & Organising_In-basket	0.50*	0.17	-0.20	0.06	-0.23	0.25	-0.02
Planning & Organising_OPQ	0.12	-0.08	0.34	0.26*	0.20	0.03	0.34*
Business Awareness_Bizcase Presentation	--	0.49	0.17	0.29	0.11	--	0.34
Business Awareness_OPQ	--	1.00	0.55*	0.07	0.50*	--	-0.20

UNIVERSITY OF CAPE TOWN

Variables	Commitment_ Career Interview	Commitment_ Group Exercise	Commitment_ OPQ	Commitment_ Role Play	Aligning Performance for success_In-basket	Aligning Performance for success_OPQ	Aligning Performance for success_Role Play
Customer focus_ In-basket	0.35	-0.08	0.02	-0.17	0.69*	0.07	-0.08
Customer focus_ BizCase Presentation	-0.03	0.02	0.06	0.44*	0.24	-0.12	0.42
Customer focus_ Written Exercise	-0.33	0.20	0.30	0.25	--	0.18	0.61*
Customer focus_OPQ	0.29	0.02	0.35*	0.43*	0.18	0.29*	0.22
Customer focus_Role Play	0.34	0.27	0.54*	0.29*	-0.17	0.17	0.49*
Building Relationships_Career Interview	0.23	0.18	0.05	0.19	0.32*	-0.04	--
Building Relationships_ OPQ	0.16	0.03	0.44*	0.30	0.24	0.23	0.42
Commitment_Career Interview	1.00						
Commitment_Group Exercise	0.12	1.00					
Commitment_OPQ	0.23	0.26*	1.00				
Commitment_Role Play	0.29	0.18	0.43*	1.00			

University of Cape Town

Variables	Commitment_ Career Interview	Commitment_ Group Exercise	Commitment_ OPQ	Commitment_ Role Play	Aligning Performance for success_In-basket	Aligning Performance for success_OPQ	Aligning Performance for success_Role Play
Aligning Performance for success_In-basket	0.12	0.05	-0.01	-0.04	1.00		
Aligning Performance for success_OPQ	-0.06	0.23	0.37*	0.13	0.30*	1.00	
Aligning Performance for success_Role Play	0.07	0.27	0.44*	0.53*	0.02	0.40*	1.00
Facilitating Change_ Career Interview	0.10	0.30	0.25	0.24	0.16	0.18	0.42
Facilitating Change_ Group Exercise	0.05	0.62*	0.20	0.21	0.28	0.06	-0.25
Facilitating Change_ Bizcase Presentation	--	0.31	0.26	0.28	0.00	0.20	0.35
Facilitating Change_OPQ	0.08	0.15	0.27*	0.09	0.24	0.32*	0.34
Facilitating Change_ Role Play	0.53	-0.04	0.53*	0.50*	0.34	0.38*	0.70*
Thinking skills_ In-basket	0.01	0.06	0.19	0.12	0.65*	0.42*	0.37
Thinking skills_ Group Exercise	0.28	0.54*	0.11	0.08	0.29*	0.31*	0.14

University of Cape Town

Variables	Commitment_ Career Interview	Commitment_ Group Exercise	Commitment_ OPQ	Commitment_ Role Play	Aligning Performance for success_In-basket	Aligning Performance for success_OPQ	Aligning Performance for success_ Role Play
Thinking skills_ Bizcase Presentation	-0.06	0.27*	0.22	0.14	0.19	0.18	0.33*
Thinking skills_Written Exercise	0.50	0.48*	0.25	0.50*	-0.82	0.24	0.53*
Thinking skills_OPQ	0.08	-0.02	0.19	0.19	0.12	0.16	0.32
Thinking skills_Role Play	-0.05	-0.18	0.01	0.44*	0.22	-0.10	0.22
Decision making_In-basket	0.36	0.06	0.13	0.06	0.69*	0.31*	-0.04
Decision making_Group Exercise	0.32	0.44*	0.04	0.09	0.52*	0.19	0.32
Decision making_Bizcase Presentation	0.25	0.07	0.46*	0.31	0.11	0.20	0.29
Decision making_Written Exercise	--	0.34	-0.21	0.60*	-0.45	-0.03	0.44
Decision making_OPQ	0.07	0.06	0.37*	0.18	0.04	0.35*	0.40*

