Psychomotor ability and learning potential as predictors of driver and machine operator performance in a road construction company

by

LOUIS PETRUS OLIVIER

submitted in accordance with the requirements for the degree of

MASTER OF ARTS

in the subject

INDUSTRIAL AND ORGANISATIONAL PSYCHOLOGY

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: PROF M DE BEER

JUNE 2015
I declare that the dissertation “PSYCHOMOTOR ABILITY AND LEARNING POTENTIAL AS PREDICTORS OF DRIVER AND MACHINE OPERATOR PERFORMANCE IN A ROAD CONSTRUCTION COMPANY” is my own work, and that all the sources that I have used or have quoted from have been indicated and acknowledged by means of complete references.

________________________________________

SIGNATURE
(Louis Olivier)

________________________________________

DATE
ACKNOWLEDGEMENTS

First and foremost I acknowledge my Father who gives me the strength and courage to walk these rough seas.

I also would like to acknowledge and express gratitude to the following persons:

- Sumari Olivier, for her endless support, example and unconditional love during this difficult time
- Matthew, Lilé, Yishay and Rainin Olivier for your understanding, patience and forgiveness
- My supervisor, professor Marié de Beer for her specialist support and guidance
- Mr James Stamp for his expert assistance with the statistical analyses of the data, and
- Grant Daly who afforded me the opportunity to further my studies and to follow my dream
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DECLARATION</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>vii</td>
</tr>
<tr>
<td>SUMMARY</td>
<td>viii</td>
</tr>
<tr>
<td><strong>CHAPTER 1</strong></td>
<td></td>
</tr>
<tr>
<td><strong>SCIENTIFIC ORIENTATION TO THE RESEARCH</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1 <strong>BACKGROUND AND MOTIVATION</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.2 <strong>PROBLEM STATEMENT</strong></td>
<td>9</td>
</tr>
<tr>
<td>1.2.1 Research Question and Hypotheses</td>
<td>10</td>
</tr>
<tr>
<td>1.3 <strong>RESEARCH AIMS</strong></td>
<td>11</td>
</tr>
<tr>
<td>1.3.1 Specific Literature Aims</td>
<td>11</td>
</tr>
<tr>
<td>1.3.2 Specific Empirical Aims</td>
<td>11</td>
</tr>
<tr>
<td>1.4 <strong>PARADIGM PERSPECTIVE</strong></td>
<td>12</td>
</tr>
<tr>
<td>1.5 <strong>THE RESEARCH DESIGN</strong></td>
<td>13</td>
</tr>
<tr>
<td>1.5.1 Research Approach</td>
<td>13</td>
</tr>
<tr>
<td>1.5.2 Unique Conditions of the Study</td>
<td>14</td>
</tr>
<tr>
<td>1.5.3 Research Method</td>
<td>15</td>
</tr>
<tr>
<td>1.5.4 Research Participants</td>
<td>15</td>
</tr>
<tr>
<td>1.5.5 Ethical considerations</td>
<td>16</td>
</tr>
<tr>
<td>1.6 <strong>MEASURING INSTRUMENTS</strong></td>
<td>17</td>
</tr>
<tr>
<td>1.6.1 Vienna Test System</td>
<td>17</td>
</tr>
<tr>
<td>1.6.1.1 Cognitrone</td>
<td>17</td>
</tr>
<tr>
<td>1.6.1.2 Determination Unit</td>
<td>18</td>
</tr>
<tr>
<td>1.6.1.3 Two-Hand Coordination Speed and Accuracy</td>
<td>18</td>
</tr>
<tr>
<td>1.6.1.4 Time Movement/Anticipation Test (ZBA)</td>
<td>19</td>
</tr>
<tr>
<td>1.6.2 TRAM Assessment Battery</td>
<td>19</td>
</tr>
<tr>
<td>1.6.3 Performance Appraisals</td>
<td>20</td>
</tr>
<tr>
<td>1.6.4 Supervisor Rankings</td>
<td>23</td>
</tr>
<tr>
<td>1.7 <strong>RESEARCH PROCEDURE</strong></td>
<td>23</td>
</tr>
<tr>
<td>1.7.1 Phase 1: Literature Review</td>
<td>23</td>
</tr>
<tr>
<td>1.7.2 Phase 2: Empirical Study</td>
<td>24</td>
</tr>
<tr>
<td>1.8 <strong>STATISTICAL ANALYSIS</strong></td>
<td>24</td>
</tr>
</tbody>
</table>
CHAPTER 2
LITERATURE REVIEW

2.1 INTRODUCTION

2.2 PSYCHOMOTOR ABILITY
  2.2.1 Defining Psychomotor Ability
  2.2.2 The Historical Development of Psychomotor Ability
  2.2.3 The Relationship between Psychomotor Ability and Cognitive Ability
  2.2.4 Psychomotor Ability as a Predictor of Job Performance

2.3 LEARNING POTENTIAL
  2.3.1 An Introduction to Intelligence
  2.3.2 Intelligence Theories
  2.3.2.1 The Structural Approach
  2.3.2.2 The Information Processing Approach
  2.3.2.3 Cognitive Components Approach
  2.3.2.4 The Learning/Dynamic Approach
  2.3.2.5 Taylor’s (1994) Integrated Theory

2.4 COGNITIVE ASSESSMENT IN THE SOUTH AFRICAN CONTEXT
  2.4.1 Learning Potential and Dynamic Assessment in South Africa

2.5 DYNAMIC ASSESSMENT
  2.5.1 Dynamic Assessment and Intelligence
  2.5.2 A Comparison between Traditional and Dynamic Cognitive Assessments
  2.5.3 Dynamic Assessment in South Africa

2.6 JOB PERFORMANCE AND PERFORMANCE MANAGEMENT
  2.6.1 Introduction
  2.6.2 Conceptualising Job Performance
  2.6.2.1 A Definition of Job Performance
  2.6.2.2 Factors Affecting Individual Performance
  2.6.3 Measuring Job Performance
  2.6.3.1 Defining Performance Measurement
2.6.3.2 The Value of Performance Measurement 68
2.6.3.3 Performance Measurement Criteria 69
2.6.3.3.1 Dimensions of Performance Criteria 69
2.6.3.3.2 Objective Measurement Criteria 70
2.6.3.3.3 Subjective Measurement Criteria 70
2.6.3.4 The Performance Appraisal 71
2.6.3.5 Supervisor Ranking 74
2.6.4 Performance Management 76
2.6.4.1 The Performance Management System 77
2.6.4.2 Different Approaches to Job Performance Management 82
2.6.4.2.1 The Systems Approach 82
2.6.4.2.2 Sink and Tuttle’s Approach 83
2.6.4.2.3 The Balanced Scorecard Approach 83
2.6.4.2.4 The Performance Pyramid 84
2.6.4.2.5 Medori and Steeple’s Approach to Job Performance Management 85
2.6.5 Performance Management in the Research Organisation 86
2.6.6 The Future of Performance Management 88
2.6.7 The Role of Industrial Psychologists in Managing Employee Performance 91

2.7 CHAPTER SUMMARY 93

CHAPTER 3

CHAPTER 4
CONCLUSIONS, LIMITATIONS AND RECOMMENDATIONS 146

4.1 INTRODUCTION 146

4.2 THE HYPOTHESES 146

4.3 CONCLUSIONS 146

4.4 LIMITATIONS 148
4.4.1 Limitations Pertaining to the Criteria 148

5. RECOMMENDATIONS 151
5.1 Practice 154
5.2  Further Research     154

6.  CHAPTER SUMMARY     155

REFERENCES     156

LIST OF TABLES

Table 1.1:  Job Performance Dimensions     64

ARTICLE (CHAPTER 3) TABLES

Table 1:  Descriptive Statistics     112
Table 2:  Intercorrelations and P-values for VTS Sub-Scores     114
Table 3:  Intercorrelations and P-values for TRAM Sub-Scores     114
Table 4:  Intercorrelations and P-values for Monthly
           Performance Scores     115
Table 5:  TRAM Mean Scores of the Comparison Group (Group 2)     116
Table 6:  TRAM Scores According to the Level of Education for
           the Sample Group     117
Table 7:  TRAM Mean Scores for the Second Comparison
           Group (Group 3)     119
Table 8:  TRAM and VTS Test Results Split According to
           Age Group     120
Table 9:  Results Obtained from the Shapiro-Wilk Normality Test     121
Table 10: Correlation Coefficients and P-values     124
Table 11: Sample Group Rankings According to Mean Scores     126
Table 12: Ranking Movement According to Mean Scores     127
SUMMARY

PSYCHOMOTOR ABILITY AND LEARNING POTENTIAL AS PREDICTORS OF DRIVER AND MACHINE OPERATOR PERFORMANCE IN A ROAD CONSTRUCTION COMPANY

By: Louis Olivier
Institution: University of South Africa
Student number: 45951020
Degree: MA
Subject: Industrial and Organisational Psychology
Supervisor: Professor Marié de Beer

The changing nature of work and its competitive characteristics are global phenomena and are mainly fuelled by ongoing technological advancement. This creates unique challenges for talent attraction and the retention of high performing individuals. In addition, the global workforce is becoming more diverse due to demographic, societal and cultural changes and companies are placing greater demands on employee competency and performance. Managing the human factor as a strategic asset in organisations remains a primary challenge in securing a competitive advantage.

The road construction industry in South Africa is no different. There is growing competition between civil engineering contractors to secure tenders and to maximise profitability. This is only possible with a sufficient and sustainable labour force. Valid selection processes are therefore required to ensure that the most productive individuals are selected for the most suitable jobs. Reliable and valid performance predictors will assist employers in making appropriate selection decisions. Selecting high performing individuals will support and enhance overall organisational performance.
In this study the investigation focused on whether psychomotor ability and learning potential are statistically significant predictors of work performance - with specific reference to drivers and machine operators in a road construction company. A quantitative approach was followed to investigate the relationships between variables, or then the prediction of one dependent variable (driver and machine operator performance) by means of two independent variables (psychomotor ability and learning potential).

Results from the study did not indicate any statistically significant relationships between the variables. Only scientifically validated assessment instruments were used in the study - which means the findings led to a renewed focus on the importance of performance measurement and the psychometric quality (reliability and validity) of performance data.

Key Terms

Psychomotor ability, learning potential, work performance, intelligence, dynamic assessment, cognitive ability, cognitive assessment, Vienna Test System (VTS), Transfer-Automisation-and-Memory (TRAM), psychometric test
CHAPTER 1
SCIENTIFIC ORIENTATION TO THE RESEARCH

The focus of this research was on psychomotor ability and learning potential as predictors of work performance. Chapter 1 contains the background and motivation, the problem statement, the aims, paradigm perspective, research design and method as well as the chapter layout.

1.1 BACKGROUND AND MOTIVATION

Managing employee performance became a strategic approach in the 21st century (Shamsi, 2010). Methods of attracting and securing future talent should therefore have higher levels of good work performance as one of its primary objectives. Reliable and valid performance predictors will assist employers in making appropriate selection decisions (Gilmore, 2008; Nzama, De Beer, & Visser, 2008; Pulakos, 2005; Schmidt & Hunter, 1998). Selecting high performing individuals will increase overall organisational performance, whereas an average or poor performer can drastically increase cost to the employer (Brudan, 2010; Shamsi, 2010).

In performance management the focus is on the broader spectrum of challenges facing organisations in today’s turbulent and changing work environment. The performance of employees enables a business to function effectively in its environment and contributes to the general direction in which it intends to go to achieve future goals. It is vital that organisations should employ individuals who correctly complete and manage the right tasks at the right time. Recruiting staff could be a very costly exercise (Cascio & Aguinis, 2010), but it is also an essential part of any business and it pays therefore to do it properly. When organisations choose the right people for the right job, train them well and treat them appropriately, these individuals not only produce good results but also tend to stay with the organisation for longer (Collings & Mellahi, 2009; Girard & Fallery, 2010). In such circumstances, the organisation’s initial and ongoing investment in its employees is well rewarded. An organisation may have all of the latest technology and the best physical
resources, but if it does not have the right people, it will struggle to achieve the results it requires (Chabault, Hulin & Soparnot, 2012). This is true across the whole range of business activity. It follows then that appropriate selection based on effective assessment methods can greatly enhance the quality and productivity of an organisation’s workforce (Murphy & Maree, 2006; Pulakos, 2005).

In this study the investigation was focused on whether psychomotor ability and learning potential are statistically and practically significant predictors of work performance - with specific reference to drivers and machine operators in a road construction company. The nature of the investigation accentuated the practical importance of the research. The identification (and appropriate utilisation) of effective predictors of performance could be used to refine and expand the selection methodology of drivers and machine operators. Furthermore, information obtained on the predictive validity of these predictors could inform other companies in the industry and could empower them to enhance their effectiveness to select high performing individuals. The study contributes to a renewed focus on employee performance and the effective management thereof.

There is a recent and growing body of literature focussing on the importance of employee performance at work (Aarabi, Subramaniam & Akeel, 2013; Brudan, 2010; Chabault et al., 2012; Meister & Willyard, 2010; Parry & Tyson, 2011; Schläfke, Silvi, & Möller, 2013; Schat & Frone, 2011). Performance of employees is significant for organisations (Aarabi et al., 2013) because of its direct impact on overall organisational productivity (Aarabi et al., 2013; Chabault et al., 2012; Gilmore, 2008). Job performance is central to the success of an organisation (Brudan, 2010; Chabault et al., 2012; Meister & Willyard, 2010; Parry & Tyson, 2011; Schläfke et al., 2013). Employee job performance represents the primary contribution of individuals to organisational effectiveness and the primary reason individuals are employed by organisations in the first place (Schat & Frone, 2011).

Additionally, the effect of performance of employees may easily extend to customer satisfaction, loyalty, brand image, and even purchase intention (Tsui, Lin &
Researchers have shown that performance is closely related to customer satisfaction and financial performance (Studer, 2008; Tsui et al., 2013). Consequently, defining and managing employee performance has become a strategic approach in the 21st century (Brudan, 2010; Meister & Willyard, 2010; Parry & Tyson, 2011; Schläfke et al., 2013), in the sense that performance management is concerned with the broader issues facing businesses. For instance, effective performance management is of strategic importance if businesses are to function effectively in a competitively defined environment (Chabault et al., 2012), and also with regard to the general direction in which the business is intended to go to achieve longer-term goals (Chabault et al., 2012; Shamsi, 2010). Appropriate selection and effective assessment methods can therefore greatly enhance the quality and productivity of an organisation’s workforce (Pulakos, 2005). A bad hiring decision can cost as much as five times the employee’s salary through a number of factors including poor employee performance, lost productivity of peers, unmet customer expectations and subsequent labour turnover costs (Jackson & Schuler, 2003; Pillay, 2009; Pulakos, 2005; Shamsi, 2010).

Personnel resource professionals and industrial psychologists have conducted empirical studies in an attempt to understand employee performance (Chabault et al., 2012; Meister & Willyard, 2010; Parry & Tyson, 2011; Schat & Frone, 2011). In this regard, assessment and selection of high potential employees is essential as they contribute toward organisational success (Aarabi et al., 2013; Brudan, 2010; Chabault et al., 2012; Gilmore, 2008; Meister & Willyard, 2010; Parry & Tyson, 2011; Schat & Frone, 2011; Schläfke et al., 2013). This research is therefore important in terms of contributing to the larger body of knowledge and of particular importance to the company involved in the research that wished to evaluate the predictive validity of its current assessment measures. Consequently, the study was focused specifically on psychomotor ability and learning potential (as predictors of performance) by investigating the statistical and practical significance of the relationships between these variables and work performance and to evaluate whether they can statistically significantly predict work performance.
Individual work performance was defined by Campbell (1990) as behaviours or actions that are relevant to the goals of the organisation. However, this definition meant that the focus of work performance is on behaviours or actions of employees, rather than the results of these actions. In addition, behaviours should be under the control of the individual, thus excluding behaviours that are influenced and constrained by the environment (Chabault et al., 2012; Galinsky & Matos, 2011). Traditionally, the main focus of the work performance construct has been on task performance, which can simply be defined as the proficiency with which individuals perform the core substantive or technical tasks central to their jobs. Behaviours used to describe task performance often include work quantity and quality, job skills, and job knowledge (Campbell, 1990; Cascio & Aguinis, 2010; Rotundo & Sackett, 2002).

It is now generally agreed that, in addition to task performance, the individual work performance domain consists of contextual work-related performance and counterproductive work behaviour (Koopmans et al., 2012). Contextual performance can be defined as behaviour that supports the organisational, social, and psychological environment (Koopmans et al., 2012) in which the technical core must function (Aguinis, 2013; Cascio & Aguinis, 2010). Behaviours used to describe contextual performance include, for example, employees demonstrating effort, the facilitation of peer and team performance, cooperation, and effective communication in a defined work environment (Campbell, 1990; Rotundo & Sackett, 2002). Counterproductive work behaviour can be defined as behaviour that harms the well-being of the organisation (Rotundo & Sackett, 2002). It includes behaviours such as absenteeism, off-task behaviour, theft, and substance abuse which have a negative impact on work performance (Koopmans et al., 2012).

A recent review by Koopmans et al. (2012) has identified the new and upcoming dimension of adaptive work performance. This dimension focuses on the growing interdependency and uncertainty of work systems and the corresponding change in the nature of work performance (Griffin, Neal & Parker, 2007; Pulakos, Arad, Donovan & Plamondon, 2000; Sinclair & Tucker, 2006). Adaptive performance can
be defined as the extent to which an individual adapts to changes in the work role or environment (Griffin et al., 2007; Koopmans et al., 2012).

Job performance concerns individual output in terms of quality and quantity expected from every employee in a particular job (Aguinis, 2013; Cascio & Aguinis, 2010). Performance is associated with quantity of output, quality of output, timeliness of output, presence / attendance on the job, efficiency of the work completed and effectiveness of work completed (Mathis & Jackson, 2009; Tinofirei, 2012). Employee performance, therefore, is about the timely, effective and efficient completion of mutually agreed tasks by the employee, as set out by the employer (Tinofirei, 2012). For Koopmans et al. (2012), individual work performance is based on a four-dimensional conceptual framework, consisting of task performance, contextual performance, adaptive performance, and counterproductive work behaviour.

Performance management refers to a continuous process of identifying, measuring and developing performance in organisations by linking each individual’s performance and objectives to the organisation’s overall mission and goals (Aguinis, 2013). Taticchi, Tonelli and Cagnazzo (2010) expanded on work done by Neely (2005) to complete a comprehensive summary of the use and development of work performance and the management thereof over the last two decades. It was found that interest in performance measurement and management (PMM) has notably increased in the last 20 years (Taticchi et al., 2010). Particularly, it is important to note the evolution from focusing performance on a financial perspective to a non-financial perspective. Since the mid 1980’s, companies emphasised the growing need of controlling production business processes (Aguinis, 2013; Cascio & Aguinis, 2010). Companies have understood that for competing in continuously changing business environments, it is necessary to monitor and understand job performance. Consequently, measurement has been recognised as a crucial element to improve business performance (Sharma, Bhagwat & Dangayach, 2005).
A performance measurement and management system (PMS) is a balanced and dynamic system that enables support of decision-making processes by gathering, elaborating and analysing information (Neely, 2005). The concept of balance refers to the need for using different measures and perspectives that when tied together give a holistic view of the organisation (Neely, 2005; Taticchi et al., 2010). The concept of dynamicity refers instead to the need of developing a system that can be used to continuously monitor the internal and external context and reviews objectives and priorities (Aguinis, 2013; Griffin et al., 2007; Taticchi et al., 2010). Increasingly competitive environments, the evolution of the quality concept, an increased focus on continuous improvement and the significant developments in information and communication technologies are the most important changes in recent years that have created a favourable context for the implementation of PMS’s in small and medium-sized enterprises (SME's), particularly in the manufacturing sector (Taticchi et al., 2010). Although extensive research has been carried out to investigate the needs and characteristics of PMSs in large organisations, there is a distinct lack of published research on issues related to SMEs (Hudson, Smart & Bourne, 2001) similar to the company involved in this research.