University of Cape Town

Variables	Commitment_ Career Interview	Commitment_ Group Exercise	Commitment_ OPQ	Commitment_ Role Play	Aligning Performance for success_In-basket	Aligning Performance for success_OPQ	Aligning Performance for success_Role Play
Decision making_Role Play	0.24	0.19	0.24	0.48*	0.07	0.14	0.33*
Initiating Action_In-basket	0.07	0.01	0.02	0.03	0.73*	0.18	-0.05
Initiating Action_OPQ	-0.14	0.11	0.38*	0.19	0.25	0.37*	0.34*
Initiating Action_Bizcase Presentation	0.08	0.11	0.29*	0.40*	--	0.05	0.21
Initiating Action_ Written Exercise	--	-0.05	0.08	0.31	--	0.20	0.49*
Planning & Organising_In-basket	0.13	0.02	-0.03	0.09	--	0.24	-0.27
Planning & Organising_OPQ	0.15	0.07	0.41*	0.35	--	0.42*	0.27
Business Awareness_ Bizcase Presentation	--	-0.19	0.01	-0.04	0.58	0.29	0.34
Business Awareness_OPQ	--	0.29	-0.14	0.37	0.71	0.22	0.60*

University of Cape Town

Variables	Facilitating Change_Career Interview	Facilitating Change_Group Exercise	Facilitating Change_Bizcase Presentation	Facilitating Change_OPQ	Facilitating Change_Role Play	Thinking skills_In-basket	Thinking skills_Group Exercise
Customer focus_ In-basket	--	0.23	0.50	0.10	0.50	0.50*	0.33*
Customer focus_BizCase Presentation	-0.19	0.25	0.46	-0.10	0.39	0.19	0.04
Customer focus_Written Exercise	0.82	0.33	0.49*	0.48	0.58*	-0.11	0.40
Customer focus_OPQ	0.10	0.20	0.47*	0.12	0.39*	0.24	0.12
Customer focus_Role Play	0.22	0.18	0.43*	0.16	0.57*	0.20	0.45*
Building Relationships_Career Interview	0.16	0.23	--	0.37*	0.82*	0.12	0.12
Building Relationships_ OPQ	0.24	0.10	0.75*	0.23	0.83*	0.25	0.04
Commitment_Career Interview	0.10	0.05	--	0.08	0.53	0.01	0.28
Commitment_Group Exercise	0.30	0.62*	0.31	0.15	-0.04	0.06	0.56*
Commitment_OPQ	0.25	0.20	0.26	0.27*	0.53*	0.19	0.11

University of Cape Town

Variables	Facilitating Change_Career Interview	Facilitating Change_Group Exercise	Facilitating Change_Bizcase Presentation	Facilitating Change_OPQ	Facilitating Change_Role Play	Thinking skills_In-basket	Thinking skills_Group Exercise
Commitment_Role Play	0.24	0.21	0.28	0.09	0.50*	0.12	0.08
Aligning Performance for success_In-basket	0.16	0.28	0.00	0.24	0.34	0.65*	0.29*
Aligning Performance for success_OPQ	0.18	0.06	0.20	0.32*	0.39*	0.42*	0.31*
Aligning Performance for success_Role Play	0.42	-0.25	0.35	0.34	0.70*	0.37	0.14
Facilitating Change_Career Interview	1.00						
Facilitating Change_Group Exercise	0.29	1.00					
Facilitating Change_Bizcase Presentation	--	--	1.00				
Facilitating Change_OPQ	--	0.32*	0.58*	1.00			
Facilitating Change_Role Play	0.45	-0.19	0.57*	0.66*	1.00		
Thinking skills_In-basket	--	0.20	0.10	0.22	0.57	1.00	