Employees with high performance potential selected to operate large road construction machinery require a specific set of abilities - none of which is more important than psychomotor ability. Psychomotor ability can be defined as the process of interaction between the perceptual systems (or five senses), the brain (where perceptual information is interpreted) and the body (where the individual reacts to such perceptual stimuli) (Tan, 2006). The concept therefore refers to the processing of information, making of decisions and putting them into action (usually by means of specific movements). Tan (2006) explained that psycho in this regard refers to the mind or psyche, and motor to the physiological body. More generally psychomotor can be seen as the mind-body interaction, and psychomotor abilities, as those capacities which allow for effective interaction between the two and within a specific environment (Tan, 2006). Mnguni (2011) noted that measures of psychomotor ability are commonly included in the selection batteries for two apparent reasons. Firstly, they have an obvious relation to the task and, secondly,
the results of validation research support their inclusion in selection batteries (Carretta & Ree, 2000; Keyser, 2012; Mnguni, 2011; Pelser, 2002) within particular contexts. For Keyser (2012), the term “psychomotor” denoted a combination of physical and psychological activities. She explained that research findings generally support the validity of psychomotor tests in the prediction of job performance in positions requiring operating or driving skills (Carretta & Ree, 2000; Martinussen, 1996; Keyser, 2012; Pelser 2002; Schoeman, 1995; Wheeler & Ree, 1997).

Russian psychologist, Vygotsky, is generally regarded as the founder of the learning potential concept and his zone of proximal development has become a key concept in learning potential assessment (De Beer, 2006; Hamers & Resing, 1993). According to De Beer (2000a), learning potential refers to an overall cognitive capacity and includes both present and potential or projected improved future performance. This implies that cognitive ability is dynamic and changeable and can therefore be increased.

According to Schoeman, De Beer and Visser (2008), learning potential has been developed as an alternative strategy for the assessment of cognitive functioning to address the inadequacy of conventional intelligence testing. Of particular importance to this study, Gilmore (2008) found that a statistically significant relationship exists between learning potential, as a predictor of performance, and work performance. Gilmore (2008) defined learning potential measures as an individual’s present level of ability as well as the potential for improvement.

A few dynamic tests have been developed and standardised in South Africa, namely: the Ability Processing of Information and Learning Battery (APIL-B), developed by Taylor (1994, 1999), the Transfer, Automisation, Memory and Understanding Learning Potential Battery (TRAM-I and TRAM-II) developed by Taylor (Taylor, 1999) and the Learning Potential Computerised Adaptive Test (LPCAT) developed by De Beer (2000b). The battery developed by Taylor (viz. TRAM-I, TRAM-II and APIL-B) is a pen-and-paper application. These assessments are time-consuming whereas the LPCAT is computer based and the time taken to complete the
assessment is roughly 30 to 45 minutes. The marking of the LPCAT is automatic and therefore eradicates marker error. The rationale for using Taylor’s batteries exclusively is that the company involved in the research is currently using the TRAM-I and TRAM-II and specifically required an evaluation of its effectiveness to predict performance.

General cognitive ability, or “g”, has consistently been found to be a valid predictor of job performance when compared to other potential predictors (De Beer, 2010; Keyser, 2012; Pelser, 2002, Taylor, 1999). However, culture-fairness in the assessment of cognitive potential is a particularly important issue in the South African context (Pelser, 2002; Taylor, 1999).

Taylor (1999) distinguished four elements in learning potential. Fluid intelligence is the type of ability used in the solution of novel intellectual problems, where no off-the-shelf or taught solution strategies are available (Taylor, 1999). Fluid intelligence involves conceptual thinking: the capacity of the individual to classify things or events into categories based on some underlying principle. Thinking in concepts or categories is much more powerful than thinking in terms of specific objects or events, and fosters the development of new competencies (Taylor, 1999). Information processing efficiency is the capacity to take in and handle information according to set procedures in a speedy and accurate manner (Taylor, 1999). Fast and accurate processing facilitates the solution of problems because information is not lost out of short term memory. Efficient processing also reserves mental capacity for the conceptual part of problem solving (Taylor, 1999). Automatisation, according to Taylor (1999) is the rate at which learning takes place in a novel task. All individuals, when they encounter a new task are slow and inefficient. With practice they become more efficient, but with some people the rate of improvement is just faster than with others (Taylor, 1999). Learning Rate or Automatisation can be investigated by giving individuals repeated chances to do a task or by giving them learning inputs such as a lesson (Taylor, 1999). Taylor’s (1999) fourth element, transfer, is the capacity to adapt current knowledge or procedures to new ones of a related nature. Transfer is a critical ability, especially in more challenging work, but necessary at all levels.
Transfer demands are placed on people almost every day. Examples are becoming competent in the operation of a new machine - given that one knows how to operate the old one (Taylor, 1999). Learning potential has been shown to be a good indicator of performance in the South African work environment (De Beer, 2006; Gilmore, 2008; Keyser, 2012; Pelser, 2002; Taylor, 1999).

1.2 PROBLEM STATEMENT

Very few studies have investigated both psychomotor ability and learning potential as predictors of work performance of drivers and machine operators (Keyser, 2012; Pelser, 2002). The measurement of learning potential has, however, received considerable attention over the years (Caffrey, Fuchs, & Fuchs, 2008; De Beer, 2010; Grigorenko, 2009; Haywood, 2008; Lidz, 2009; Murphy & Maree, 2006). The concept of learning ability or learning potential has come into its own as a result of dissatisfaction with the outcomes of the geneticist and environmentalist approaches to intelligence testing (Gilmore, 2008). Instruments used to measure learning potential are regarded by some as being less culture biased (De Beer, 2010; Taylor, 1999) than current IQ tests, because they measure the potential for modifiability through learning, rather than an innate and immutable trait (Taylor, 1999). Learning potential is thus of particular relevance in developing countries where measurements for employee selection purposes are made especially difficult by virtue of high levels of illiteracy and vastly unequal living and development conditions (Taylor, 1999). The current climate of change and equal employment opportunity in the South African workplace has produced the need for selection instruments, which are less affected by a lack of education and an impoverished upbringing and can highlight undeveloped potential (Gilmore, 2008; Taylor, 1999). Such measures could, among other things, be used for more accurate selection of employees in terms of future work performance.

The difference between cognitive and psychomotor tests has caused several researchers to consider them unrelated to one another (Carretta & Ree, 1997;
Keyser, 2012). Essentially, there is a paucity of research studies conducted in the South African context on the effectiveness of these selection instruments.

This study was intended to benefit Industrial and Organisational Psychologists, Human Resources Practitioners and road construction companies in understanding the validity of psychomotor skills and learning potential as predictors of driver and operator performance. In this particular research context a definite gap exists as there is currently no such information available. The study was therefore conducted with the aim of addressing this gap. It was anticipated that the research findings may inform more accurate selection methods and procedures for the particular company involved and could inform the broader road construction industry.

In the road construction industry, there is growing competition between the various civil engineering contractors to secure tenders and maximise the profitability of road construction sites (Gabriel, 2011). To ensure this and to achieve upper limit profitability, the cost curve needs to decline. This translates into a need for a more productive workforce, working more efficiently. Valid selection processes are therefore required to ensure that the most productive individuals are selected for the most suitable jobs. In order to investigate whether the organisation’s preferred measurement instruments do significantly predict job performance, the following question for the research study was formulated:

1.2.1 Research Question and Hypotheses

Do psychomotor ability and learning potential statistically significantly predict the work performance of drivers and machine operators in a road construction company? Flowing from the background and problem identified the following research hypotheses were posed and tested empirically:

H1: Psychomotor ability and learning potential are statistically significantly and positively related to work performance.
H2: Psychomotor ability statistically significantly predicts work performance
H3: Learning potential statistically significantly predicts work performance
H4: Psychomotor ability and learning potential jointly statistically significantly predict work performance

1.3 RESEARCH AIMS

The general aim of the study was to investigate whether psychomotor ability and learning potential statistically significantly predict work performance of operators and drivers in a road construction company.

1.3.1 Specific Literature Aims

The specific literature aims of the research are to:
- Conceptualise psychomotor ability and learning potential
- Conceptualise job performance
- Investigate the theoretical relationship between psychomotor ability and learning potential in predicting job performance, with specific reference to drivers and operators.

1.3.2 Specific Empirical Aims

The specific empirical aims of the research are to:
- Investigate the statistical and practical significance of the relationship between psychomotor ability, learning potential and job performance
- Investigate whether psychomotor ability and learning potential statistically significantly predict work performance amongst driver and machine operators in a road construction company
- Formulate recommendations towards optimising the selection, training and development of drivers and machine operators as well as for future research.
1.4 PARADIGM PERSPECTIVE

This study was conducted from a humanistic paradigmatic stance. The stance emphasises an individual's inherent drive towards self-actualisation and creativity (Aanstoos, 2003; DeRobertis, 2013; Morgan, 1980; Pirson & Lawrence, 2010). The paradigm typically holds that people are inherently good and adopts a holistic approach to human existence and pays special attention to such phenomena as creativity, free will, and human potential (Aanstoos, 2003).

With specific reference to learning potential and dynamic assessment (De Beer, 2006), the TRAM assessment tool presents participants with an opportunity to change, adapt and develop (De Beer, 2000b; Keyser; 2012; Meyer, Moore & Viljoen, 1989; Pelser, 2002; Taylor, 1999). Similarly, knowledge and self-awareness of their psychomotor ability will allow drivers and machine operators to develop and more accurately define their competencies in terms of strengths and areas of development. The humanistic paradigm is therefore deemed appropriate for this study.

The humanistic psychology perspective is summarised by five core principles first articulated in an article written by James Bugental (1964) and adapted by Greening (2010) who was a psychologist and long-time editor of the Journal of Humanistic Psychology. The five basic principles of humanistic psychology are:

1. Human beings, as human, supersede the sum of their parts. They cannot be reduced to components.
2. Human beings exist in a uniquely human context, as well as in a cosmic ecology.
3. Human beings are aware and are aware of being aware - they are conscious. Human consciousness always includes an awareness of oneself in the context of other people.
4. Human beings have some choice and, with that, responsibility.
5. Human beings are intentional, aim at goals, are aware that they cause future events, and seek meaning, value, and creativity.

While humanistic psychology is a specific division within the American Psychological Association, humanistic psychology is not so much a discipline within psychology as a perspective on the human condition that informs psychological research and practice (Greening, 2010).

1.5 THE RESEARCH DESIGN

The design is presented according to the research approach and research methodology.

1.5.1 Research Approach

A quantitative approach (Howell, 2004; Terre Blanche, Durrheim & Painter, 2006; Welman, Kruger & Mitchell, 2005) was followed and a cross-sectional field survey design (Howell, 2004; Terre Blanche, et al., 2006; Welman et al., 2005) used. This approach was preferred because a quantitative research approach allowed the researcher to measure and analyse data statistically - especially the relationship between independent and dependent variables (Mouton & Marais, 1996; Terre Blanche, et al., 2006). In the current study, this is of particular importance because the research aims were determined and based on an investigation of the relationships between variables, or then the prediction of one dependent variable (driver and machine operator performance) by means of two independent variables (psychomotor ability and learning potential). In addition, the quantitative approach is advantageous because the researcher can be more objective about the findings of the research (Terre Blanche, et al., 2006). Quantitative data can be used to test the hypotheses of the current study using statistical analysis.

A cross-sectional field survey design involves the observation of all of a population, or a representative subset, at one specific point in time (Welman et al., 2005). The
aim is to provide data on the population under study. Data is collected to make inferences about a population of interest at the particular point in time (Howell, 2004; Terre Blanche et al., 2006; Welman et al., 2005). Cross-sectional surveys have been described as snapshots of the populations about which data is gathered (Welman et al., 2005). Cross-sectional studies are observational in nature and are known as descriptive research, not causal, as researchers typically record the information that is present in a population or sample, but they do not manipulate variables. This type of research can be used to describe characteristics that exist in a population, but not to determine cause-and-effect relationships between different variables (Welman et al., 2005). However, this method is often used to make inferences about possible relationships between variables - which was the case in the current study.

1.5.2 Unique Conditions of the Study

The parameters and unique setting of the study allowed for the evaluation of both predictive and concurrent validity. The company involved in the research conduct performance appraisals on an ongoing monthly basis. However, the psychological assessments were facilitated at a particular point in time and during the participants’ ongoing employment. This meant that criterion data could be collected at the same time as the assessment scores (concurrent data), or, the criterion data could be collected after some time has lapsed (predictive data) from the date of the assessments.

According to Cascio and Aguinis (2010), concurrent designs for obtaining evidence of criterion-related validity are useful to HR researchers in several ways. Criterion measures usually are substitutes for other more important, costly, or complex performance measures. However, predictive designs for obtaining evidence of criterion-related validity are the cornerstone of individual differences measurement (Cascio & Aguinis, 2010; Gregory, 2007). When the objective is to predict behaviour on the basis of scores on a predictor measure, there is simply no substitute for it (Aguinis, 2013; Cascio & Aguinis; 2010). Predictive studies demonstrate in an objective, statistical manner the relationship between predictors and criteria in a
particular context. For this reason, the researcher opted to evaluate predictive validity and not concurrent validity. However, the research conditions were unique in that it allowed for both concurrent and predictive validity because the assessments were done whilst the workers were already employed. Usually these types of assessments are done to select potential workers, i.e. before they commence work. In the present study, however, individuals were employed already, when the assessments were completed and only after a period of six months was criterion data collected - in order to determine if psychomotor ability and learning potential were statistically significantly and positively related to work performance.

1.5.3 Research Method

In adhering to the quantitative research approach, numerical data was obtained from the research sample. All drivers and machine operators working in the company were assessed on the Vienna Test System (Schuhfried, 1996) and the TRAM assessment battery (Taylor, 1999). Primary data collected included psychomotor ability (VTS) and learning potential (TRAM) assessment results, performance appraisals and supervisor rankings. The data was statistically analysed by making use of correlation coefficients and regression analyses to test the stated hypotheses and answer the research question.

1.5.4 Research Participants

The research organisation employs approximately 650 employees, depending on contract availability. One-hundred-and-thirty of the company’s permanent members of staff are dedicated drivers and machine operators. This section of employees constituted the target sample in terms of the current study. The research population constitute the drivers and machine operators in the wider road construction industry of South Africa. In this particular case, therefore, the target sample was a convenience sample.
At the onset of the study, the researcher was cognisant of the fact that the target sample may not be equal to the realised sample. Although a hundred percent response rate was the aim, some individuals were absent at the time of the assessments and a few others resigned from the company either before the research started or before the performance data was gathered.

1.5.5 Ethical Considerations

The entire research project was conducted in the light of clear ethical guidelines as prescribed by the Health Professions Council of South Africa (HPCSA). These guidelines include: the principle of best interest or well-being (non-maleficence and beneficence), the principle of respect for persons (autonomy and confidentiality) and the principle of justice.

More specific ethical considerations included the fact that written permission and authorisation were obtained from the company’s chief executive officer (CEO) to conduct the research and to access all the necessary data. All participants signed an informed consent form and no individual person could be identified in the final research document – all personal data was removed. Paper-based records were kept in a secure, locked location. Only the researcher and his supervisor, a registered Industrial Psychologist had access to paper-based records. Similarly, all computer-based records were stored on an external storage device and controlled by the use of access privileges and passwords. All identifiable data was kept secure. Personal identifiers were then removed from research-related information and only encrypted data used. However, all participants in the study were advised in different focus groups that confidentiality cannot be assured.

Prior to the research it was decided that data will not be kept for longer than two years and that all information will be properly disposed of once the time limit has been reached. All data sets and information pertaining to the research will be shredded and/or permanently deleted from the external storage device. Provisions were made to accommodate any form of distress that could possibly come as a result of the
study, although it was highly unlikely. None of the participants reported any form of discomfort, psychological side-effects, persecution, stigmatisation or negative labelling as a result of this study. Should anything out of the ordinary occur, even after the completion of the study, the matter will be referred to appropriate professionals; the University together with the researcher’s supervisor will also be notified immediately. The researcher undertook to disclose the final result of the study to the participants and together with the supervisor put time aside should any of the participants wish to consult on the matter.

1.6 MEASURING INSTRUMENTS

The measuring instruments used in the research are the Vienna Test System (VTS) (Schuhfried, 1996) to measure psychomotor ability and the TRAM learning potential test battery (Taylor, 1999). The research company’s performance appraisals for drivers and operators and supervisor rankings were used to measure the performance of drivers and machine operators.

1.6.1 Vienna Test System

The Vienna Test System (VTS) is a computer-assisted application of a large number of highly diverse psycho-diagnostic tests, measuring reaction times in tasks that require choosing among complex stimuli (Schuhfried, 1996). The following subtests are incorporated in the Vienna Test System.

1.6.1.1 Cognitrone - This subtest of the Vienna Test System assesses the candidate’s ability to concentrate and to adjust his/her work tempo to different stimuli patterns (Schuhfried, 1996; Schuhfried, 2000a). It was included because of its logical conceptual link with road construction drivers and machine operator performance. These drivers and operators are required to demonstrate sustained concentration throughout their shift, taking into consideration the demands that the continually changing operating environment places on them. In the subtest, candidates are required to indicate as fast, and as accurately as possible, whether any of four
figures presented on a computer screen is similar to the figure in the test question. The test yields various options in terms of differentiated results (Schuhfried, 1996; 2000a).

1.6.1.2 Determination Unit - This subtest assesses a candidate’s reaction speed, reactive stress tolerance and ability to demonstrate sustained multiple-choice reactions to rapidly changing stimuli (Schuhfried, 1996). Like the Cognitrone, this sub-test was administrated because of its conceptual links to road construction driver and operator requirements (Schuhfried, 1996; Schuhfried, 2000a). Its focus is on the operators' appropriate and fast responses in rapidly changing environments that may involve various stressors and stimuli (Schuhfried, 1996). The Determination Unit requires the discrimination of colours and acoustic signals, memorisation of the relevant characteristics of stimulus configurations and response buttons, and also memorisation and application of assignment rules (Schuhfried, 1996). Individuals have to react to differently coloured visual stimuli as well as acoustic stimuli that require either finger or foot responses. The test starts off slowly, gains speed to a very fast response requirement (approximating high stress situations, such as accident or near-accident situations) and then slows down marginally (approximating the period just after the accident/near accident) (Schuhfried, 1996). Schuhfried (1996) reported an internal consistency reliability of 0.99 for the Determination Unit. In various criterion-related validity studies significant correlations between results on the Determination Unit and driving performance criteria were obtained, for instance in a study with driving behaviour during a test drive as well as results of a driving test as the criteria (Schuhfried, 1996). In another example, driving safety criteria, frequency of accidents and driver errors were used as criterion measures (Schuhfried, 1996).

1.6.1.3 Two-hand Coordination Speed and Accuracy - This subtest assesses hand-eye and hand-hand coordination (Schuhfried, 1996; Schuhfried, 2000c). It was included as a predictor in this study because of the two hand coordination requirements of driving and operating activities. Candidates are required to move a cursor on a given track with the aid of two joysticks, one that can move forward and
backward and one that can move right and left (Schuhfried, 1996; 2000c). Candidates must therefore use both hands in a coordinated way to move the cursor along the track within acceptable accuracy limits (Schuhfried, 1996; Schuhfried, 2000c). The track consists of three sections varying in the demands made on the left and right hands. The scores yielded are total mean duration (the speed dimension) and total percentage error duration (the accuracy dimension). Reported internal consistency reliabilities of the measures varied from 0.85 to 0.97 (Schuhfried, 1996).

1.6.1.4 Time Movement/Anticipation Test (ZBA). The ZBA sub-test assesses an individual’s ability to imagine the effect of a movement and correctly estimate the movement of objects in space (Schuhfried, 1996; Schuhfried, 2000c). As the individual watches a ball move across the computer screen, the ball suddenly disappears and they are required to indicate when and at what position the ball would have crossed a line. Data is recorded on the time and position accuracy.

1.6.2 TRAM Assessment Battery

The TRAM assessment battery is a cognitive measure of the respondents’ overall learning potential (Taylor, 1999). The TRAM Learning Potential Test Battery (Taylor, 1999) was selected as a predictor in this study and was developed in South Africa by a South African, which enhances the overall face validity of the instrument. Essentially, the TRAM-I is a learning potential assessment instrument for candidates who fall in the illiterate and semi-literate ranges or who have had formal schooling lower than and up to Grade 7 (Taylor, 1999). The TRAM-II is intended for application to testees with education ranging from Grade 8 to Grade 12.