University of Cape Town

Variables	Facilitating Change_Career Interview	Facilitating Change_Group Exercise	Facilitating Change_Bizcase Presentation	Facilitating Change_OPQ	Facilitating Change_Role Play	Thinking skills_In-basket	Thinking skills_Group Exercise
Thinking skills_Group Exercise	0.40*	0.58*	--	0.09	0.54	0.43*	1.00
Thinking skills_Bizcase Presentation	0.31	0.37	0.63	0.40	0.29	0.50	0.41
Thinking skills_Written Exercise	0.87	1.00*	0.56*	0.31	0.49*	-0.87	0.58
Thinking skills_OPQ	-0.04	0.41	0.17	0.24	0.29	0.19	0.26
Thinking skills_Role Play	0.29	0.26	0.78*	-0.13	0.26	0.06	0.04
Decision making_In-basket	0.16	0.31	0.25	0.22	0.50	0.81*	0.48*
Decision making_Group Exercise	0.38*	0.58*	0.64*	0.33*	0.42*	0.46*	0.72*
Decision making_Bizcase Presentation	0.13	0.23	0.51*	0.14	0.46*	0.36*	0.25
Decision making_Written Exercise	--	0.50	0.50*	0.34	0.29	--	-0.50
Decision making_OPQ	0.10	0.18	0.34	0.26*	0.42*	0.22	0.03
Decision making_Role Play	0.16	0.27	0.40*	0.27	0.47*	0.37	0.34

UNIVERSITY OF CAPE TOWN

Variables	Facilitating Change_Career Interview	Facilitating Change_Group Exercise	Facilitating Change_Bizcase Presentation	Facilitating Change_OPQ	Facilitating Change_Role Play	Thinking skills_In-basket	Thinking skills_Group Exercise
Initiating Action_In-basket	0.20	0.21	-0.41	0.20	0.49	0.60*	0.15
Initiating Action_OPQ	-0.01	0.26	0.35	0.35*	0.33	0.36*	-0.09
Initiating Action_Bizcase Presentation	0.32	0.36	0.50*	0.31*	0.46*	0.23	0.33
Initiating Action_Written Exercise	--	--	0.64*	0.64*	0.59*	--	--
Planning & Organising_In-basket	0.12	0.28	0.38	0.17	-0.25	0.55*	0.27
Planning & Organising_OPQ	0.13	0.44*	0.53	0.25	0.34	0.45*	0.12
Business Awareness_Bizcase Presentation	--	--	0.47*	0.34	0.48*	0.30	--
Business Awareness_OPQ	--	--	0.46*	0.25	0.46*	-0.65	--

University of Essex Library

Variables	Thinking skills_ Bizcase Presentation	Thinking Skills_ Written Exercise	Thinking skills_OPQ	Thinking skills_ Role Play	Decision making_ In-basket	Decision making_ Group Exercise	Decision making_ Bizcase Presentation
Customer focus_In-basket	0.39*	--	0.22	0.05	0.59*	0.46*	0.27
Customer focus_BizCase Presentation	-0.10	-0.61	0.22	0.52*	0.13	0.23	0.19
Customer focus_Written Exercise	0.57*	0.66*	0.27	0.50	-0.20	0.58*	0.60
Customer focus_OPQ	0.26*	0.56*	0.52*	0.21	0.23	0.11	0.29*
Customer focus_Role Play	0.48*	0.84*	0.50*	-0.07	0.23	0.47*	0.56*
Building Relationships_Career Interview	-0.07	-0.87	-0.12	-0.08	0.09	0.40*	-0.12
Building Relationships_ OPQ	0.35*	0.80	0.23	0.19	0.29*	-0.10	0.32*
Commitment_Career Interview	-0.06	0.50	0.08	-0.05	0.36	0.32	0.25
Commitment_Group Exercise	0.27*	0.48*	-0.02	-0.18	0.06	0.44*	0.07
Commitment_OPQ	0.22	0.25	0.19	0.01	0.13	0.04	0.46*