The TRAM assessment battery was included in this study as a culture-fair measure of learning potential, which also portrayed an indication of fluid intelligence ($g_f$) and general cognitive ability, or “$g$” (Jensen, 1986; Taylor, 1994). The test requires candidates to translate symbols into other symbols, using a dictionary. The symbols are pictorial or quasi-geometric (Taylor, 1999). The symbols are translated using some underlying rule (such as opposites – sun/moon; or the symbols being used
together – such as teacup/teapot). In Phase A1 of the test, candidates first complete the translation process by themselves. Thereafter they are given a lesson to explain the underlying rules, followed by the completion of Phase A2. Then they are given another test book and another dictionary to assess the transfer of skills. The final step is the completion of a memory test (Taylor, 1999).

Taylor (1999) explained that scores are provided on the following TRAM dimensions: Automatisation, Transfer, Memory and Understanding and Speed and Accuracy. Composite scores of respondents’ overall performance are also generated. Only the overall assessment rating was used as a predictor in this study. Taylor (1999) reported reliability coefficients ranging from 0.62 to 0.95 for the various dimensions. In terms of validity, Taylor (1999) found that composite scores on the TRAM correlated significantly \( r = 0.59; p= 0.01 \) with academic performance in an ABET course and also with academic performance \( r = 0.51; p= 0.01 \) in N1 studies (NQF level 2, or grade 10) (Taylor, 1999).

1.6.3 Performance Appraisals

Work performance was measured using the company’s performance appraisal system and the results of a separate supervisor ranking exercise. The company for which the research was conducted uses a top-down performance appraisal approach and all appraising managers received unit-standard aligned training on performance management as part of the organisation’s management development programme. Appraisals are done on an on-going monthly basis and the scores are linked to production bonuses which are paid out on a monthly basis. Appelbaum, Gilliland and Roy (2011) noted that adequate training must be provided to both the appraiser and the appraisee in order to avoid the many rating errors that are common in performance appraisals. Training should include cultural, legal and customer differences by providing managers with the tools to improve on the process. Managers must also be given the opportunity to build the required relationship with these employees (Appelbaum et al., 2011).
The company’s performance appraisal criteria included productivity, care for resources and a combined criterion called attitude and safety. It is a customised system designed specifically for the company - taking into account the particular needs within the industry. This customisation contributes to the system’s user-friendly interface (Sillup & Klimberg, 2010) and face validity among the company’s managers. Senior site management facilitated the driver and operator performance evaluations.

Performance appraisals are a foundation element of human resource management: the results of appraisals are used as the basis for many HR decisions (Brown, Hyatt & Benson, 2010). Performance appraisal is potentially a key tool for organisations to make the most of their human resources (Prowse & Prowse, 2009). One of the main advantages to a forced choice system is that it can help alleviate some of the more common rater errors that may occur when managers are rating employee performance, such as the severity error (i.e. when all employees are rated poorly) and leniency error (i.e. when all employees are rated high) (Steward, Gruys & Storm, 2010). A second advantage of this type of system is that its use insures that employees are evaluated using the same criteria so that the outcome of the process is more objective (Spector, 2012; Steward et al., 2010). It is better to have all employees evaluated using the same criteria, rather than having individual managers use their own criteria to rate and rank employees. Having managers use the same criteria brings more objectivity to the performance evaluation process (Steward et al., 2010). A third advantage to forced choice systems is that they facilitate more candid and open communication between managers and employees so that employees know where they stand and know what they need to do to improve (Steward et al., 2010) - which also increases the face validity and general acceptance of the appraisal system. A fourth advantage of a forced choice system is that it may help employers identify their top performers and allow for them to be rewarded (Guralnik & Wardi 2003; Spector, 2012; Steward et al., 2010). Performance ratings can be used to allocate merit increases (e.g., compensation budget is more accurately estimated) and identify appropriate employees for promotions. Distinctive pay rewards and differentiation of the level of rewards between excellent performers and
those who are performing less well can help a company hold onto their best and brightest (Guralnik & Wardi, 2003; Steward et al., 2010). Turnover of an organisation’s critical talent pool may be minimized since the employees that help move the organisation forward can be clearly identified, appropriately rewarded and retained. This is another argument for using a forced choice system (Guralnik & Wardi, 2003; Steward et al., 2010). Forced distribution systems can be an effective tool for eliminating poor performers and keeping employees on their toes (Spector, 2012; Steward et al., 2010).

O’Sullivan (2009) suggested three key choices to be made regarding how a given performance dimension should be measured. The first key choice has to do with the type of judgment (i.e., relative vs. absolute) that will be required of raters. The research company makes use of absolute judgments. These types of judgements are generally regarded more positively by employees, because they compare an employee’s performance to a performance standard, not to the performance of other employees (O’Sullivan, 2009). This can lead to more reliable results (Guralnik & Wardi, 2003; O’Sullivan, 2009; Spector, 2012; Steward et al., 2010), provided that those rating the employee actually use common performance standards (O’Sullivan, 2009). Secondly, a choice must be made about the measurement approach (O’Sullivan, 2009; Spector, 2012). The company uses a combination of behaviourally oriented and outcome-oriented approaches to measure performance dimensions. This dual-approach is generally viewed as having higher utility than other methods (O’Sullivan, 2009). The third key choice in measuring performance dimensions is the type of scale to be used. Here the company opted for ease of use because it facilitates performance evaluations on a monthly basis. Each performance criterion to be measured is represented by a scale on which a rater indicates the extent to which the employee possesses/demonstrates that characteristic (and this applies whether the characteristic is a trait, behaviour, or an outcome). Such scales are commonly used for their ease of implementation (Guralnik & Wardi, 2003; O’Sullivan, 2009; Spector, 2012).
1.6.4 Supervisor Rankings

To further ensure the accuracy of work performance information, the researcher obtained such data based on the facilitation of supervisor rankings (in addition to the performance appraisals). Forced ranking (FR) is a performance intervention, which can be defined as an evaluation method of forced distribution, where managers are required to distribute ratings for those being evaluated into a pre-specified performance distribution ranking (Cooper & Argyris, 1998). These rankings were done by the plant operations manager who oversees all drivers and operators, as opposed to the performance evaluations done by senior site management. The plant operations manager ranked each driver and operator, using the paired comparisons method (Cascio, 1998). Procedurally, he was asked to decide which operator in every pair of operators they would select if the working conditions were particularly difficult (due to congestion in the loading areas, wet road conditions, project completion time-constraints etc.) with number 1 being the best operator, number 2 the second best operator and so forth.

1.7 RESEARCH PROCEDURE

This study consisted of two distinct phases, namely an explorative literature review and an empirical study.

1.7.1 Phase 1: Literature Review

In the explorative literature review, the researcher endeavoured to determine the following:

1. The theory of psychomotor ability and the conceptualisation of psychomotor ability as a measurement construct.
2. The theory of learning potential and the conceptualisation of learning potential as a measurement construct.
3. The exploration of job performance and the measurement thereof.
1.7.2 Phase 2: Empirical Study

The empirical study involved a quantitative investigation into the statistical relationship between psychomotor ability and learning potential assessment results and job performance.

In step 1 of the empirical study, all drivers and machine operators were assessed on the Vienna Test System (Schuhfried, 1996) and the TRAM assessment battery (Taylor, 1999). In step 2, the researcher gathered the psychomotor ability and learning potential data necessary to conduct the empirical study. In step 3, data from the drivers and machine operator’s performance appraisal forms and supervisor rankings was collected from official company documentation. Notably, assessment data was gathered first and, six months thereafter, the criterion data was collected – as the study was concerned specifically with predictive validity and not with concurrent validity. In phase 4, the collected data was statistically analysed with computer software (SPSS ver. 20) and results interpreted.

1.8 STATISTICAL ANALYSIS

The study followed a descriptive approach (Terre Blanche et al., 2006) and was aimed at describing the relationships between the selected variables. Correlations (Terre Blanche et al., 2006) were used to report on whether, and to what extent, a statistically significantly (and positive) relationship between variables exist. Regression analysis (Terre Blanche et al., 2006) was used to report on whether the independent variables (psychomotor ability and learning potential) statistically significantly predicted the dependent variable (work performance). The correlation coefficient was used as a method to interpret the results and to specifically evaluate the statistical significance of the relationship (Howell, 2004; Terre Blanche et al., 2006) between psychomotor ability, learning potential, and work performance. Subsequent to determining the correlation coefficients, the validity of psychomotor ability and learning potential as predictors of work performance was evaluated by statistical regression.
Regression refers to the prediction of one dependent variable based on knowledge of the levels of one or more independent variable (Howell, 2004). Singular regression was used to respectively examine the predictive validity of psychomotor ability and learning potential in terms of work performance. Multiple regression was used to analyse the predictive validity of both psychomotor ability and learning potential for the work performance. Multiple regression is a method of studying the separate and collective contributions of several independent variables (psychomotor ability and learning potential) to the variation of a dependent variable (work performance) (Howell, 2004; Terre Blanche et al., 2006). The empirical study is explicated in chapter 3.

1.9 RESULTS

The data were analysed and the results are reported and presented, in Chapter 3. The relationship between the variables is interpreted. The statistical analyses indicate whether the relationship is statistically and practically significant. The statistical and practical significance of the various analyses is considered in the interpretation of the results. Practical significance refers to the effect size – which is related to the magnitude (or size) of the correlation (Cohen, 1992).

1.10 CHAPTER LAYOUT

The chapters are presented in the following manner.

Chapter 2 Literature review
Chapter 3 Article
Chapter 4 Conclusions, Limitations and Recommendations

1.11 CHAPTER SUMMARY

In Chapter 1 the scientific orientation to the research was discussed. This contained the background and motivation, the research problem, aims, the paradigm perspective and the research design. The chapter ended with the chapter layout.
CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

The business environment in which organisations have to operate has become increasingly complex (Chabault et al., 2012; Luthans, 2008; Meister & Willyard, 2010). Business organisations are faced with ever-increasing uncertainty, turbulence and changes in the external environment. These changes are due to increased international and local competition, technological advances and increased stakeholder and customer expectations, among other factors (Luthans, 2008; Meister & Willyard, 2010). To survive and remain profitable in this ever-changing external environment, management must use its awareness of these forces to improve its internal business operations and productivity (Chabault et al., 2012; Luthans, 2008; Meister & Willyard, 2010). To accomplish this, companies need to attract, select and retain effective and committed employees (Chabault et al., 2012; Collings, & Mellahi, 2009; Girard & Fallery, 2010; Mendes & Stander, 2011).

In order to select individuals who will contribute to the effectiveness of organisations, certain selection methodologies are employed. These selection methods need to remain within the scope of legislation to potentially select the most productive employees for an organisation (Gilmore, 2008; Schmidt & Hunter, 1998). Valid selection procedures should translate into productivity and improved performance levels, and ultimately, into the overall effectiveness of any organisation (Pulakos, 2005; Schmidt & Hunter, 1998). The ability to assess and identify employees who have the potential to learn new tasks, rather than only being able to demonstrate the skills they have learnt, should have an impact on the prediction of the manner in which employees ultimately perform in the workplace (Sternberg & Grigorenko, 2002). However, given the socio-economic and educational disadvantages that many individuals in South Africa did and still do experience, there is a need for
assessment strategies that focus on future potential rather than current ability (Murphy & Maree, 2006).

Against this backdrop, the study was aimed at investigating the utility of psychomotor ability and learning potential assessment results as predictors of work performance. The aim of the literature review is to investigate and conceptualise these concepts within the road construction environment.

2.2 PSYCHOMOTOR ABILITY

The need for psychomotor skills as a driver or machine operator is apparent. High performing drivers and operators demonstrate practical evidence of having learnt complex sequences of actions that require use of perceptual information and control of specific muscles.

2.2.1 Defining Psychomotor Ability

Psychomotor behaviours are performed actions that are neuromuscular in nature and demand certain levels of physical movement and dexterity (De Kock & Schlechter, 2009; Jacobs, MacKenzie & Botma, 2013; Weckowicz & Liebel-Weckowicz, 1990). The term psychomotor denotes a combination of physical and psychological activities (Plug, Meyer, Louw & Gouws, 1989) and involves the process of receiving sensory messages from the environment and then producing a response thereto (Jacobs et al., 2013; Singer, 1972). To be proficient, sensory information needs to be organised and processed (Ackerman, 1988) to enable a person to produce an appropriate motor, or movement response (Guilford, 1956; Gregory, 2003; Hergenhahn, 2009; Reynolds & Adams, 1953).

According to Sullivan (2009) psychomotor ability refers to the capacity to physically manipulate objects in the environment by making use of available perceptual information. Psychomotor performance abilities typically include two types of component skills: production of motor actions and recognition of environmental
conditions that trigger these actions (Hergenhahn, 2009; Sullivan, 2009). Production and recognition skills are often intertwined in a seamless cycle of adaptive action that appears effortless when observed in an expert performer (Fadde, 2007). Therefore, for the purpose of this study, the concept specifically refers to the relationship between cognitive functions and physical movement and is demonstrated by physical skills such as movement, coordination, manipulation, dexterity, strength, speed and fine motor skills - such as those needed for precision instruments and operating complex machinery (Fernández-Ballesteros, 2003; Gibb & Dolgin, 1989; Keyser, 2012; Pelser, 2002).

For Johnston and Catano (2002), psychomotor abilities include skills such as hand-eye coordination, balance, and reaction time that arise from a unity of cognitive and physical functions. The concept can further be defined as the process of interaction between the perceptual systems (or five senses), the brain (where perceptual information is interpreted) and the body (where the individual reacts to such perceptual stimuli) (Mohan, Srivastava & Srivastava, 1984; Tan, 2006; Vesia Esposito, Prime & Klavora, 2008). Tan (2006) explained that the word ‘psycho’ refers to the mind or psyche, and the word ‘motor’ to the physiological body. More generally ‘psychomotor’ can thus be seen as the mind-body interaction, and ‘psychomotor abilities’, as those capacities which allow for effective interaction between the two and the environment (Johnston & Catano, 2002; Stuart, 2013; Tan, 2006).

### 2.2.2 The Historical Development of Psychomotor Ability

The first theorist to coin the term ‘psychomotor’ was Carl Wernicke (1848-1905). Wernicke worked on the inability (or impaired ability) to understand or produce speech, as a result of brain damage (Fernández-Ballesteros, 2003; Gregory, 2007; Hergenhahn, 2009). Through his studies Wernicke demonstrated how human functioning can be explained by psycho-sensory input, psychomotor output, and the intra-psychic functions which coordinate the two (Fernández-Ballesteros, 2003; Weckowicz & Liebel-Weckowicz, 1990). In doing so, Wernicke explained how
humans interact with their environment, by making sense of it, and interpreting such sensory information in an integrated manner (Bergh, Pelser & Visser, 2005; Hergenhahn, 2009).

With the development of new psychomotor theories, a need arose to measure psychomotor abilities. Theorists and researchers such as Wilhelm Wundt (1832 – 1920), Francis Galton (1822 – 1911), Clark Wissler (1870 – 1947), and James McKeen Cattell (1860 – 1944) contributed greatly in this (Gregory, 2007; Hergenhahn, 2009).

Wundt was responsible for the development of numerous psychological instruments, most of which can be considered psychomotor instruments (Fernández-Ballesteros, 2003; Gregory, 2007; Hergenhahn, 2009). The thought meter, a device composed of swinging pendulums and calibration needles was one such instrument. Wundt would measure the degree to which the individual was capable of anticipating the movements of the pendulum and then compared these results with the real calibrations. By contrasting the observed and real positions of the pendulum, Wundt measured a basic type of mental processing speed (Gregory, 2007).

Galton, considered the father of psychological testing, developed thousands of measurement instruments (Fernández-Ballesteros, 2003; Gregory, 2007). Some of the psychomotor assessments Galton developed included the tint discrimination instrument, which determined to what degree an individual could discriminate between visual tints of light, and reaction time instruments, which measured the delay between the presentation of a stimulus and reaction to it. Galton proposed that mental speed was greatly related to intelligence, a construct not well understood at the time (Fernández-Ballesteros, 2003; Gregory, 2007; Hergenhahn, 2009).

Cattell was also fascinated by a possible correlation between mental speed and intelligence. He eventually teamed up with Wundt, and together they assessed thousands of individuals for their reaction speed in relation to different stimuli (Fernández-Ballesteros, 2003). Cattell discovered that certain people have small, but
consistent differences in their reaction speed. Consequently he proposed that there were individual differences in the processing speed of human beings (Fernández-Ballesteros, 2003; Gregory, 2007; Hergenhahn, 2009).

With the advent of psychomotor theory and the development of a better understanding of job-related requirements, testing and assessment of the psychomotor abilities for job-fit started to become prominent (Bergh et al., 2005; Carretta & Ree, 2000; De Kock & Schlechter 2009; Jacobs et al., 2013; Jelovsek, Kow & Diwadkar, 2013; Vesia et al., 2008). Many tests were developed to measure the domain and its possible impact on job-related requirements from the First and Second World War to the late 1970’s (Bergh et al., 2005; Hergenhahn, 2009). Aspects such as two-hand coordination, visual processing, strength, stamina, reaction time, integrated perception, auditory reaction, leg strength, speech formation, concentration, and intelligence were all measured within this domain of psychomotor ability (Bergh et al., 2005; Fleishman & Hempel, 1954; Hall, Echt, Wolf, & Rogers, 2011). However, a dedicated domain structure or factor structure for psychomotor abilities had not yet been developed during this time. The development of such a structure was imperative – this is because a domain structure for psychomotor abilities (or skills) would unify the seemingly separate aspects of psychomotor abilities and as a result lend validity to the measurement and use of such constructs (Bergh et al., 2005; Johnston & Catano, 2002; Mohan et al., 1984; Reynolds & Adams, 1953; Vesia et al., 2008).

The two theorists most notably responsible for the establishment of this domain or field of study were Joy Guilford (1897 – 1987) and Edwin Fleishman (1927 – present) (Fernández-Ballesteros, 2003; Gregory, 2007). Guilford found that psychomotor abilities shared much of the variance of ‘pure’ intelligence constructs such as ‘g’ (Gregory, 2007). It was theorised that perception and processing speed were related significantly to both general intelligence and psychomotor ability (Gregory, 2007). Fleishman also worked on numerous psychomotor measurement instruments (Bergh et al., 2005). He compared these measurement tools with regard to their validity, how well they measured a specific psychomotor ability, and then
grouped those abilities which inter-correlated (Fleishman & Hempel, 1954). Fleishman and Guilford’s work validated each other, both describing similar psychomotor factors (Singer, 1972).

Modern psychomotor tests and assessments are generally computer based, and make use of advanced sensory equipment and simulation exercises (the Vienna Test System is one example). However, most of the constructs they measure can still be traced to the psychomotor domains of Guilford and Fleishman. In today’s era, psychomotor assessments are much more advanced. Schuhfried, for instance, has developed the Vienna Test System (VTS), an assessment system, which has integrated numerous forms of psychometric tests in a computerised format (Bergh et al., 2005). Recent developments of various VTS tools can now be successfully applied to areas as diverse as clinical psychology, personnel selection and development, aviation psychology, traffic psychology and sport psychology. This is made possible through the symbiosis of hardware, software and scientifically validated theory in the field of psychology (Bergh et al., 2005).

Many psychomotor tests proposed by early theorists have now become virtual (Vorster, 2012) and although these assessments are much more advanced in comparison to their early counterparts, they still measure similar constructs (Pelser, 2002).

2.2.3 The Relationship between Psychomotor Ability and Cognitive Ability

Keyser (2012) and Bergh et al. (2005) noted that general cognitive ability and psychomotor ability were frequently regarded as independent. However, these and other researchers explained that numerous recent studies have found a statistically significant positive relationship between cognitive and psychomotor ability (Bedell, Van Eeden & Van Staden, 1999; Bergh et al., 2005; Carretta & Ree, 1997; Chaiken, Kyllonen & Tirre, 2000; Fernández-Ballesteros, 2003; Hunter & Hunter, 1984; Jacobs, et al., 2013; McHenry, Hough, Toquam, Hanson & Ashworth, 1990; Rabbitt, Banejeri & Scymnaski, 1989; Ree & Carretta, 1994; Tirre & Raouf, 1998; Vesia et al.,
hence it is concluded that these factors are, in fact, not independent. 

Research done by Ree and Carretta in 1994 was specifically aimed at explaining this relationship between cognitive ability and psychomotor ability (Bergh et al., 2005). They administered multiple aptitude and psychomotor test batteries to 354 United States Air Force recruits. The researchers found that the average multiple correlation of the cognitive tests and each psychomotor score as a criterion was 0.34 (Bergh et al., 2005; Ree & Carretta, 1994).

According to Keyser (2012), individuals' psychomotor performance and cognitive abilities vary from person to person and can be affected by certain factors. These abilities can also slow down over time due to factors such as aging. In their study on aging, motivation and the positivity effect on attention and memory, Mather and Carstensen (2005) found that cognitive capacities in older age groups were lower than in younger age groups.

It is generally agreed that cognitive abilities have a hierarchical structure with general cognitive ability, $g$, at the top and lower-order common factors such as verbal, mathematics and spatial test scores at the lowest level (Bergh et al., 2005; Caretta & Ree, 1997; Keyser, 2012). The hierarchical model demonstrates that $g$ accounts for a major portion, frequently more than half, of the variance of the lower-order common factors and test scores. Caretta and Ree (1997) also found that the hierarchical structure did not differ across gender and racial/ethnic groups.

Various studies focusing on the direct correlation between psychomotor tests and general (cognitive) ability tests have found significant correlations (Bergh et al., 2005; Rabbit et al., 1989, Tirre & Raouf, 1998). In addition, different factor analytical studies have pointed to a relationship between psychomotor ability and general (cognitive) ability (Carretta & Ree, 1997; Chaiken et al., 2000; Ree & Carretta, 1994). Bergh et al. (2005) concluded that the literature points to three possible explanations for the correlation between psychomotor and intelligence measures.
Firstly, performance on both measures requires the ability to reason. In this regard, Ree and Carretta (1994) suggested that psychomotor and intelligence factors correlate due to the fact that performing both types of tests require a certain amount of reasoning. In a later publication these writers went so far as to suggest that the measurement of $g$ is actually unavoidable in all measures of ability (Carretta & Ree, 1996a). They reasoned that when responding to test material, regardless of whether the test requires a psychomotor response, specialised knowledge or verbal skills, reasoning is unavoidable and consequently results in $g$ being measured (Carretta & Ree, 1996a). Jensen (1980) also suggested that most ability tests would be reduced to irrelevant and unpractical exercises if $g$ is factored out of the equation.

Secondly, performance on both measures requires a certain degree of learning to take place (Bergh et al., 2005; Jacobs et al., 2013). The correlation between $g$ and psychomotor ability may therefore be a reflection of the fact that a certain degree of learning is required in order to perform well in psychomotor tests (Carretta & Ree, 1996b). Chaiken et al. (2000) suggested that this correlation is due to the fact that subjects differ in terms of working memory capacity ($g$), and that this difference impacts on learning complex and novel tasks such as those that are often involved in psychomotor tests.

Thirdly, measures such as information processing speed, working memory capacity, and reaction time, which often underlie good performance in many psychomotor tests, have also been identified as measures of cognitive ability (Jensen, 1982; Jensen, 1993; Kranzler & Jensen, 1991; Kyllonen & Christal, 1990; Miller & Vernon, 1992), hence, it should not be surprising that statistically significant positive correlations are found. In his research Jensen (1993) found that tasks which require a simple reaction time show only low positive correlation with $g$. However, more complex tasks where a choice between alternatives is required show moderate correlation with $g$ (Bergh et al., 2005; Jensen, 1993). A statistically significant positive correlation between general (cognitive) ability or $g$ and psychomotor ability is therefore to be expected (Bergh et al., 2005).
Like cognitive ability, psychomotor ability has been studied for about a century (Bergh et al., 2005; Carretta & Ree, 1997; Keyser, 2012). Fleishman and Quaintance (1984) identified eleven conceptually separate domains of psychomotor performance. Historically, psychomotor abilities have been seen as lower order factors not influenced by a higher order factor (Cronbach, 1970; Fleishman, 1964). Ree and Carretta (1994) examined the relationship of a limited battery of cognitive tests and psychomotor tracking tests. They found both lower-order and higher-order cognitive and lower-order and higher-order psychomotor inter-correlations. However, their cognitive tests were limited to verbal and mathematical abilities and their psychomotor tests to tracking tasks only (Bergh et al., 2005; Carretta & Ree, 1997).

Cognitive tests and psychomotor tests bear little superficial similarity (Carretta & Ree, 1997). Cognitive tests require answering questions on an answer sheet while psychomotor tests are usually computer-administered and use control sticks, a computer pointing device (or mouse), and foot pedals (Pelser, 2002). The dissimilarity between cognitive and psychomotor tests has caused several researchers to consider them unrelated to one another (Carretta & Ree, 1997; Pelser, 2002). In this regard, Tun and Luchman (2010) studied the association between computer use and cognition across adulthood. Their findings offer new insight into the association between computer activity and cognition across the lifespan in adults (Keyser 2012). Tun and Luchman (2010) found that frequent computer use is associated with better overall cognitive performance across adulthood, from younger adults through middle-aged and older adults. Furthermore, they found a positive association between computer use and executive function that was seen even after controlling for basic intellectual ability. Specifically, more frequent computer use was associated with better task-switching performance and this association was strongest in adults with lower general cognitive ability (Keyser 2012).
2.2.4 Psychomotor Ability as a Predictor of Job Performance

Psychomotor abilities, skills, and constructs are measured for a number of practical purposes, including the understanding of neurological deficits, appropriate perception-stimuli interactions, safety, intelligence, emotional wellbeing, stress tolerance (Vorster, 2012) and employment selection (Keyser, 2012; Pelser, 2002). Research has indicated that psychomotor ability is an important indicator of performance, especially for pilots, drivers and machine operators (Bergh et al., 2005; Carretta & Ree, 1996b; Fernández-Ballesteros, 2003; Johnston & Catano, 2002; Keyser, 2012; Mohan et al., 1984; Reynolds & Adams, 1953; Vesia et al., 2008).

Psychomotor tests were used in the armed forces from as far back as the early 1900’s (Bergh et al., 2005). Research in this area started during the First World War, as part of a general drive to improve pilot selection techniques in the United States of America (Carretta & Ree, 1996b; De Kock & Schlechter, 2009; Flotman, 2003; Martinussen, 1996; Mnguni, 2011; Ree & Carretta, 1996a). Early in the war, candidates were selected for flight training on a volunteer basis. When casualty reports started coming in, it was noted that accidents were not always due to equipment error or enemy action and that many casualties were actually due to human error (Bergh et al., 2005; Griffin & Koonce, 1996). Human error, concern for safety, high pilot training costs and a growing interest in scientific selection methods justified the use of psychomotor tests, especially in aviation (De Kock & Schlechter, 2009; Duke & Ree, 1996; Griffin & Koonce, 1996; Martinussen, 1996; 2; Ree & Carretta, 1998).

Early studies indicated that psychomotor tests had considerable validity for predicting both Air Force and Navy pass or fail rates (Bergh et al., 2005) but also with regards to the success of pilots in training programmes before and during World War II (Bergh et al., 2005; Fleischman, 1988). During World War II, research in the psychomotor field gained momentum (Bergh et al., 2005; Duke & Ree, 1996, Griffin & Koonce, 1996) and numerous psychomotor tests were developed during this time with encouraging validities for various criteria (Bergh et al., 2005; Duke & Ree, 1996).
in the USA, Canada (Bergh et al., 2005; Duke & Ree, 1996), the United Kingdom (Bergh et al., 2005; Duke & Ree, 1996), and Germany (Bergh et al., 2005; Duke & Ree, 1996).

With the advent of the use of computer technology for the assessment of psychomotor skills, interest in this form of assessment was again revived (Duke & Ree, 1996; Fernández-Ballesteros, 2003; Fleischman; 1988; Griffin & Koonce, 1996; Pelser, 2002; Ree & Carretta, 1998; Vesia et al., 2008). The development and rapid distribution of computers made it possible to assess abilities that were not previously possible with paper and pencil tests (Bergh et al., 2005). Some examples of these newly assessable abilities include divided attention, the ability to concentrate and the ability to function under pressure (Fernández-Ballesteros, 2003; Fleischman; 1988; Pelser, 2002). This development was mainly due to the dynamic capabilities of the personal computer, which lends itself better to the assessment of constructs such as perceptual speed, spatial visualisation and reaction time (Bergh et al., 2005; Vorster, 2012). Further advantages are the improvement of the reliability of the presentation of the tests as well as the accuracy of measurement and data collection, and the reduction in the equipment failure that is typical of older apparatus tests (Fernández-Ballesteros, 2003; Griffin & Koonce, 1996; Maguire, Smith, Brallier, & Palm, 2010; Piaw, 2012; Pelser, 2002; Ree & Carretta, 1998; Thurlow, Lazarus, Albus & Hodgson, 2010; Tippins, 2009).

By far the greatest amount of research in the use of psychomotor ability as predictor in selection exercises within the last three decades has centred on pilot selection in the military context (Bergh et al., 2005; De Kock & Schlechter, 2009; Keyser, 2012). This research has been mostly conducted in the United States of America. However, in a South African validity and utility study for the selection of train drivers using the Vienna Test System as predictor and simulator operating performance as criterion, Schoeman (1995) reported a validity coefficient of 0.50, which is reported to increase to 0.70 once corrected for shrinkage and 0.87 once corrected for unreliability of the criterion (Schoeman, 1995). These coefficients seem high. Pelser (2002) noted that, bearing in mind the small sample (N=62), these results need to be interpreted with
caution. Bouwer’s (1984) study (also with a small sample, N=58), investigating the differences between heavy duty vehicle drivers who had been divided into strong and weak performing groups, reported that in terms of the psychomotor tests used, significant differences were found between the two groups (Bergh et al., 2005). However, no significant differences were found between the two groups in terms of their performance on the apparatus tests designed to measure psychomotor skills, more specifically reaction time, two hand coordination, and information processing (Bergh et al., 2005; Bouwer, 1984).

2.3 LEARNING POTENTIAL

The hallmark of a learning potential measure is that testees learn a new skill or competency in the process of doing the tasks. Some individuals become more competent than others; it is these differences that are captured in the scores of a learning potential test or battery (Taylor, 1999).

2.3.1 An Introduction to Intelligence

According to Carretta and Ree (2000) the measurement and structure of abilities have been a topic of speculation and study since the time of Aristotle, who distinguished ability from emotional and moral faculty. Within the field of personnel selection, intelligence has been studied scientifically for more than a century and is probably the most researched predictor in the field (Bergh et al., 2005; Gilmore 2008; Keyser; 2012; Salgado, 2000, Schmidt & Hunter, 1998). In addition, intelligence has been credited with being the most valid predictor of job performance in numerous studies (Hunter & Hunter, 1984; Ree & Carretta, 1996a; Ree, & Earles, 1992; Schmidt & Hunter, 1998).

While intelligence is one of the most talked about subjects within psychology, there is no standard definition of what exactly constitutes intelligence. Some researchers have suggested that intelligence is a single, general ability; while other believe that intelligence encompasses a range of aptitudes, skills and talents (Bergh et al., 2005). The theories on intelligence differ markedly in terms of the nature of the construct,
definitions, assumptions and the measures used to assess the construct (Suzuki, Meller & Ponteretto, 1996). Taylor (1994) identified three main schools of thought on the nature of intelligence and the methodology of its assessment, namely the structural approach, the information processing approach and the learning or dynamic approach. These are used as a framework for discussing cognition as it relates to this study.

2.3.2 Intelligence Theories

The study and measurement of intelligence has been an important research topic for nearly 100 years. Much of the excitement among investigators in the field of intelligence derives from their attempts to determine exactly what intelligence is. Different investigators have emphasized different aspects of intelligence in their definitions. While there has been considerable debate over the exact nature of intelligence, no definitive conceptualization has emerged. Today, psychologists often account for the many different theoretical viewpoints when discussing intelligence and acknowledge that this debate is ongoing.

2.3.2.1 The Structural Approach to Intelligence

The structural approach is also sometimes referred to as the individual differences approach and was the first school of thought on the nature of intelligence and its implications for assessment (Bergh et al., 2005; Sternberg, Kaufman, & Grigorenko, 2008). In this approach, performance is measured along those dimensions which are contended to form the fundamental structure of the construct, intelligence. According to Taylor (1994), in structuralism, extensive use is made of correlational and factor analytical techniques to resolve theoretical and empirical questions about intelligence. The major theoretical assumptions in this approach are based on sensory responses theory (Galton & Cattell), intelligence quotient theory (Binet & Simon), Spearman’s two-factor theory, primary mental abilities theory (Thurstone) and the hierarchical theories (Taylor, 1994).
Sensory responses theory - Galton (1822-1911) was the first to propose the concept of general mental ability – a sphere in which he proposed that individuals differ widely (Walsh & Betz, 1990). Galton contended that since all information that reaches the individual, does so through the senses, the differences in general mental ability could be ascribed to differences in sensory function, hence Galton defined intellect as the sum of the simple component parts of sensory functioning (Taylor, 1994; Walsh & Betz, 1990). Cattell (1860-1944) joined Galton (1822-1911) and together they developed basic measures for sensory capacity, which Cattell referred to as mental tests (Bergh et al., 2005). This involved the measurement of psychomotor ability (encompassing measures such as reaction time) and perception (involving, among other things, the measurement of the ability of candidates to perceive differences in size and colour) (Bergh et al., 2005; Sternberg et al., 2008). Galton and Cattell’s contribution introduced the world to the concept of general mental ability (GMA).

Intelligence quotient theory - By the turn of the century, research started indicating that the scores on the various mental tests from the sensory response theory did not correlate sufficiently with one another to be measuring one concept (Bergh et al., 2005). Furthermore the scores did not seem to correlate with criteria with which they could logically be expected to correlate such as teachers’ ratings and school results (Bergh et al., 2005; Sternberg et al., 2008; Walsh & Betz, 1990). As the popularity of the sensory response theory diminished, Binet (1857-1911) and Simon (1873-1961) started developing their theory of intelligence, which was based on the higher mental processes of judgment and reasoning as opposed to the lower order sensory-motor capabilities of the Galton-Cattell approach (Fernández-Ballesteros, 2003). Over and above their emphasis on judgment and reasoning, Binet and Simon made an important contribution in terms of a hypothesis on the process of the development of these so called higher mental processes. They proposed that the capacity to demonstrate higher mental processes would increase as a child increased in age and used this hypothesis as the foundation of their intelligence test (Gregory, 2007). Their theory forms the base of the Stanford-Binet intelligence scale, which yields an intelligence quotient or, as it is more popularly known, an IQ score (Fernández-
Ballesteros, 2003; Gregory, 2007). IQ as an expression of intelligence is one of the most widely used intelligence measures in the world (Fernández-Ballesteros, 2003; Gregory, 2007; Walsh & Betz, 1990).

**Spearman’s g and s factors** - Spearman (1904) is credited with being the great pioneer in the development of scientific methodology in the intelligence research field (Bergh et al., 2005; Jensen, 1986). He devised factor analysis, which made it possible to study the factors that make up intelligence. He confirmed Galton’s proposition of the existence of a general mental ability component and proposed that all cognitive tests had a general g and several specific components (Bergh et al., 2005). According to Spearman (1904), g was a component of all cognitive tests, but other specific abilities were test unique (Carretta & Ree, 2000). The general ability referred to by Spearman became much narrower than the general ability referred to by Galton (Bergh et al., 2005). Galton referred to general ability in relatively broad terms – essentially in biological and evolutionary terms (Bergh et al., 2005). Spearman derived his conception of g exclusively from factor analysis (Jensen, 1986). This gave rise to one of the biggest debates in the field of psychometrics, namely whether g is merely a methodological artefact or a consequence of the mathematical manipulation (using factor analysis) of inter-correlations between various tests, or whether it reflects a real construct, independent of psychometric tests and factor analysis (Jensen, 1986).

**Multiple aptitude theories** - Since many theorists felt that there were factors of mental ability somewhere between the generality of g and the uniqueness of the specific factors, they developed theories in which intelligence was postulated to be constituted of a number of group factors (Bergh et al., 2005; Jensen, 1986; Taylor 1994). Psychologist Louis L. Thurstone (1887-1955) offered a different theory of intelligence. Instead of viewing intelligence as a single, general ability, Thurstone’s theory focused on seven different primary mental abilities. The abilities that he described were: verbal comprehension, reasoning, perceptual speed, numerical ability, word fluency, associative memory and spatial visualization (Carretta & Ree, 2000). This led to the development of several ability taxonomies and many multiple
aptitude batteries (Bergh et al., 2005; Carretta & Ree, 2000, Taylor, 1994) such as the General Aptitude Test Battery (GATB) (Carretta & Ree, 2000). Pelser (2002) explained that a major contribution of this school of thought was its ability to identify the pattern of abilities for specific individuals. Candidates with equal amounts of $g$ could, for instance, differ in terms of their particular strengths and weaknesses as far as verbal and numerical ability are concerned (Bergh et al., 2005; Taylor, 1994). This, in turn, could be relevant to the individual’s success or non-success in various intellectual pursuits some of which could relate to job performance (Bergh et al., 2005). Taylor (1994) contended that most test constructions are based on the Thurstonian model. In the South African context, there are numerous tests that measure verbal ability, numerical ability and mechanical ability, such as the Senior Aptitude Test (SAT) and the Intermediate Battery (Bergh et al., 2005). These are strongly related to Thurstone’s Primary Mental Abilities (Taylor, 1994).

Hierarchical theories - The advent of multiple aptitude theories was later shadowed by research findings supporting a hierarchical structure of intelligence (Bergh et al., 2005; Taylor, 1994). Hierarchical theories imply one or more higher order scores with several lower order scores (Carretta & Ree, 1996a; Carretta & Ree, 2000; Gustafsson, 1993; Ree & Carretta, 1994; Taylor, 1994; Vernon, 1969). To a certain extent, the advent of the hierarchical theories can be seen as a move back to Spearman’s model, but with more specific focus being placed on the hierarchical nature of the relationship between $g$ and the specific factors and the dynamics between $g$, specific factors and the prediction of criteria (Bergh et al., 2005).

Vernon’s description of different levels of intelligence may fill the gaps between two extreme theories of the two-factor theory of Spearman, which did not allow for the existence of group factors, and the multiple-factor theory of Turstone, which did not allow for a $g$ factor (Bergh et al., 2005; Fernández-Ballesteros, 2003; Sternberg et al., 2008; Taylor, 1994; Walsh & Betz, 1990). According to Vernon, intelligence can be described as comprising abilities at varying levels of generality: Firstly, the highest level: $g$ (general intelligence) factor with the largest source of variance between individuals. The next level was the major group factors such as verbal-
numerical-educational and practical-mechanical-spatial-physical ability. (Bergh et al., 2005). Thirdly, minor group factors are divided from major group factors. Vernon's fourth level represents the “s”, or specific factor (Bergh et al., 2005; Fernández-Ballesteros, 2003; Sternberg et al., 2008; Taylor, 1994; Walsh & Betz, 1990). Beginning in 1969, Vernon became increasingly involved in studying the contributions of environmental and genetic factors to intellectual development (Sternberg et al., 2008). Vernon continued to analyse the effects of genes and the environment on both individual and group differences in intelligence. He concluded that individual differences in intelligence are approximately 60 percent attributable to genetic factors, and that there is some evidence implicating genes in racial group differences in average levels of mental ability (Sternberg et al., 2008; Taylor, 1994).