University of Cape Town

Variables	Thinking skills_Bizcase Presentation	Thinking skills_Written Exercise	Thinking skills_OPQ	Thinking skills_Role Play	Decision making_ In-basket	Decision making_Group Exercise	Decision making_Bizcase Presentation
Commitment_Role Play	0.14	0.50*	0.19	0.44*	0.06	0.09	0.31
Aligning Performance for success_In-basket	0.19	-0.82*	0.12	0.22	0.69*	0.52*	0.11
Aligning Performance for success_OPQ	0.18	0.24	0.16	-0.10	0.31*	0.19	0.20
Aligning Performance for success_Role Play	0.33*	0.53*	0.32	0.22	-0.04	0.32	0.29
Facilitating Change_Career Interview	0.31	0.87	-0.04	0.29	0.16	0.38*	0.13
Facilitating Change_Group Exercise	0.37	1.00*	0.41	0.26	0.31	0.58*	0.23
Facilitating Change_Bizcase Presentation	0.63*	0.56*	0.17	0.78*	0.25	0.64*	0.51*
Facilitating Change_OPQ	0.40*	0.31	0.24	-0.13	0.22	0.33*	0.14
Facilitating Change_Role Play	0.29	0.49*	0.29	0.26	0.50	0.41*	0.46
Thinking skills_In-basket	0.49*	-0.87	0.19	0.06	0.81*	0.46*	0.36

University of Cape Town

Variables	Thinking skills_Bizcase Presentation	Thinking skills_Written Exercise	Thinking skills_OPQ	Thinking skills_Role Play	Decision making_In-basket	Decision making_Group Exercise	Decision making_Bizcase Presentation
Thinking skills_Group Exercise	0.41*	0.58	0.26	0.04	0.48*	0.72*	0.25
Thinking skills_Bizcase Presentation	1.00						
Thinking skills_Written Exercise	0.62*	1.00					
Thinking skills_OPQ	0.20	0.62*	1.00				
Thinking skills_Role Play	-0.10	--	0.17	1.00			
Decision making_In-basket	0.50*	-0.33	0.04	0.02	1.00		
Decision making_Group Exercise	0.51*	0.44*	0.22	0.08	0.53*	1.00	
Decision making_Bizcase Presentation	0.66*	0.41	0.14	0.05	0.42*	0.39*	1.00
Decision making_Written Exercise	0.72*	0.62*	0.22	--	0.50	0.55*	0.69
Decision making_OPQ	0.25*	0.52*	0.45*	0.05	0.27	0.17	0.49*

University of Cape Town

Variables	Thinking skills_Bizcase Presentation	Thinking skills_Written Exercise	Thinking skills_OPQ	Thinking skills_Role Play	Decision making_In-basket	Decision making_Group Exercise	Decision making_Bizcase Presentation
Decision making_Role Play	0.41*	0.54*	0.22	0.26	0.47*	0.40*	0.12
Initiating Action_In-basket	0.46*	-0.87	-0.09	0.21	0.72*	0.42*	0.21
Initiating Action_OPQ	0.29*	-0.03	0.22	-0.07	0.28*	0.14	0.38*
Initiating Action_Bizcase Presentation	0.39*	0.49*	0.11	0.02	0.46*	0.36*	0.41*
Initiating Action_Written Exercise	0.70*	0.55*	0.25	--	--	0.51*	--
Planning & Organising_ In-basket	0.21	-1.00	-0.10	0.20	0.67*	0.43*	0.20
Planning & Organising_OPQ	0.34*	0.25	0.21	0.15	0.45*	0.30	0.24
Business Awareness_Bizcase Presentation	0.21	0.17	0.22	0.42	--	0.23	-0.49
Business Awareness_OPQ	0.53*	0.65*	0.40	0.87	--	0.53*	-0.33