According to Taylor (1994), Cattell’s (1971) theory, which is also a hierarchical theory, is one of the structural theories that is better established theoretically because it also deals with the nature rather than just the structure of intelligence. This theory of intelligence consists of two main components, namely fluid and crystallized intelligence (Bergh et al., 2005; Sternberg et al., 2008; Taylor, 1994). Fluid intelligence is defined as the inherited capacity which is developed by interaction with the environment (which is not culture-specific) and can be seen as a purer content-free reflection of reasoning ability (Taylor, 1994; Walsh & Betz, 1990). Crystallized intelligence is defined as the specialized skill and knowledge required and promoted by a specific culture and is related to acquired knowledge (Taylor, 1994). Whilst crystallized intelligence is usually assessed with tests with informational content which draws on previously acquired knowledge (Vocabulary and numerical ability for instance), fluid intelligence is usually assessed by tests with as little as possible informational content, assessing the ability to see relationships, for example pattern or series completion (Pelser, 2002; Walsh & Betz, 1990).

Although crystallized and fluid intelligence have been shown to be highly correlated and often indistinguishable in test populations that are homogenous in terms of culture and education levels (Walsh & Betz, 1990), the implication is that this correlation is not necessarily prevalent in test populations where there is
heterogeneity in these variables (Bergh et al., 2005; Budoff, 1987; Laughton, 1990; Taylor, 1994). For this reason, the argument that this theory presents is that the assessment of fluid intelligence will give a clearer picture of intellectual potential, particularly in culturally and educationally diverse test populations (Taylor, 1994). This educational diversity is expected to be relevant in a country such as South Africa where educational opportunities were assigned on the basis of race for a great part of its history (Taylor, 1994).

Cattell’s culture-fair tests were developed to measure fluid intelligence (Pelser, 2002). Test items are presented in abstract-diagrammatical form and involve universal activities such as odd-one-out, pattern completion, series completion and identification of conceptual relationships (Bergh et al., 2005; Pelser, 2002). There is evidence for the culture-fairness of this theory (Taylor, 1994). Cattell’s model also forms the foundation of many other of the learning/dynamic theories applied today (Pelser, 2002).

Currently the trend in the research in the structural approach to intelligence supports the hierarchical model, postulating that the higher order factor, general ability, or $g$, is at the apex, followed by lower order factors such as verbal, mathematics and spatial ability, followed by test scores at the lowest level (Carretta & Ree, 1996a, 1997, 2000, Ree & Carretta, 1994). However, Bergh et al. (2005) explained that the structural theories have several weaknesses: (1) their concepts ($g$, for example) are dependent on factor analysis and are data-driven at the expense of being based on theory (Jensen, 1986); (2) other than Cattell’s theory, they are relatively theory-poor (Taylor, 1994); and (3) the risk of possible cultural bias inherent in tests based on the structural theories (Jensen, 1986; Taylor 1994). These weaknesses led to the investigation of other theories and measures of intelligence. More particularly the information processing and learning/dynamic approaches were developed (Taylor, 1994). Both these approaches focus on the nature (rather than the structure) of intelligence and on the development of the theory that underpins the concept (Taylor, 1994).
2.3.2.2 The Information Processing Approach

The information processing approach had its advent in the 1960’s as access to the processing capacity of computers became more readily available to researchers (Taylor, 1994). It can be argued that this school of thought is the ultimate operationalisation of Galton’s sensory responses theory in that it sees the human being as a general-purpose information processor that obtains information from, and produces output to the environment (Taylor, 1994). In this way people’s relationship to the outside world is formed and maintained (Sternberg, 1984). There are two main schools of thought in this approach, namely the limited capacity theory of cognitive competence and the cognitive components approach (Taylor, 1994).

Limited capacity theory of cognitive competence - The limited capacity theory of cognitive competence holds that the human information processing system contains one or more bottlenecks which limit the flow of information, and that individuals who are able to process information faster in these bottlenecks are also more competent at problem solving and other real-life tasks (Taylor, 1994). According to this theory, the measurement of the receiving, processing and retrieval speeds of information links to intelligence. Bergh et al. (2005) noted that, due to the nature of the stimulus material (diagrammatical objects on a computer screen), the measurement will probably be relatively free from the impact of prior knowledge and other environmental variables. There is evidence for working memory to correlate well with fluid intelligence measures (Baddeley, 1986; Bergh et al., 2005; Larson & Saccuzzo, 1989).

2.3.2.3 Cognitive Components Approach

The major theories in the cognitive components approach, according to Taylor (1994) are:

- automatisation theory (Sternberg, 1984);
- radex theory (Taylor, 1994);
• the circular cognitive model (Snow, Kyllonen & Marshalek, 1984); and
• the cyllindrical cognitive model (Ackerman, 1988).

Automatisation theory (Sternberg, 1984) holds that there are two main cognitive processes fundamental to intelligence, namely the response to novelty and automatisation. The way in which individuals respond to novelty, the process of mastering it, automatising it and moving towards efficiency, which, in turn, frees them to apply their mind to more novel tasks, is seen to be indicative of intelligence (Taylor, 1994).

Radex theory (Taylor, 1994) holds that intelligence can be seen to be reflected by a radex with more complex tasks in the centre and less complex ones on the circumference. The actual placement of the tests along the circumference of the circular space will depend on the content of the test (e.g. verbal, numerical or spatial) (Bergh et al., 2005).

The circular cognitive model is an extension of Radex theory (Taylor, 1994). Snow et al. (1984) found in a study based on this model, novel rule-induction items are consistently shown to be more complex and g-saturated. Test content seems to have no relevance in terms of g-loadings (Bergh et al., 2005; Taylor, 2004). These g-loaded abilities are shown in the centre of the radex clustering around fluid intelligence with specific, more specialized rule applying activities plotted on the external boundaries of the radex (Bergh et al., 2005; Taylor, 2004).

The cyllindrical cognitive model (Ackerman, 1988) expands even further on Snow et al’s (1983) model, by adding the notion of speed on a vertical plane, effectively transforming the circle to a cylinder. Hence the generality-specificity dimension is represented by the horizontal cross-section of the cylinder, while the power-speed dimension is demonstrated by the vertical dimension (Bergh et al., 2005; Taylor, 2004). It is argued that as the movement outwards from the core and downwards on the vertical plain progresses, previously acquired skills and knowledge play an ever greater role in the acquisition of new skills (Bergh et al., 2005; Taylor, 2004).
has obvious implications for fairness in South Africa where, as mentioned earlier, there has been great disparity in the provision and quality of education (Taylor, 2004).

Although the information processing approach theory is much more defined and its procedures are so simple that it seems as though there should be little bias, little cross-cultural research has been conducted to confirm this empirically (Bergh et al., 2005). Furthermore, despite having shown promising correlations with measures of fluid intelligence (Jensen, 1982, Vernon, 1990), little research has indicated correlations with criteria indicating differential performance in the real world (Taylor, 1994).

2.3.2.4 The Learning/Dynamic Approach

According to the structural approach (and to a lesser extent the information processing approach), intelligence is stable and possibly inborn, in that people seem to differ in terms of the intelligence they have been allocated (Sternberg et al., 2008). Of greater importance to this study, the learning (or dynamic) approach seems to lend itself particularly well to cross-cultural assessment (Taylor, 1994) and because of its centrality to the current research; this approach will be discussed in detail and also linked to the South African Context.

The learning approach focuses on the capacity of people to adapt to the demands of the environment (Resnick & Neches, 1984; Vo et al., 2011), an approach that seems to be more aligned with the humanistic paradigm (Aanstoos, 2003; Bugental, 1964; DeRobertis, 2013; Greening, 2010; Pirson & Lawrence, 2010).

Vygotsky, a Russian psychologist, viewed the development of cognitive ability as a social phenomenon and defined the concept zone of proximal development (ZPD) as acknowledgement of individuals' differing ability to profit from mediated learning (Gilmore, 2008; Vygotsky, 1978). Vygotsky viewed the development of cognitive ability as a social phenomenon. Adults and older peers transfer the knowledge and skills required in a culture (inter-personal thinking), and in this way assist with

According to Vygotsky (1978) there are two distinct levels when describing learning and development. The first is the learning that has already taken place, while the second can be achieved with assistance or guidance (Vygotsky, 1978). The difference between potential and actual developmental levels is called the zone of proximal development, or ZPD (Bergh et al., 2005; Brown & French, 1979; Frisby & Braden, 1992; Kozulin, Gindis, Ageyev, & Miller, 2003; Thompson, 2013; Vygotsky, 1978). It refers to the difference between what a person can do independently and what he or she is capable of doing with specifically tailored guidance (Kozulin et al., 2003; Thompson, 2013).

In Vygotsky’s view, the intellectual environment can determine children’s intellectual growth (Vygotsky, 1978). This implies that the number and quality of the learning opportunities play an important part in the child’s cognitive development (Thompson, 2013; Vygotsky, 1978). In his famous analogy, the theorist compared the measuring of a person’s potential to a gardener judging the usefulness of a fruit tree. He argued that the gardener should not only look at the ripened apples but also at the maturing tree and consider all other aspects of the tree (Vygotsky, 1978).

Vygotsky argued that learning and instruction result in cognitive development (Boeyens, 1989). It was Vygotsky’s view (1978) that individuals differ in the capacity to benefit from mediated learning. He termed the distance between the two distinct levels as the zone of proximal development and described it as follows. According to Vygotsky (1978), this zone defines those functions that have not yet matured but are in the process of maturation, functions that will mature tomorrow but are currently in an embryonic state. In terms of his analogy, these functions could be viewed as the buds or flowers of development rather than the fruits of development. The actual developmental level characterises mental development retrospectively, while the zone of proximal development characterizes mental development prospectively (Vygotsky’s, 1978). The application of this concept therefore allows the test facilitator
or researcher to measure the potential level of reasoning a person may be able to reach (De Beer, 2000a; Vygotsky, 1978).

As mentioned, Sternberg (1984) contended that intelligence comprises of an individual’s ability to respond effectively to novelty, and to automatise. The learning approach focuses on this adaptation of individuals to the completion of novel tasks – as demonstrated by increased speed and accuracy as a result of repetition, instructions, examples and learning interventions (Taylor, 1994). To a certain extent the learning approach tells the researcher more than just what the individual’s cognitive capacity is – there is a degree of diagnosis inherent in the assessment (Sternberg, 1984; Taylor, 1994). There is a definite effort to understand the mechanics of the individual’s cognitive capability and how it operates in terms of the learning processes (Taylor, 1994). This, in turn, improves the understanding of how the individual operates cognitively in certain environments and also allows for individual diagnosis of problem areas, which opens possibilities for focused intervention (Bransford, Delclos, Vye, Burns & Hasselbring, 1987; Vo et al., 2011).

Early studies in this field reported modest correlations between learning and \( g \) (Taylor, 1994). Taylor (1994), however, contended that this is largely due to the fact that the learning tasks used in the learning potential measures were very simple. Snow et al. (1984) produced results that led to the conclusion that the more complex the learning tasks in the learning potential measures are, the higher the correlations found with \( g \). There are several approaches to the assessment of learning potential. Laughton (1990) identified three main approaches namely:

- Budoff (1968, 1974, 1987);
- Feuerstein (1980); and
- Campion, Brown, Ferrara, Jones and Steinberg (1985).

Budoff (1987) used a Ravens-type task to assess subjects ranging across the intelligence spectrum. This was followed by practice and task-specific training and the assessment was then repeated. Feuerstein (as cited in Laughton, 1990) used a similar process, but the intervention between the pre- and post-assessments
involved the development of thinking skills. Campion et al. (as cited in Laughton, 1990) focused on the transfer of learning and assessed transfer as the inverse function of the number of hints required to complete a novel task once the basic principles have been acquired doing other similar novel tasks.

When measuring learning potential, performance is assessed initially and a specific score assigned to the outcome (Gilmore, 2008; Schaap & Luwes, 2013). This first score represents the initial unassisted problem-solving attempt and is commonly known as the pre-test score (Gilmore, 2008; Keyser 2012; Taylor, 1999). A training intervention is then facilitated and the candidate is offered a second attempt at solving similar problems – a score is generated, named the post-test score (De Beer, 2006). The pre-test provides information on the current (actual) level of performance (De Beer, 2006; Keyser 2012) whereas the post-test provides information on the potential future level of performance. According to De Beer (2006), learning potential is a combination of the pre-test and post-test scores, as well as factoring in the magnitude of the difference between these two scores.

The intervention reduces the unfamiliarity of problem solving and therefore helps those individuals from disadvantaged backgrounds to better understand the appropriate problem solving strategies (Babad & Budoff, 1974). The mediated learning that occurs, aims at providing hints and guidelines that will assist the examinee to solve similar problems (De Beer, 2006).

In general, potential is defined as a person’s unrevealed capacity and it is argued that this innate capacity or ability is probably greater than the person’s manifest level of functioning (Feuerstein, Feuerstein & Gross, 1997). Learning potential assessment measures individuals’ present levels of ability as well as their potential for improvement with training interventions (Haywood & Tzuriel, 1992). Many influential theorists like Thorndike defined intelligence, or linked the concept of intelligence to the capacity to learn (Guthke, 1993; Guthke & Stein, 1996; Humphreys, 1985, Kaufman et al., 2010; Sternberg & Grigorenko, 2002;) and since 1980 there has been increasing interest in the role of intelligence in learning (Ekinci,
Learning potential is therefore concerned with what could be developed and is based on the premise that ability (that which is available on demand, namely already developed) can change over time (De Beer, 2000a).

2.3.2.5 Taylor’s (1994) Integrated Theory

Taylor’s (1994) theory integrated all three approaches mentioned above. He disagreed with the statement that processing speed and capacity is the complete foundation of intelligence (Jensen, 1982; Vernon, 1986), but contended that it is one of the two fundamental components. The other component is the ability to infer rules or abstract thinking – a concept analogous to Cattell’s fluid intelligence. Taylor contended that both of these fundamental abilities are biologically or genetically determined and set upper bounds to performance in the cognitive sphere (Pelser 2002). The learning or dynamic approach is integrated into his theory in that intelligence is hypothesized to be a product of learning and other interactions with the environment (Taylor, 1994; 1999).

Taylor (1994) used Ackerman’s cylindricial cognitive model (1988) to explain the basics of his integrated theory. He suggested that the focus should be on learning potential (with the focus on fluid intelligence or general ability, or ‘g’ - which lies in the core of the cylinder) and not on learning performance (crystallized intelligence or specific abilities which are encountered when moving towards the periphery of the cylinder). Hence, the inclusion of novel tasks where the stimulus material is unfamiliar to all testees (and previously acquired skills will be of little assistance in the mastery of the tasks) is essential (Pelser, 2002; Taylor, 1994; 1999).

Taylor (1994) furthermore emphasized that transfer has long been recognized as a primary component of learning, the rationale being that an individual needs to be able to take what he or she has learnt in one context and apply it in a different context, since the circumstances of the new challenge are very unlikely to be identical to previous problems or tasks.
Taylor (1994) contended that although it is unwise to do one-on-one mapping from cognitive abilities to learning processes, his two fundamental components of intelligence relate to learning processes in the following ways: abstract thinking capacity (fluid intelligence) is related to transfer and processing efficiency (information processing) is related to automatisation. Therefore, Taylor proposed that intelligence tests should tap the following four domains: fluid intelligence; information processing; transfer in tasks requiring learning potential; and automatisation in such tasks (Taylor, 1994; 1999).

De Goede (2007) investigated Taylor’s reason for including or combining the structural -, information processing and cognitive components approaches. He contended that Taylor’s (1999) integration stems from an observation that the psychological tests that are widely available for use in industry and education are mostly designed to measure broad-based static psychological constructs such as abilities. De Goede (2007) explained that it seems as if the two new approaches of information processing and learning and modifiability tend not to be widely used in industry, despite a need for assessment techniques of a more dynamic nature.

One of the reasons that could be contributing to this lack of usage might be the fact that the information processing and learning and modifiability approaches do not offer much in the form of practical measurement instruments, which can be used for selection or vocational guidance (Taylor, 1994). Thus, through his work Taylor (1992, 1994) attempted to relate his theoretical concepts to the mainstream of cognitive psychology while developing an assessment instrument that is suitable for practical application in an industrial or educational context (De Goede, 2007).

2.4 COGNITIVE ASSESSMENT IN THE SOUTH AFRICAN CONTEXT

South Africa is a country with complex and unique challenges that affect the practice of psychological assessment and usage of cognitive assessment. Psychological testing in South Africa cannot be investigated in isolation without taking the country’s political, economic, and social history into account (Claassen, 1997; Meiring, Van De Vijver, Rothmann & Barrick, 2005). After the first democratic elections in 1994, a new constitution was required to protect the cultural appropriateness of psychological tests and culminated in the decree of the Employment Equity Act (EE Act 55 of 1998, Section 8), which stipulates that psychological testing and other similar assessments are prohibited unless the test or assessment being used (a) has been scientifically shown to be valid and reliable, (b) can be applied fairly to all employees; and (c) is not biased against any employee or group (Government Gazette, 1998). Since 2012 a further clause was added to the Act, which states that only psychometric tests that have been certified by the Health Professions Council of South Africa, or another body which is authorised to certify such tests, may be used in tests on employees. Psychologists practising in South Africa therefore have the responsibility to only use fair and valid assessment tools to adhere to the EE Act, to ensure that no candidate of any race or creed is disadvantaged.

Assessments are used in many areas of business such as hiring, managing, restructuring, promoting, coaching, training, succession planning, counselling and ongoing development (Vaiman, Scullion & Collings, 2012). However, cognitive assessment in particular is used more and more for selection rather than training purposes (Gregory, 2007; Mouton & Marais, 1996; Vaiman et al., 2012).

After nearly a century of intelligence testing, the assessments we use have hardly changed, and our conceptions of intelligence bear a remarkable resemblance to the early theories (Schweizer & Neubauer, 2012). Nowadays, there are measures of cognitive and intellectual styles, levels of complexity, emotional intelligence, and many types of intelligence (Gregory, 2007). However, none of these has completely revolutionised our perspective on the actual definition of intelligence (Gregory, 2007;
Schweizer & Neubauer, 2012). It will be interesting to see whether technologies bring new ways of thinking about this prickly construct, and bring new methods of tapping into the human psyche.

Cognitive tests can be classified in terms of their functional characteristics into power tests and speed tests. In a speed test a person’s score is driven by the speed with which they are able to respond to a relatively simple task (Burke, 2009). In power tests a person’s score is driven by the level of difficulty at which they are able to operate (Burke, 2009).

Cognitive ability tests, and particularly power tests of reasoning, are the most consistent and strongest predictors of job performance (Burke, 2009). There is a range of various cognitive tools available nationally and internationally. These represent a variety of test methodologies, including the following (Prinsloo, 2013):

- Knowledge-based tests largely indicate learning exposure within specific content domains.
- Questionnaires often rely on the subject’s self-insight and are notoriously invalid. Results on cognitive styles are often inferred from personality-like tests.
- Unstructured or semi-structured interviews and structured interviews have various methodological weaknesses including relying on the skill and insight of the interviewer, the rapport between interviewer and interviewee and the verbal competence of the interviewee.
- Subjectivity and halo effects may bias the results of assessment centre methodologies, particularly where the observed behaviours are not clearly operationalised and inter-rater reliability issues are not addressed.
- IQ tests and figural analysis tests only measure a small aspect of thinking and are culturally loaded. Results are generally distorted by time limitations, as speed and power are separate constructs in cognitive assessment.
Simulation exercises and trainability tests that are reflective of job specific problem solving skills, seem to have better predictive validity, especially when well designed.

Hunt and Sternberg (2006) noted that intelligence is a culturally-relative concept and is thus the degree to which a person successfully adapts to cognitive tasks that are valued by members of that culture. According to Parker, Philp, Sarai and Rauf (2007), cultural, language and educational differences could contribute to lower the validity of cognitive assessment. In addition, political and ethical concerns could also have an impact on cognitive assessment (Claassen, 1997; Pretorius et al., 2009). Cultural differences could account for a large proportion of the discrepancies in cognitive assessment (Pretorius et al., 2009). South Africa, in particular, is a multicultural country which emphasises the importance of considering the influence of culture on the application and interpretation of cognitive assessment measures (Claassen, 1997; Pretorius et al., 2009).

Cognitive assessment yields relevant information used within different contexts and for a wide variety of purposes (Griffin & Christie, 2008). In the work environment cognitive assessments are a cost effective way of obtaining useful information such as an individual’s abilities and traits which are difficult to gather during an interview or reference check (Gregory, 2007; Griffin & Christie, 2008). As such, these assessments provide relevant samples of behaviour allowing organisations to make a more accurate judgement of ability (Gregory, 2007; Griffin & Christie, 2008).

Because of the diverse nature of our society, many external and internal factors surrounding the cognitive assessment can impact the eventual score – especially in the South African context (Claassen, 1997; Meiring, Pretorius et al., 2009; Schoeman, De Beer & Visser, 2008; Meiring et al., 2005), for example the perception of the testing process, self-doubt, and unfamiliarity with test taking, culturally confusing questions and the role of the examiner in the test situation may have an effect on performance (Schoeman et al., 2008). Prinsloo (2013) indicated that cognitive assessment should thus accommodate more than the linear, convergent,
logical-analytical thinking skills that are measured by IQ tests. Although the latter is important in certain contexts, it only represents a very small aspect of cognitive behaviour. Of the many factors that need to be considered in the construction and usage of assessment techniques, two are of critical importance:

- Cross-cultural applicability, given the cultural diversity of business and education;
- Accessing learning potential and adaptational flexibility, given educational and socio-economic disparities and fast changing technological and work requirements.

### 2.4.1 Learning Potential and Dynamic Assessment in South Africa

People differ in their capacity to learn (De Beer, 2006; Taylor, 1994), some learning faster than others, some mastering more difficult material than others. These differences can be measured using appropriate psychometric tools. Taylor (1999) explained that most conventional psychometric tools measure specific competencies rather than learning potential. There is a distinct difference between these competencies and learning potential. The former is an assessment of what has been acquired in the past by way of mastery of a particular cognitive domain (such as math) whereas the latter is a measure of what the person could achieve in the future with regard to mastering new material (De Beer, 2006; Taylor, 1994; 1999). Hence, the former measures the past, but the latter measures the future (Taylor, 1994; 1999).

According to Taylor (1999) and De Beer (2006) the most important reason for measuring learning potential is the fact that it assesses what the person could achieve in the future with regard to dealing effectively with cognitively challenging material, whereas tests of specific abilities assess the person’s current mastery of certain cognitive domains, which reflects the opportunities to learn that the person had in the past (Taylor, 1994; 1999). It is common knowledge that, in South Africa, the playing ground was not level in the past and it is not even properly level now Taylor (1999). Some children go to schools that are well equipped with all facilities,
including computers, and have excellent teachers teaching them in small classes, and others go to schools that are poorly equipped and are taught in large classes by teachers who are themselves poorly educated (Taylor, 1994; 1999). Education has not reached a level where we can say all individuals have an opportunity to develop their full potential. Nevertheless, the assessment of learning potential is still a preferable way to evaluate a person’s likely effectiveness in cognitively challenging environments like a new job, for example (Taylor, 1999).

Schaap and Luwes (2013) explained that learning potential, in addition to current cognitive abilities, is increasingly used in South Africa. Learning potential assessments measure individuals’ current levels of ability, as well as their potential for improvement if they are given suitable assistance. These assessments focus on existing and improved levels of functioning to evaluate a person's capacity for gaining new skills or knowledge when training is provided (De Beer 2005; Schaap & Luwes, 2013). Learning potential indicates the level to which an individual is likely to be able to learn and master cognitively demanding new challenges of an occupational or educational nature (De Beer, 2005; Taylor, 1999). It highlights fundamental capabilities that power the development of skill and knowledge (Taylor, 1994).

In an effort to conduct more equitable cognitive assessments, non-verbal reasoning assessment has received increasing attention in the past few decades, both in South Africa and internationally (Murphy, 2006). Non-verbal reasoning assessment has been intensively researched since the 1960s and 1970s and can be considered a fundamental element of learning potential tests (Schaap & Luwes, 2013). Research initiatives have focused, first, on providing more culture-fair assessment – this would be useful in comparing results obtained in culturally diverse populations; secondly, on designing measures appropriate for testing individuals with disadvantaged educational experiences and, lastly, on measuring learning potential as distinct from what has been learned – regardless of the culture, population, or social group of the individuals being tested (De Beer, 2005). Such research initiatives have contributed significantly to the development of the learning potential tests that form part of
selection batteries for the diverse South African society. The Transfer, Automisation and Memory tests (TRAM1/2), Ability of Processing Information and Learning Battery (APIL-B) and LPCAT tests are well-known, non-verbal-based learning potential tests developed in South Africa and used in industry for selection, placement and development purposes (Foxcroft & Roodt, 2009).

Dynamic assessment has shown considerable variance among individuals (Taylor, 1994). Furthermore, it has been found that candidates who perform poorly on static tests are inclined to perform considerably better on dynamic tests (Budoff, 1987; Laughton, 1990). The implication seems to be that this is the case due to the reality that dynamic assessment minimizes cultural bias. This is seen to be the main advantage of the learning potential or dynamic approach to assessing intelligence (Budoff, 1987; Taylor, 1994; 1999). It is suggested that the minimization of cultural bias occurs due to tests focusing on learning tasks which are unfamiliar to all the candidates regardless of their cultures (Sternberg, 1984; Taylor, 1994). Furthermore, improvement scores are bound to cancel out any bias that remains, because each candidate is assessed against his or her own baseline performance (Taylor, 1994).

2.5 DYNAMIC ASSESSMENT

Dynamic assessment results contribute to a more in depth understanding of both present and potential performance that is not readily obtainable from other sources.

2.5.1 Dynamic Assessment and Intelligence

Dynamic assessment is a more modern approach to assessing intelligence and learning potential in particular (Deadrick & Madigan, 1990; Murphy, 2011; Zollezzi, 1995). It is an interactive approach to conducting assessments and focuses on the learner’s ability to respond to intervention (Haywood & Lidz, 2007).

In essence, dynamic assessment is a theory of learning and a theory about how individuals acquire and express knowledge (Benjamin & Lomofsky, 2002; Murphy,
There are variations on several dimensions of the dynamic assessment model. However, there are several salient characteristics and assumptions that can be identified (Caffrey, Fuchs & Fuchs, 2008; Deadrick & Madigan, 1990; De Beer, 2006; Frisby & Braden, 1992; Grigorenko, 2009; Haywood, 2008; Haywood & Lidz, 2007; Lidz, 2009; Murphy, 2006; Zollezzi, 1995).

- Some abilities that are important for learning (in particular) are not assessed by normative, standardised intelligence tests.
- The assessor or assessment process provides active intervention during the course of the assessment with the learner with the goal of intentionally inducing changes in the learner's current level of independent functioning.
- Most people typically function at less than their intellectual capacity.
- The assessment focuses on the learner's processes of problem solving, including those that promote, as well as obstruct successful learning.
- The most unique information from the assessment is information about the learner's responsiveness to intervention.
- Teaching within the test provides a useful way of assessing potential as opposed to performance.
- Observing new learning is more useful than cataloguing (presumed) products of old learning. History is a necessary, but not sufficient predictor of future performance.
- The assessment also provides information about what interventions successfully promote change in the learner (connecting assessment with intervention).
- The assessment is most often administered in a pre-test-intervention-post-test format.
- The assessment is most useful when used for individual diagnosis, but can also be used for screening of classroom sized groups.
- Many conditions that do not reflect intellectual potential can and do interfere with one's expression of intelligence.

Intelligence is measured distinctly as either an intelligence test score or as learning potential test scores (Deadrick & Madigan, 1990; Hamers & Resing, 1993). This
distinction essentially refers to traditional static tests and dynamic tests. According to De Beer (2006), learning potential is concerned with an overall cognitive capacity, including current and projected future performance, which implies that intelligence is changeable when mediated. The interest in learning potential assessment was partly sparked by the criticism of traditional intelligence tests (Hamers & Resing, 1993; Taylor, 1994; Taylor & Richards, 1990) and this interest proved to be of particular importance to the South African assessment context (Taylor, 1999).

As mentioned, a significant criticism of traditional static intelligence tests is that not everybody has the same exposure or opportunities to learn or to gather the necessary knowledge and skills to perform as expected on the standardised tests (Claassen, 1997; Hamers & Resing, 1993; Meiring et al., 2005; Pretorius et al., 2009). Biesheuvel (1973) maintained that the traditional, or static, intelligence tests measure only learnt skill, whereas the adaptability, or dynamic tests assess what individuals can learn to do.

According to Van De Vijver (1993), traditional intelligence tests have three main areas of criticism. Firstly, the test-taker's verbal abilities are assumed. Secondly, the test items may contain (often unintended) implicit references to the cultural background of (usually) the test composer. Lastly, test-wiseness, which refers to subsidiary skills that are essential to the problem solving process (e.g. the skill to handle multiple-choice items or time limits in speed tests, etc.), are more often tested than cognitive ability per se (Van De Vijver, 1993).

De Beer (2000a) maintained that intelligence is reflected in the actual test scores of static tests and learning potential can only be measured if learning opportunities are provided as part of the assessment process – hence the term dynamic assessment. This notion is of particular interest in the South African context. In South Africa, development opportunities for the various cultures have historically been vastly different (Taylor, 1999). According to De Beer (2000a), socio-economic differences have an impact on the performance of South Africans being assessed on traditional (static) intelligence tests. This researcher found numerous discrepancies in the
various South African cultures, taking into account socio-economic and educational indicators. Assessing an individual’s potential to learn would therefore be a fairer application of cognitive assessments, especially in multi-cultural societies similar to South Africa (Taylor, 1999).

2.5.2 A Comparison between Traditional and Dynamic Cognitive Assessments

The aim of traditional assessment is to measure the actual or current performance of the testee at a particular point in time. The aim of dynamic assessment is to obtain a more realistic measurement of a person’s cognitive ability by including training as part of the assessment process (Hamers & Resing, 1993). Dynamic testing has all the psychometric properties of a static test but involves the use of standardised objective intervention or training procedures (Carlson & Wiedl, 1979; De Beer, 2006; Hamers, Hessells & Pennings, 1996; Hamers & Resing, 1993; Sternberg & Grigorenko, 2002).

There are three major differences between static and dynamic testing (Sternberg & Grigorenko, 2002): Firstly, static testing measures products formed from developed skills and knowledge, whereas dynamic testing taps into developing processes quantified through the capacity to learn. Secondly, in static testing, the testee responds to a graded sequence of problems, and no feedback is given. In dynamic testing, either implicit or explicit feedback is given on the quality of performance on items. Thirdly, the relationship between examiner and examinee should be neutral and uninvolved in static testing, while in dynamic testing, the relationship is interactive with information and mediated learning occurring in the relationship (Sternberg & Grigorenko, 2002).

There are various approaches to dynamic testing, but two common formats emerge: The first approach sees instruction sandwiched between a pre-test and a post-test (for individuals or groups). In the second approach instruction is given in response to the individual’s solution to each test item (only individuals) (Sternberg & Grigorenko, 2002). Because of its history and multicultural context, dynamic testing tends to be
attractive in South Africa (Gilmore, 2008; Keyser 2012, Taylor, 1999) as the individual historical differences can be reduced through repeated contact with the material in a teaching and supportive environment (Babad & Budoff, 1974).

2.5.3 Dynamic Assessment in South Africa

According to De Beer (2000a), Gilmore (2008) and Pelser (2002), both Feuerstein and Budoff’s approaches have been investigated successfully in South Africa. Budoff’s approach of administering a standardised measure in a dynamic test-train-test way, was researched by Shochet (1986) and Zolezzi (1995). Zollezi’s (1995) research concluded that traditional tests, which are applied dynamically, yield a better predictor of academic performance than static cognitive measures (Bergh et al., 2005; Gilmore, 2008). The South African LPCAT is similar to Budoff’s test but was specifically designed for dynamic assessment (Bergh et al., 2005; De Beer, 2000b; Gilmore, 2008). Both Budoff’s approach and De Beer’s LPCAT use nonverbal tasks to determine reasoning abilities.

Feuerstein’s approach has also been investigated in South Africa, research being performed by Van Niekerk in 1991 and Shochet in 1986 (as cited in Bergh et al., 2005). Sibaya, Hlongwane and Makunga (1996) also investigated the application of the Learning Potential Assessment Device (LPAD) for the assessment of giftedness, intelligence and other cognitive abilities. On the basis of Shochet’s research, training-enriched testing is supposed to help in the selection of disadvantaged students. However, Van Niekerk could not provide evidence that mediation improved performance on verbal or nonverbal reasoning, perceptual speed, mathematical applications, vocabulary or study habits (Bergh et al., 2005; Gilmore, 2008).

A few dynamic tests have been developed and standardised in South Africa, namely: (1) the Ability Processing of Information and Learning Battery (APIL-B), developed by Taylor (Taylor, 1997; 2003), (2) the Transfer, Automisation, Memory and Understanding Learning Potential Battery (TRAM-1 and TRAM-2) developed by Taylor (Aprolab, 1998; Taylor, 1999) and (3) the Learning Potential Computerised
Adaptive Test (LPCAT) developed by De Beer (2000b). The battery developed by Taylor (viz. TRAM-II, TRAM-II and APIL-B) is a pen-and-paper application. These assessments are time-consuming, whereas the LPCAT is computer based and the time taken to complete the assessment is roughly 30 to 45 minutes. The marking of the LPCAT is automatic and eradicates marker error. The company, in which the research is being conducted, however makes use of the TRAM battery.

2.6 JOB PERFORMANCE AND PERFORMANCE MANAGEMENT

Job performance refers to the work related activities expected of an employee and how well those activities are executed. Many businesses assess the job performance of each employee in order to help them identify suggested areas for improvement.

2.6.1 Introduction

The survival of organisations depends on the performance of their workforce. The greatest cost to most companies is labour which directly affects the profit margin. It is in organisations’ best interest to optimise their human capital in order to obtain greater gains in terms of this margin. However, encouraging employees to perform effectively, measuring and managing their job performance remains a concern.

The purpose of this section is to provide a literature review focusing on conceptualising job performance, by first describing the South African challenges in the changing work environment. Job performance is then defined and its measurement discussed. Finally, job performance management, and the various approaches to it, are highlighted and the approach to job performance in the research organisation discussed.
2.6.2 Conceptualising Job Performance

Jensen (1986) noted that intelligence is a general factor that runs through all types of performance.

2.6.2.1 A Definition of Job Performance

The word “performance” is used frequently in companies, among industrial psychologists and other human resource practitioners (Shamsi, 2010). Regardless of its importance for the company or unpopularity with employees, the concept is seldom clearly defined (Neely, Gregory & Platts, 1995). Viswesvaran, Ones and Schmidt, (1996) explained that, until now, no clear consensus exists on what exactly constitutes individual work performance. It is however becoming an increasingly important topic in the world of work (Meister & Willyard, 2010).

Job performance is usually described in terms of observable and non-observable behaviours which can be appraised or evaluated (Cascio, & Aguinis, 2010). It is generally defined as measureable actions, behaviour and outcomes that employees engage in or bring about that are linked with and contribute to organisational goals (Viswesvaran, et al., 1996). It is a multidimensional concept which describes how a worker completes a task, focusing on efficiency, skills used, initiative and utilised resources (Rothmann & Coetzer, 2003). Job performance is an action that involves a work process and contributes to the final output (Aguinis, 2013; Viswesvaran, et al., 1996). This individual process, or processes, can in turn be influenced by an organisation’s overall performance (Van der Linde, 2005). However, it is not only actions that determine an employee’s level of performance but also external factors such as resources, organisational culture and economic, political and social factors (Meister & Willyard, 2010; Studer, 2008; Van der Linde, 2005).

According to Viswesvaran (1993) job performance consists of 10 dimensions and, depending on the nature of the job, certain dimensions are more important than others. Those dimensions mostly overlap with the eight dimensions highlighted by
Campbell, McCloy, Oppler and Sager (1993) which are generally well supported in the literature. The dimensions are as follows:

Table 1.1

Job Performance Dimensions

10 Job Performance Dimensions of Viswesvaran (1993)

- Overall job performance
- Productivity
- Communication
- Effort
- Job-related knowledge
- Interpersonal skills
- Quality
- Leadership
- Rule following
- Administrative skills

8 Job Performance Dimensions of Campbell et al. (1993)

- Job-specific task proficiency
- Non-job-specific task proficiency
- Written and oral communication
- Demonstrating effort
- Maintaining personal discipline
- Facilitating peer and team performance
- Supervision/leadership
- Management/administrative

According to Ivancevich, Konopaske and Matteson (2005) the following three main factors impact on performance: the willingness to perform (motivation), opportunity (organisational factors such as resources and tasks) and the capacity to perform (skills, abilities and knowledge). Grote (2002) explained work performance from a behavioural perspective. By concentrating on work-related observable activities, one
is able to evaluate job performance more objectively, thus eliminating rating bias (Brewster, Carey, Doling, Grobler, Holland & Wärnich, 2003). Good performance is only achievable in an environment in which, the outputs and criteria are clearly defined (Mullins, 1999; Hayward, 2008). Individual performance can be defined as actions and behaviours individuals carry out which are linked to company goals (Campbell et al., 1993). In order for any company to reach its goals and strategic objectives, individual performance needs to be managed effectively and efficiently (Amos, Ristow & Ristow, 2004). It is through the individuals (their attitudes and actions) that the company is able to achieve a competitive edge (Sutherland, De Bruin & Crous, 2007). Individual performance is in fact the most critical component impacting companies’ successes (Meihem, 2004).

There are many definitions of job performance in the literature (Aarabi et al., 2013; Aguinis, 2013; Brudan, 2010; Campbell, 1990; Cascio & Aguinis, 2010; Chabault et al., 2012; Meister & Willyard, 2010; Usher, 2005). These definitions generally have in common that, in order to understand job performance, one needs to observe both the actions of the employees as well as the results or outcomes of these actions. However, it is essential to also understand possible counterproductive behaviours because these reflect employees’ attitudes, which in turn, influence the overall performance outcome (Bartlett & Ghoshal, 1995; Studer, 2008; Tsui et al, 2013; Usher, 2005). According to Usher (2005), when considering employees’ job performance; it is critical to focus on the behaviours as well as the outputs or results of behaviour. Usher (2005) argued that employees can be held individually accountable for the behaviours, which are in their control, as opposed to their limited accountability for the department’s sales figures, for example. Also, should the employees only be held accountable for results (e.g. sales figures) certain behaviours (e.g. politeness) may be compromised in order to achieve those results at all cost (Usher, 2005).

Job performance should be viewed in an organisational context in terms of resources, policies, practices and so forth, which are made available to employees to perform their tasks (Koopmans et al., 2012; Pulakos, 2005). Job performance
descriptions should include the quality and quantity of work required from employees to meet organisational objectives (Ivancevich et al., 2005; Luthans, 2008). From this standpoint, job performance can be viewed in terms of the systems approach (Van der Linde, 2005). The systems approach includes inputs (personality, experience, knowledge, behaviours, etc.), organisation (processes, resources available to do the work, etc.) and outputs (job performance).

2.6.2.2 Factors Affecting Individual Performance

Individual performance is influenced not only by employees’ behaviour and personal characteristics but also by external circumstances. Six main factors influence job performance (Cascio & Aguinis, 2010):

- Environmental and organisational characteristics (situational factors): factors, such as turnover or absenteeism, empowerment, policies and role clarity can potentially have a significant impact on one’s performance, especially those in leadership positions (Cascio & Aguinis, 2010; Funder, 1994; Sutherland et al., 2007)
- Environmental safety: accidents and injuries may affect outputs
- Life space variables: circumstances (such as life stability or the employee’s personality traits) that surround an employee inside and outside the work context
- Job and location: factors such as policies and practices of an organisation
- Leadership: managers may impact on individual performance by encouraging competence and creating a culture in which competence is valued (Cascio & Aguinis, 2010).
- Dispositional factors such as personality, attitude, motivation, ability, emotional intelligence and behaviour also play a significant role in employees’ performance. Research shows that where personality is concerned, conscientiousness has the greatest impact on a person’s performance (Johnson, 2003; Sutherland et al., 2007). In a study of organisational factors,
empowerment was identified as one of the main factors influencing individual performance (Liden, Wayne & Sparrowe, 2000).

Awareness of the different factors that impact on work performance helps one to consider different ways of managing and measuring performance.