University of Cape Town

Variables	Decision making_ Written Exercise	Decision making_ OPQ	Decision making_ Role Play	Initiating Action_ In-basket	Initiating Action_ OPQ	Business Awareness_ Bizcase Presentation	Business Awareness _OPQ	Initiating Action_ Bizcase Presentation	Initiating Action_ Written Exercise	Planning & Organising _ in-basket	Planning & Organising _OPQ
Customer focus_ In-basket	--	0.04	0.02	0.47*	0.20	--	--	0.39*	--	0.50*	0.12
Customer focus_ BizCase Presentation	-0.58	0.16	-0.07	0.14	0.17	0.49	1.00*	0.31*	--	0.17	-0.08
Customer focus_Written Exercise	0.50*	0.24	0.35	0.48	0.24	0.17	0.55*	0.23	0.52*	-0.20	0.34
Customer focus_OPQ	0.34	0.31*	0.51*	0.10	0.34*	0.29	0.07	0.24	0.54*	0.06	0.26*
Customer focus_Role Play	0.60*	0.38*	0.42*	0.01	0.14	0.11	0.50*	0.38*	0.44	-0.23	0.20
Building Relationships_Career Interview	-0.50	-0.17	-0.04	0.27	0.32*	--	--	-0.01	--	0.25	0.03
Building Relationships_ OPQ	0.47	0.34*	0.34	0.19	0.50*	0.34	-0.19	0.38*	--	-0.02	0.33*
Commitment_ Career Interview	--	0.07	0.24	0.07	-0.14	--	--	0.08	--	0.13	0.15

University of Cape Town

Variables	Decision making_ Written Exercise	Decision making_ OPQ	Decision making_ Role Play	Initiating Action_ In-basket	Initiating Action_ OPQ	Business Awareness_ Bizcase Presentation	Business Awareness _OPQ	Initiating Action_ Bizcase Presentation	Initiating Action_ Written Exercise	Planning & Organising _ In-basket	Planning & Organising _OPQ
Commitment_ Group Exercise	0.34	0.06	0.19	0.01	0.11	-0.19	0.29	0.11	-0.05	0.02	0.07
Commitment_ OPQ	-0.21	0.37*	0.24	0.02	0.38*	0.01	-0.14	0.29*	0.08	-0.03	0.41*
Commitment_ Role Play	0.60*	0.18	0.48*	0.03	0.19	-0.04	0.37	0.40*	0.31	0.09	0.35
Aligning Performance for success_ In-basket	-0.45	0.04	0.07	0.73*	0.25	0.58	0.71	0.31*		0.79*	0.41
Aligning Performance for success_ OPQ	-0.03	0.35*	0.14	0.18	0.37*	0.29	0.22	0.05	0.20	0.24	0.42*
Aligning Performance for success_ Role Play	0.44	0.40*	0.33*	-0.05	0.34*	0.34	0.59*	0.21	0.48*	-0.27	0.27
Facilitating Change_ Career Interview	--	0.10	0.16	0.20	-0.01	--	--	0.32	--	0.12	0.13
Facilitating Change_ Group Exercise	0.50	0.18	0.27	0.21	0.26	--	--	0.36	--	0.28	0.44*