### 2.6.3 Measuring Job Performance

Performance appraisal has become a continuous process by which an employee’s understanding of a company’s goals and his or her progress toward contributing to them are measured. Performance management is an integral part of the workplace as it provides a platform for supervisors and managers to measure employee performance and determine whether employees are meeting the company's expectations. The method of performance measurement varies according to the work environment, type of business and, to some extent, the employee's occupation.

#### 2.6.3.1 Defining Performance Measurement

Performance measurement is generally defined as regular measurement of outcomes and results (Cascio & Aguinis, 2010), which generates reliable data on the effectiveness and efficiency of tasks performed by an organisation (Brudan, 2010) and its employees (Schat & Frone, 2011). It relates to the measurement of performance information that quantifies input, output, and other performance dimensions of processes, products and services (Campbell, 1990). It is an ongoing progress monitoring and reporting process, particularly regarding progress towards pre-determined goals (Aguinis, 2013).

Kagioglou, Cooper and Aouad (2001) and Harry (2006) noted that performance measurement is an evaluation process used to determine how successful organisations or individuals have been in achieving their goals. Armstrong (2006) defined performance measures as a process that is concerned with inputs and outputs: inputs involve the knowledge, skills and competencies, which are required
to reach the expected results; and outputs involve meeting targets, standards or indicators.

Performance measurement can be categorised as contextual performance or task performance (Aguinis, 2013; Cascio & Aguinis, 2010). Contextual performance involves the activities one engages in that are not part of the job description (e.g. volunteering to perform extra tasks, supporting others and following procedures) but essential for accomplishing the company’s goals (Aguinis, 2013; Beatty, Murphy & Cleveland, 2001; Cascio & Aguinis, 2010). It can be predicted by personality characteristics, such as conscientiousness and emotional intelligence (Aguinis, 2013; Cascio & Aguinis, 2010; Sutherland et al., 2007). Task performance refers to job-related activities, performed to transform inputs into outputs (products and services) (Aguinis, 2013; Cascio & Aguinis, 2010; Beatty et al., 2001; Murphy, 2002). The type of tasks and activities being evaluated will determine the performance evaluation method. According to Furnham (1997), these tasks and activities may relate to quality (the excellence of the products and services delivered), quantity (how much output is generated) or accidents and rejects.

2.6.3.2 The Value of Performance Measurement

Accurate performance measurement supports business policy, procedure and future strategy. Performance measurement informs management on challenges and development opportunities for employees (Brudan, 2010; Chabault et al., 2012; Pulakos, 2005). It guides overall improvement and forms the basis for employee rewards and achievement recognition. Assessment feedback sessions provide management teams with the data needed to support promotion (Aguinis, 2013) and person-job-fit decisions (Cascio & Aguinis, 2010; Chabault et al., 2012). Performance measurement reveals the strengths, weaknesses and potential of employees and assists managers with career development decisions that will be to the benefit of both the organisation and the individual employee (Pillay, 2009). Performance measurement informs the company’s strategic objectives aimed at increased profitability and sufficiency (Studer, 2008). It also may assist employee engagement (Macey & Scheider, 2008; Meister & Willyard, 2010) in that it clearly
informs the employee on tasks to be completed and how these tasks contribute to the organisation’s end product or service (Cascio & Aguinis, 2010). Evaluations offer management the opportunity to recognise employees who perform well and this recognition of success and achievement, in turn, builds staff morale, and employees with high morale are generally more productive (Macey & Scheider, 2008; Meister & Willyard, 2010; Mendes & Stander, 2011).

2.6.3.3 Performance Measurement Criteria

Measurement criteria can be divided into two broad categories namely objective (also referred to as nonjudgmental) measures of performance and subjective (or judgmental) measures of performance (Cascio, 1998; Kaplan & Norton 1996; Landy & Farr, 1983).

Ratings are examples of subjective criteria, and this process requires one individual to make a judgment about another’s performance level (Furnham, 1997). Examples include supervisory ratings, peer ratings or self evaluation ratings. Objective criteria consist of measures that do not require a judgment, and these data consist of any data that can be counted, seen, and compared directly from one employee to another (Aguinis, 2013; Landy & Farr, 1983). Objective measures can be directly related to job performance. Examples of such measures have traditionally been production output, scrap rate, time taken to complete the task, and so forth. There are less obvious objective measures that can still directly influence performance - examples of these are absenteeism, turnover, job knowledge, accidents or grievances (Aguinis, 2013; Cascio & Aguinis, 2010; Landy & Farr, 1983).

2.6.3.3.1 Dimensions of performance criteria - In brief, criteria dimensionality refers to the reviewer’s approach to job performance and performance information (Cascio & Aguinis, 2010). From a static dimension, performance is viewed at a single point in time, making use of specified criteria and with the underlying assumption that the ratee has full and accurate knowledge of what was expected of him or her. For dynamic dimensionality, it is assumed that employees learn as they gain job
experience and should therefore be able to enhance their performance as they gain more work-related experience. Individual dimensionality looks at the unique contribution an employee makes to a company’s overall performance – it relates more to the nature of an employee’s contribution (Cascio & Aguinis, 2010).

In order to define accurate performance criteria, the following three challenges must be taken into account (Cascio & Aguinis, 2010): Firstly, those working with performance information must take job performance reliability into account. It refers to the consistency with which an employee performs over time. A focus on reliability will guide the rater to also account for intrinsic (personal factors) and extrinsic (external sources affecting the job demands on one’s behaviour) factors affecting reliability. Secondly, the consistency of the methodology used to observe and monitor performance needs to be addressed. This is important because different methods of observation may lead to inconsistent performance evaluations and results. Lastly, the dimensionality must be taken into account in order to accurately analyse job performance according to the different levels of performance on each criteria.

2.6.3.3.2 Objective measurement criteria - Objective measurement data are generally more accessible and more readily available than subjective criteria. This type of data can be collected and discrete comparisons made. However, Bommer et al. (1995) argued that performance constructs that can be measured objectively tend to be narrow in focus and are typically low-order organisational goals. The higher up in the organisation the employee is, the more difficult it is to measure relevant goals objectively. Some of the variables that lend themselves to objective measurement are absenteeism, turnover and job knowledge (Gilmore, 2008).

2.6.3.3.3 Subjective measurement criteria - Subjective measurement criteria involve ratings, rankings and paired comparisons of employees (Landy & Farr, 1983). The process involves collecting information, considering its value and using it to draw conclusions about the ratee’s performance. It is therefore based on perceptions. Individuals’ perceptions of their own and their colleagues’ work-related competences
have a huge impact in this process. Arnold and Davey (1992) found that employees continuously rated themselves far higher than their supervisors’ rating of them. There are several criticisms of the use of ratings as the only measurement when assessing employees’ performance. Bommer et al. (1995) contended that ratings as a performance measure are subject to systematic bias and random error, which objective measures seem to be less prone to. Posthuma (2000) maintained that supervisors’ subjective evaluations may be clouded by interpersonal behaviour and employees may therefore be able to influence the supervisors’ opinions of their performance without increasing their workload. For example, employees with positive willing attitudes, who work after hours, may not necessarily be highly productive, but they may be rated higher on their performance appraisals than sullen, quiet employees who are productive, but because of their lack of interpersonal skills, scored lower on their performance ratings. Despite all the arguments against judgmental criteria, the use of subjective measuring is a convenient way of summing up judgments of behaviour, which otherwise may not be discussed openly. These criteria provide a method of identifying and putting in place plans, for both exceptional performers and under-performers (Armstrong, 1996).

Armstrong (1996) argued that owing to the subjective nature of these criteria, achieving consistency is a concern. Because of the multidimensionality of performance, he also highlighted the danger of oversimplification by summing up an employee in a single rating. Supervisor ratings call for judgments about potential, and he believes that labelling employees can be both dangerous and demeaning. Fink and Longenecker (1998) maintained that performance measurement fails in most companies because of poor rater skills. Hence to ensure efficient use of judgmental measurement criteria, rater skills should be refined – especially in terms of the appraisal process.

2.6.3.4 The Performance Appraisal

Millmore et al. (2007) defined performance appraisal as a process that entails the evaluation of an employee’s performance and progress. Based on the employee’s
Performance appraisal involves two distinct processes: observation and judgement (Aguinis, 2013). Observation processes are fundamental and include the detection, perception, and recall or recognition of specific behavioural events (Aguinis, 2013; Grote, 2005). Judgement processes include the categorisation, integration, and evaluation of information (Williams, 2006). In practice, observation and judgement represent the final stage of a three-part performance measurement sequence (Grote, 2005; Williams, 2006):

- **Job Analysis** - Describes the work and personal requirements of a particular job
- **Performance Standards** - Translate job requirements into levels of acceptable and unacceptable performance.
- **Performance Appraisal** - Describes the job-relevant strengths and weaknesses of each individual

Performance appraisal, the last of the three steps in the sequence, is the actual process of gathering information about individuals based on critical job requirements (Bussin, 2013; Grote, 2011). Gathering job performance information is accomplished by observation while evaluating the adequacy of individual performance is an exercise of judgment (Aguinis, 2013).

The evaluative and developmental components of performance appraisals are equally important (Bussin, 2013). It will therefore not be surprising that the
performance appraisal is an ideal opportunity for the employee and his or her manager to (Aguinis, 2013; Bussin, 2013):

- Summarise the employee’s actual performance compared to the objectives that were agreed during the “Reaching an Agreement Discussion”
- Recognise the employee’s major strengths
- Agree on ways to improve performance that falls short of the agreed objectives
- Identify ways to further develop the employee in the next performance cycle

Bussin (2013), Grote (2005) and Williams (2006) mentioned features that would contribute to a positive and collaborative appraisal discussion:

- The employee leads the discussion
- Communication is two way
- There are no surprises as performance has been tracked throughout the year, and both parties have met to discuss progress throughout the cycle
- The review is based on hard data, and not on feelings or opinions
- Ratings are discussed by both the employee and the manager and confirmed after moderation has taken place
- The employee’s strengths and areas that need improvement are discussed
- The employee’s continued and planned development for the future is also discussed

It is extremely important to measure performance as accurately as possible. These measurements are essentially the most visible indicators of success in the total performance management system (Aguinis, 2013; Grote, 2011). Honesty and objectivity are essential to a successful reviewing performance discussion. If this is lacking then the review has no value to the employee or the organisation (Grote, 2011). Generally the intention of such reviews are to be honest and objective about a performance review, but feelings, opinions, biases and single events sometimes do colour perceptions (Bussin, 2013). Common rating errors that should be avoided include:
• Halo effect refers to favourable ratings to all job duties based on impressive performance in one or a few performance indicator(s).

• Horn effect refers to the opposite of the halo effect, namely downgrading an employee across objectives based on one performance indicator.

• First impression - developing negative or positive impressions based on impressions early on in the review process

• Recency error - allowing impressions close to the review period to negatively or positively influence assessment

• Leniency/strictness error - consistently rating someone higher than is deserved or consistently rating someone lower than is deserved.

• Central tendency error - avoiding extremes in ratings across employees

• Clone error - giving better ratings to those who are like the rater in behaviour

2.6.3.5 Supervisor Rankings

Supervisory ranking is a simplified method of appraising performance, but it is not without its own problems or potential errors (Grote, 2011; Krainer & Lopez, 2004). Ranking is simplified because the set of ranks is not required to contain more than about four or five choices. It is common to ask respondents to rank, say, their best four from a list of ten, with 1 = best, etc. (Cooper & Argyris, 1998). Although supervisor rankings are not used in the company where the research was conducted, it is discussed here because supervisor rankings were included in the study as an additional performance measurement tool. This was done in an attempt by the researcher to enhance the quality of the criterion data (Harris & Schaubroeck, 1988). Low reliability of the criterion measure may lead to a type 2 error, namely in this case, missing a significant validity coefficient that is, in fact, present (Carretta & Ree, 2000, Cascio, 1998).

To facilitate supervisor rankings, individuals or groups are asked to rank a set of options to elicite specific type of data – performance data in this case (Borman, White, Pulakos & Oppler, 1991; Cooper & Argyris, 1998; Myford & Wolfe, 2002). However, a researcher’s decision to use ranks in the first place means that results
are less informative than scoring, especially if respondents are forced to choose between some nearly-equal alternatives and some very different ones (Arvey & Murphy, 1998). Yet, supervisor ratings are often used as the only means of assessing performance in the workplace (Heneman 2006, Tziner, Murphy & Cleveland (2002). Supervisor characteristics (e.g. skill in rating, attitude towards a staff member) have an impact on the accuracy of supervisor ratings (Arvey & Murphy, 1998). Krainer and Lopez (2004) found that supervisor ratings were related to the knowledge and situational judgment aspects of performance to some degree, but supervisors were not very good at assessing employees’ competency in performing their primary technical task. Technical proficiency and ratee problem behavior had substantial direct effects on supervisory ratings. Ratee ability, job knowledge, and dependability played strong indirect (Krainer & Lopez, 2004).

A study was concluded by Myford and Wolfe (2002) to examine a procedure for identifying and resolving discrepancies in supervisory ratings. Their results suggest that, while it is important for an assessment program to identify cases in which there is obvious disagreement in the ratings assigned and have a policy to resolve those disagreements, implementing a discrepancy resolution procedure is not sufficient in and of itself for quality control monitoring. Often times, there are other anomalous ratings that discrepancy resolution procedures may miss. Fit analysis can provide a valuable adjunct to a discrepancy resolution procedure; flagging suspect rating profiles in need of expert review before a final score report is issued (Myford & Wolfe, 2002). In other studies reviews of self–supervisor, self–peer, and peer–supervisor ratings have generally concluded that there is at best a modest correlation between different rating sources Heneman (2006). Nevertheless, there has been much inconsistency across studies. Accordingly, a meta-analysis was conducted. The results indicated a relatively high correlation between peer and supervisor ratings ($\rho = .62$) but only a moderate correlation between self-supervisor ($\rho = .35$) and self-peer ratings ($\rho = .36$). While rating format (dimensional versus global) and rating scale (trait versus behavioral) had little impact as moderators, job type (managerial/professional versus blue-collar/service) did seem to moderate self-peer and self-supervisor ratings (Heneman, 2006). Interestingly, rating standards
varied over time (Krainer & Lopez, 2004). Supervisors seem to have applied more stringent rating standards from 1989 to 1992, a period marked by a recession and a large degree of distress in the banking sector. Rating standards then eased during the economic recovery from 1993 to 1998, before showing increasing signs of toughness again from 1999 through 2004 (Krainer & Lopez, 2004).

Guidelines for effective supervisor ratings and the potential positive effects thereof are important considerations for the facilitation thereof (Arvey & Murphy, 1998; Harris & Schaubroeck, 1988). To get it right, some researchers warn not to expect supervisors to be able to accurately rate all aspects of job performance (Heneman, 2006; Simmons, 2003). One should therefore first consider which aspects of job performance supervisors are equipped to assess (Grote, 2011; Harris & Schaubroeck, 1988). In addition, tasks which supervisors are equipped to accurately assess are likely to vary from job to job and will also depend on how much direct involvement supervisors have with a staff member’s daily work routines (Cooper & Argyris, 1998; Simmons, 2003; Tziner et al., 2002). Supervisor ratings can give supervisors a clearer understanding of employee concerns (Krainer & Lopez, 2004; Myford & Wolfe, 2002). Even if a supervisor encourages employees to communicate with her openly at all times, many of them do not (Heneman, 2006). By giving employees a safe context in which to communicate their opinions and concerns, the company can improve morale and strengthen relationships between managers and employees (Arvey & Murphy, 1998; Grote, 2011; Heneman, 2006). In addition, employee evaluations provide accurate information about a supervisor’s interpersonal skills that may not be available from any other source (Cooper & Argyris, 1998; Grote, 2011; Simmons, Tziner et al., 2002).

2.6.4 Performance Management

Performance measurement is crucial to employee optimisation – hence the need to manage it effectively (Dixon, Nanni & Vollmann, 1992; Lewis, Goodman & Fandt, 1998; Meister & Willyard, 2010; Shank, 1989). Traditionally, this process has been one of control and authoritarian actions that may result in disciplinary action or, in
some extreme cases, termination of the employee’s services (Amaratunga & Baldry, 2002; Tangen, 2004). However, more contemporary performance management processes include the measurement of behaviour as the key to change (Gattorna, Ogulin & Reynolds, 2003), and this is no longer used to influence employees negatively (Mendes & Stander, 2011).

There is agreement in the literature that most performance management processes include translation of overall company vision and mission into the smallest individual goals and measures. This ensures that each employee has the relevant resources, current policies and procedures to support the acquisition of goals and sharing of results (Macey & Scheider, 2008; Meister and Willyard, 2010; Mendes & Stander, 2011). Hence the performance management information can subsequently be utilised to effect positive change in the culture and practices of the organisation in order to achieve the shared targets (Becker, Huselid & Ulrich, 2001; Gerber et al., 1999; Van der Linde, 2005).

According to Pickett (2000), performance management, in its broadest context, is a managerial process that links strategic planning, performance standards, individual objectives, performance evaluation, training and individual development. Underlying this definition is the assumption that if the performance management of an organisation includes elements of control, feedback and improved communication, optimised business performance will ultimately result. In order for performance management to be effective, each organisation should have a framework in place to provide guidance.

2.6.4.1 The Performance Management System

Various authors use different terms to describe the performance management process, but generally they concur that there are four stages: planning, monitoring (through coaching), review and reward (Armstrong, 1996; Cascio, 1998; Fink & Longenecker, 1998; Landy & Farr, 1983; Spangenberg, 1994). In order to be effective, performance needs to be monitored across all facets of the organisation,
from the broadest (organisation-wide, strategic) view right through to individual level. According to Tangen (2004) a performance management system (PMS) should:

- Support strategic objectives. A PMS should be derived from the overall organisational objectives and strategies. As these change, the PMS should be adjusted accordingly.
- Have an appropriate balance. Performance should be viewed from various perspectives, not only financial. Many contemporary systems include quantity, quality, cost, people and environment measures (Cascio, 1998). The organisation in this research project included these five measures in their PMS.
- Guard against sub-optimisation. The assertion *you get what you measure* applies here because the goals and targets that are put in place have a huge impact on the behaviour of the employees being measured. Hence the correct measurements should be put in place to ensure that employees perform optimally.
- Have a limited number of performance measures. Employees function better if their targets are focused and not diluted. It is therefore crucial for the measurement criteria to be specific and relevant to the desired outcome.
- Be easily accessible. One of the main goals of a PMS is to relay relevant and important information, and it should therefore be designed in such a way that this information is available and understood by those who require it.
- Contain performance measures that have comprehensible specifications. The purpose of the measures should be defined in such a way that the relevant parties understand them. There should be clarity on the collection of data, the frequency of collection and the use of information. Furthermore, it is vital for each measure to be understood in terms of how to act in order to achieve the measure, the impact of not achieving target and the timeframe within which the targets should be reached. To ensure effectiveness, there are various different theoretically based approaches to performance management. A discussion of some of the most commonly used approaches follows.
Aguinis (2013) defined performance management as a continuous process of identifying, measuring and developing the performance of individuals and teams and aligning performance with the strategic goals of the organisation. According to Aguinis (2013), an ideal performance management system should include the following characteristics:

- **Strategic Congruence** - The system should be congruent with the unit and organization’s strategy. In other words, individual goals must be aligned with unit and organizational goals.

- **Context congruence** - The system should be congruent with the organization’s culture as well as the broader cultural context of the region or country.

- **Thoroughness** - The system should be thorough regarding four dimensions. Firstly, all employees should be evaluated (including managers). Secondly, all major job responsibilities should be evaluated (including behaviour and results). Thirdly, the evaluation should include performance spanning the entire review period, not just a few weeks or months before the review. Finally, feedback should be given on positive performance aspects as well as those that are in need of improvement.

- **Practicality** - Systems that are too expensive, time consuming, and convoluted will obviously not be effective. Good, easy-to-use systems (e.g., performance data are entered via user-friendly software) are available for managers to help them make decisions.

- **Meaningfulness** - The system should be meaningful in several ways. Firstly, the standard and evaluation conducted for each job function must be considered important and relevant. Secondly, performance assessment must emphasize only those functions that are under the control of the employee. Thirdly, evaluations must take place at regular intervals and at appropriate moments. Fourthly, the system should provide for the continuing skill development of evaluators. Finally, the results should be used for important personnel management decisions.