University of Cape Town

Variables	Decision making_ Written Exercise	Decision making_ OPQ	Decision making_ Role Play	Initiating Action_ In-basket	Initiating Action_ OPQ	Business Awareness_ Bizcase Presentation	Business Awareness _OPQ	Initiating Action_ Bizcase Presentation	Initiating Action_ Written Exercise	Planning & Organising _ In-basket	Planning & Organising _OPQ
Facilitating Change_ Bizcase Presentation	0.50*	0.34	0.40*	-0.41	0.35	0.47*	0.46*	0.50*	0.69*	0.38	0.53
Facilitating Change_OPQ	0.34	0.26*	0.27	0.20	0.35*	0.34	0.25	0.31*	0.63*	0.17	0.25
Facilitating Change_Role Play	0.30*	0.42*	0.47*	0.49	0.33	0.48*	0.46*	0.46*	0.59*	-0.25	0.34
Thinking skills_In-basket	--	0.22	0.37	0.60	0.36*	0.30	-0.65	0.23	--	0.55*	0.45*
Thinking skills_Group Exercise	-0.50	0.03	0.34	0.15	-0.09	--	--	0.33	--	0.27	0.12
Thinking skills _Bizcase Presentation	0.72*	0.28*	0.41*	0.46*	0.29*	0.21	0.53*	0.39*	0.70*	0.21	0.34*
Thinking skills_Written Exercise	0.62*	0.52*	0.54*	-0.87	-0.03	0.17	0.65*	0.49*	0.55*	-1.00	0.25
Thinking skills_OPQ	0.22	0.45*	0.22	-0.09	0.22	0.22	0.40	0.11	0.25	-0.10	0.21
Thinking skills_Role Play	--	0.05	0.26	0.21	-0.07	0.42	0.87	0.02	--	0.20	0.15
Decision making_In-basket	0.50	0.27	0.47*	0.72*	0.28*	--	--	0.46*	--	0.67*	0.45*

University of Cape Town

Variables	Decision making_ Written Exercise	Decision making_ OPQ	Decision making_ Role Play	Initiating Action_ In-basket	Initiating Action_ OPQ	Business Awareness_ Bizcase Presentation	Business Awareness _OPQ	Initiating Action_ Bizcase Presentation	Initiating Action_ Written Exercise	Planning & Organising _ In-basket	Planning & Organising _OPQ
Decision making_ Group Exercise	0.55*	0.17	0.40*	0.42*	0.14	0.23	0.53*	0.36*	0.51*	0.43*	0.30
Decision making_ Bizcase Presentation	0.69	0.49*	0.12	0.21	0.37*	-0.49	-0.33	0.41*	--	0.20	0.24
Decision making_Written Exercise	1.00					-0.01	0.53	0.27	0.38	-1.00	0.76
Decision making_OPQ	0.18	1.00				0.09	0.32	0.36*	0.24	0.23	0.31*
Decision making_ Role Play	0.83*	0.26	1.00			0.07	0.29	0.32*	0.44	0.24	0.39*
Initiating Action_ In-basket	--	0.10	0.35	1.00		-0.58	--	0.30	--	0.66*	0.42*
Initiating Action_OPQ	0.20	0.40*	0.21	0.21	1.00	0.20	0.20	0.22	0.38	0.14	0.49
Initiating Action_Bizcase Presentation	0.27	0.36*	0.32*	0.30	0.22	0.33	0.23	1.00		0.34*	0.06
Initiating Action_Written Exercise	0.38	0.24	0.44	--	0.38	0.65*	0.54*	0.68*	1.00	--	--
Planning & Organising_ In-basket	-1.00	0.23	0.24	0.66*	0.14	--	--	0.34*	--	1.00	

University of Otago Thesis

Variables	Decision making_ Written Exercise	Decision making_ OPQ	Decision making_ Role Play	Initiating Action_ In-basket	Initiating Action_ OPQ	Business Awareness_ Bizcase Presentation	Business Awareness _OPQ	Initiating Action_ Bizcase Presentation	Initiating Action_ Written Exercise	Planning & Organising _ In-basket	Planning & Organising _OPQ
Planning & Organising_ OPQ	0.76	0.31*	0.39*	0.42*	0.49	0.19	-0.11	0.06	--	0.39*	1.00
Business Awareness_ Bizcase Presentation	-0.01	0.09	0.07	-0.58	0.20	1.00		0.33	0.67*	--	0.19
Business Awareness_ OPQ	0.53*	0.32	0.29	--	0.20	0.46*	1.00	0.22	0.54*	--	-0.11

Note. Unavailable data is indicated by -- ; an empty cell indicates that the inter-correlation has already been computed and has been illustrated in a different cell; a shaded cell illustrates the Heterotrait-heteromethod approach; the unshaded cells illustrate the Heterotrait-monomethod approach; and the underlined cells illustrate convergent validity. Bizcase Presentation stands for Business Case Presentation

*$p > 0.05$

University of Cape Town