- **Specificity** - A good system should be specific: it should provide detailed and concrete guidance to employees about what is expected of them and how they can meet the expectations.
• Identification of effective and ineffective performance - The performance management system should provide information that allows for the identification of effective and ineffective performance.

• Reliability - A good system should include measures of performance that are consistent and free of error.

• Validity - The measures of performance should also be valid. In this context, validity refers to the fact that the measures should include all relevant performance facets and should not include irrelevant performance facets.

• Acceptability and fairness - A good system is acceptable and is perceived as fair by all participants. Perceptions of fairness are subjective and the only way to know if a system is seen as fair is to ask the participants about the system. Such perceptions include four distinct components, namely distributive justice, procedural justice, interpersonal justice and finally informational justice.

• Inclusiveness - Good systems include input from multiple sources on an ongoing basis. Firstly, the evaluation process must represent the concerns of all the people who will be affected by the outcome. Secondly, input about employee performance should be gathered from the employees themselves before the appraisal meeting.

• Openness - Good systems have no secrets. Firstly, performance is evaluated frequently and performance feedback is provided on an on-going basis. Secondly, the appraisal meeting consists of a two-way communication process during which information is exchanged. Thirdly, standards should be clear and communicated on an ongoing basis. Finally, communications are factual, open, and honest.

• Correctability - The process of assigning ratings should minimize subjective aspects; however, it is virtually impossible to create a system that is completely objective because human judgment is an important component of the evaluation process.

• Standardization - As noted earlier, good systems are standardized. This means that performance is evaluated consistently across people and time

• Ethicality - Good systems comply with ethical standards. This means that the supervisor suppresses his/her personal self-interest in providing evaluations.
Both Aquinis (2013) and Tangen’s (2004) explanation of the characteristics of performance management is thoroughly detailed and addresses key aspects within an organizational setup. However, it does not specifically address the changing nature of the work world. According to Schläfke et al. (2013), performance management and performance analytics are becoming more relevant and is in accordance with the increased competitive nature of the 21st century world of work. For example, by making use of available technology businesses can electronically gather, disseminate and integrate performance management data much more effectively (Brudan, 2010). It was also recently found that employees responded more favourably toward online performance management systems as opposed to the traditional paper-based system (Parry & Tyson, 2011). Technology is increasingly being used in performance management which means that a stronger focus is needed on how humans interact with information technology (IT) and software systems (Schreuder & Coetzee, 2010). Performance management data can be used as important business signals. However, the effective extraction and utilisation of this data might mean that those managing the performance management system should acquire new skills. Some of these skills include: IT-, mathematical-, statistical- and economical skills (Schläfke et al., 2013). Effectively analysing performance management data can guide a deeper understanding of business dynamics and help to control key performance drivers (Brudan, 2010; Schläfke et al., 2013).

Apart from the changing work environment, companies must be cogniscent of the fact that employees are dynamic too and that most of them strive toward self-actualisation (Greening, 2010). Furthermore, the results of a study conducted by Rothmann and Coetzee (2003) showed that personality dimensions (emotional stability, extraversion, openness to experience and conscientiousness) were related to task performance and creativity. Research findings like these encourage organisations to define new perspectives on performance management. For example, some companies are revamping their performance management processes to set career expectations in increasingly flattened organisations. Meister and Willyard (2010) explain that the presence of millennials in the workplace has added a layer of complexity to an already complex work environment. According to
research, one answer could be to encourage an open career dialogue between a millennial and his or her manager. For example, younger employees may have unreasonable but strongly held career expectations. Having a dialogue forces the manager to explain the gap and start a conversation with the employee. These programmes have been successful because everyone walks away knowing where they stand (Meister & Willyard, 2010). Lievens, Conway and De Corte (2008) stressed the importance of directly linking performance management rating systems to an organisation’s culture. Performance management is a multi-faceted concept. It is important to understand the different theoretical approaches to performance management in order to select the most appropriate system.

2.6.4.2 Different Approaches to Job Performance Management

Job performance is influenced by, and in turn influences, many factors apart from only the actual tasks performed. In understanding job performance, it is vital to grasp the interdependency of these factors (Gilmore, 2008). Many of the various approaches to the administration of performance management take cognisance of all the factors involved, and a discussion of five well known approaches follows (Cascio & Aguinis, 2010; Gilmore, 2008; Spangenberg; 1994; Tangen; 2004).

2.6.4.2.1 The systems approach - According to Spangenberg (1994), the importance of the systems approach to performance management cannot be overemphasised. The aim of this approach is to illustrate performance management in its entirety, including the relationships between the various elements (Cascio & Aguinis, 2010). The systems approach makes provision for those factors that can make or break the system.

This approach is extremely comprehensive and illustrates the numerous elements that influence the implementation and effectiveness of a PMS (Gilmore, 2008; Spangenberg, 1994). The systems approach takes into account all the possible areas which, if not designed accurately, may cause the system to fail, and
emphasises that it is crucial to consider not only financial outcomes when assessing performance (Cascio & Aguinis, 2010; Spangenberg, 1994).

2.6.4.2.2 Sink and Tuttle’s approach - According to the Sink and Tuttle approach, (Tangen, 2004; Van der Linde, 2005), organisational performance is an intricate interrelationship between seven performance criteria. Criteria of the Sink and Tuttle model include:

- Effectiveness involves doing the correct things, at the right times, with the right quality - and is expressed as a ratio of actual versus expected output.
- Efficiency involves doing things right - and is expressed as a ratio of expected resources consumed versus actual resources consumed.
- Quality is measured at six checkpoints in order to make an extremely broad concept more tangible.
- Productivity is expressed as the ratio of output versus input.
- Quality of work life is essential to ensure the success of the system.
- Innovation is an element that is crucial for sustainable performance and improvement.
- Profitability is the goal of any organisation.

The seven elements are important in any organisation (Gilmore, 2008). However, according to Tangen (2004), there are several limitations to this approach, including the lack of attention to the need for flexibility and consideration of the customers’ perspective. Cascio and Aguinis (2010) also reiterated the importance of flexibility within any performance management system.

2.6.4.2.3 The balanced scorecard approach - The balanced scorecard framework was developed by Kaplan and Norton (Becker et al., 2001; Gilmore, 2008; Maylor, 2003; Stein & Book, 2006). This framework incorporates measures that describe an actual value-creation process instead of focusing on only the financial results (Cascio & Aguinis, 2010). This approach suggests that organisations should utilise a balanced set of measures that allows a brief view of the business (Cascio & Aguinis,
A key feature of the balanced scorecard is that it is tailored to what employees can control (Chase, Aquilano & Jacobs, 2001). These measures should stem from four perspectives and provide answers to the following questions (Tangen, 2004):

- Financial perspective: How do we look to our shareholders?
- Internal business perspective: What must we excel at?
- Customer perspective: How do our customers see us?
- Innovation and learning perspective: How can we continue to improve and create value?

The balanced scorecard approach limits the number of measures and therefore provides a focused view of critical areas, thereby limiting information overload (Becker et al., 2001). According to Tangen (2004), it compels managers to consider all four perspectives, and not only focus on one. The main limitation to this approach is that it has been designed for use by managers and does not prove useful to the organisation's lowest level employee (Gilmore, 2008). It has also been criticised because it has been designed to monitor and control, rather than to be used as an improvement tool (Cascio & Aguinis, 2010). In addition, it does not provide sufficient guidelines on how to identify, introduce and ultimately use measures to manage the business (Cascio & Aguinis, 2010; Gilmore, 2008; Tangen, 2004).

2.6.4.2.4 The performance pyramid - Tangen, (2004) proposed an approach that clearly links the performance measures at the various hierarchical company levels to ensure that each facility in the organisation works towards the same ultimate goal.

This approach is also referred to as the SMART system. According to Tangen (2004), it translates the organisation's objectives from the top down, and then measures them from the bottom up to ensure that all targets are met, culminating in the eventual achievement of the overall business plan. Four levels of objectives that address external effectiveness and internal efficiency are included (Gilmore, 2008). The organisation’s vision is at the first level, and is then translated into relevant objectives for each of the levels below. Each level is essential, and the targets build
up from the day-to-day measurement (at the lowest level) to the longer time-span measurement (at the higher levels).

According to Tangen (2004), the strength of the model lies in the integration of the vision with the lowest operational objective. The chief limitation, however, is the failure to include any mechanism, either to identify key performance indicators or to integrate the concept of continuous business improvement.

2.6.4.2.5 Medori and Steeple’s approach to job performance management -

According to Tangen (2004), Medori and Steeple’s approach consists of six detailed stages and is an integrated framework for auditing and enhancing performance management systems. A description of the stages follows:

Stage 1 involves the definition of the organisation’s strategy and success factors.

Stage 2 involves matching the strategic requirements with six defined priorities (quality, cost, flexibility, time, delivery and future growth).

Stage 3 involves selecting suitable measures through the use of checklists.

Stage 4 involves an audit of the existing measures in order to ascertain which, if any, should be transferred to the new PMS.

Stage 5 involves the implementation of the new measures, where each measure is described in terms of title, objective, benchmark, equation, frequency, data source, responsibility and improvement.

Stage 6: involves the periodic review of the organisation’s PMS.

A strong advantage of this approach is that it can be designed as a new PMS or used to exploit an existing system, and also contains descriptions of how measures should be identified and implemented (Gilmore, 2008). According to Tangen (2004), a limitation of this approach is that little guidance is given on how the performance measurements grid is created, as well as the fact that only six priorities are used in the construction of the grid. There are many other categories into which measures can be divided (Cascio & Aguinis, 2010).
2.6.5 Performance Management in the Research Organisation

The company where this study was conducted falls within the civil engineering industry and the main organisational function is the construction and maintenance of public roads. The performance management system in the company is extremely user friendly and kept as uncomplicated as possible. The reason for this being that many of its employees work on remote construction sites, long distances and far away from its head offices where performance information is kept and evaluated.

The company makes use of a subjective appraisal system and draws a clear distinction between hourly-paid employees and employees who earn a monthly salary. For the purpose of this study only the hourly paid employees were considered as they make up the bulk of the company’s drivers and machine operators. The drivers and operators are also the only group of employees being assessed on the instruments under investigation in this research. Hourly paid employees are appraised on three generic criteria: productivity, care of resources and attitude and safety.

The conceptual reasoning that support the company’s choice of these constructs as performance criteria is closely linked to the company’s culture, specifically on construction sites. For instance, as a criterion construct, productivity refers to much more than the industrious completion of assigned tasks. It refers to every effort taken to add value to the end product. If, for example, a machine is temporarily under maintenance, it is expected of the operator to find something to do that will contribute to the completion or value of the project. Tasks may include cleaning, picking up and storing sign boards, loading or reporting to the foreman to be assigned a task for the few hours his vehicle is idle. The second criterion is care of resources and has to do with utilising material in the most efficient manner possible. Apart from taking care of expensive trucks and machines, this criterion will also include simple acts like properly closing a half-used bag of cement, washing and cleaning tools or efficient stock-piling. Care of resources also has to do with the documentation, security, waste management and quality of any material (or resource) handled on site. The
third criterion is a combined concept called attitude-and-safety. The company values and rewards a positive attitude, the willingness to go the extra mile and a can-do attitude. This type of attitude is of particular importance when there is emergency work to be done, or when an employee is expected to work overtime. Safety is part of the culture of the company but the construct encompass an awareness of site activity and the whereabouts of fellow construction workers at any given time.

The three performance criteria constructs associate with the job performance dimensions listed in *Table 1.1* The company’s conceptualisation of productivity coincides with Viswesvaran’s (1993) dimensions of overall job performance, productivity and effort and Campbell’s *et al.* (1993) dimensions of job-specific task proficiency, non-job-specific task proficiency and demonstrating effort. Similarly, the company’s definition of care of resources as a performance measurement criterion relates to Viswesvaran’s (1993) dimensions of quality and administrative skills and Campbell’s *et al.* (1993) dimension of management/administrative. Lastly, the company’s construct of attitude and safety can be linked to Viswesvaran’s (1993) dimensions of interpersonal skills and rule following and to Campbell’s *et al.* (1993) dimension of maintaining personal discipline.

It could also be mentioned that the company’s innovative definition of its performance criteria allows its performance management system to be more dynamic and to adapt to changing environments. The company’s definition of productivity and its link to the work culture on construction sites can be applied in different work environments. For instance, the company recently embarked on new form-work projects which focus primarily on structures as opposed to road-works. The same performance management criteria can be effectively applied on these new for-work sites because the company’s definition of performance criteria can be adapted to different work environments.

Each employee is appraised by his or her line manager and assigned a score out of ten for each criterion. Appraisals occur on a monthly basis. The total score, out of a maximum of thirty, determines the amount of hours the employee is entitled to as a
production bonus. For example, an employee who scored eleven out of a total of thirty performance points in the month of August will receive a bonus in September equal to eleven normal working hours. These performance assessments, therefore, form the basis for the calculation of monthly production bonuses. Thus, a higher performance score means a bigger bonus at month end.

The company pays production bonuses on a monthly basis and payment is therefore directly related to the performance management system. The production bonus structure was established to reward those drivers and machine operators who perform well. It was developed as part of an attempt to stimulate efficiency and productivity to ultimately become more competitive in an ever changing industry.

2.6.6 The Future of Performance Management

Organisations all over the world are changing rapidly (Meister & Willyard, 2010). The nature of work and the work-place itself, the traditional employment contract, and the composition of the workforce are all dramatically changing (Luthans, 2008). These changes are in terms of structure, workforce composition, reward systems, service contracts, technology and information, and are the results of technological, economic and political developments (Chabault et al., 2012; Luthans, 2008; Meister & Willyard, 2010). Competition is increasing and the global economy brings new international competition (Schreuder & Coetzee, 2007; 2010). At an organisational level advanced information technology, globalisation, diversity, and trying to solve ethical problems and dilemmas come to the fore. These are unquestionably major issues facing contemporary organisations. However, according to Luthans (2008) the basic premise and assumption is that managing employees have been, are, and will continue to be, the major challenge and critical competitive advantage.

Meister and Willyard (2010) contended that the workplace will increasingly focus on the performance of people as a core company asset. Models of flexible working conditions will continue to evolve in response to the changing needs of Baby Boomers and the preferences of Generation X and Generation Y in the workplace.
(Meister & Willyard, 2010). As a result, greater attention will be given to measuring and improving the performance of people – as opposed to the traditional one-dimensional focus on finance and physical assets.

In terms of talent, Pandey (2012) explained that problem solving and creativity skills will become increasingly important. Furthermore, the performance of highly promising individuals should be integrated with the objectives and overall performance of the organisation (Chabault, et al., 2012; Pandey, 2012; Shank, 1989). Innovative organisations driven by the 21st century’s knowledge-economy need individuals with original and varied competencies to bring creative flair. In this context, talent management (and performance management in particular) becomes a matter of strategic and critical importance (Chabault, et al., 2012; Meister & Willyard, 2010; Parry & Tyson, 2011).

Research by Meister and Willyard (2010) showed that successful companies are revamping their performance management processes. They found that these organisations deliberately set career expectations in increasingly flattened organisational structures. Meister and Willyard (2010) explained that the presence of Millennials in the workplace has added a layer of complexity to an already complex work environment. To overcome this challenge one answer could be to encourage an open career and performance dialogue between a Millennial and his or her manager. For example, younger employees may have unreasonable but strongly held career expectations. Having a dialogue forces the manager to explain the gap and start a conversation with the employee to explain the performance necessary for the employee to reach the next level. These programmes have been successful because everyone walks away knowing where they stand (Meister & Willyard, 2010).

Furthermore, organisations should be in a position to offer a workplace environment that encourages individual performance (i.e. best working tools made available along with access to expert colleagues in the field) (Meister & Willyard, 2010; Parry & Tyson, 2011). What is required here is the creation of a stimulating and dynamic environment for the employee on an intellectual and emotional level, as well as
exciting challenges and opportunities (Chabault, et al., 2012; Schlafke et al., 2013). Organisations should recognise that in the Web community, status is built upon making meaningful contributions (Meister & Willyard, 2010), so companies should ensure the inclusion of ratings by peers (Meister & Willyard, 2010; Parry & Tyson, 2011). Organisations should examine how they can integrate employee expectations for social media usage in the company’s performance management practices so the quality of the organisation’s online contributions is part of the overall performance management system (Meister & Willyard, 2010). Research found that the utilisation of e-HRM is positively related to perceptions of general HRM effectiveness in line managers and employees alike. It was also found that reactions to an online performance management system were more positive than those to a paper-based version of the same system (Parry & Tyson, 2011).

The rise of business intelligence software products over the last ten years has had a profound impact on how companies manage their operational performance (Brudan, 2010; Galinsky & Matos, 2011). Enterprise resource planning software, combined with business intelligence software, enabled companies to reach new levels of data integration, by making the data gathering and reporting process more streamlined (Brudan, 2010). In addition, Mass Career Customization (MCC) processes have created a structured performance appraisal process for every employee and manager that considers the career aspirations of men and women of all ages in the context of their work, personal, and family responsibilities (Brudan, 2010; Galinsky & Matos, 2011). Making this or similar processes universal, not just aimed at employees of specific ages or backgrounds, it has the potential to become the “new normal” way of doing business and managing performance (Galinsky & Matos, 2011). As a part of every employee’s performance appraisal, the MCC process guides employees and their managers to make choices around four major dimensions of career progression — role, pace, location/schedule and workload - calibrating each based on their current aspirations and life circumstances. Since the system is modular, with the ability to personalise each arrangement at any point in time, the MCC is uniquely suited to a workplace where both men and women are looking to manage work and family (Galinsky & Matos, 2011).
Increased business competition requires even more rapid and sophisticated information and data analysis (Gratton, 2011; Meister & Willyard, 2010; Zsolnai, Junghagen, & Tencati, 2012). These requirements challenge performance management systems to effectively support the decision making process (Brudan, 2010; Cascio & Aguinis, 2011; Chabault et al., 2012). Business analytics is an emerging field that can potentially extend the domain of performance management to provide an improved understanding of business dynamics and lead to better decision making (Meister & Willyard, 2010; Schreuder, & Coetzee, 2010). The increasing relevance of performance analytics is undoubtedly due to the fast-growing hyper-competition effect in today’s world of work (Meister & Willyard, 2010). According to this effect, companies more rapidly and increasingly compete to provide lower costs and better quality with better know-how to create competitive advantages (Schläfke et al., 2013) – this will not be possible without strategically informed employees who consistently perform in accordance with the company’s ever changing objectives (Cascio & Aguinis, 2011; Meister & Willyard, 2010; Schläfke et al., 2013).

2.6.7 The Role of Industrial Psychologists in Managing Employee Performance

Industrial psychologists are equipped to partner with organisations in order to guide, assist and manage performance management systems which are suitable and sustainable in the very dynamic 21st century world of work (Barnard & Fourie, 2007; Benjamin & Louw-Potgieter, 2008; Kelly & Finkelman, 2011). An emerging systems-based approach to strategic performance management (Shamsi, 2010) is represented by strategy dynamics. Strategic management dynamics is concerned with understanding and managing performance through time, focusing on the factors that explain why performance is as it is today, and how it might be managed into the future (Brudan, 2010). Contemporary trends in global competition, rapid technological developments and increased use of management information systems and the Internet, developments in planning and control and management thinking, and changing demographics are putting pressures on both profit and non-profits
organisations (Chabault et al., 2012; Luthans, 2008; Meister & Willyard, 2010). As a consequence, companies are having more and more difficulty in achieving sustained performance (Grigore, Bagu & Radu, 2009). A well-designed performance management process stimulates managers to develop high-quality strategic plans, set ambitious targets, and track performance closely - all activities which help to achieve strategic objectives and consequently sustained value creation (Grigore et al., 2009; Lievens et al., 2008). Performance raters must, however, match their rating policies to the organisation’s culture and strategic objectives (Chabault et al., 2012; Lievens et al., 2008).

Performance management is in fashion among new-age corporations (Cascio, & Aguinis, 2010; Chabault et al., 2012; Luthans, 2008; Meister & Willyard, 2010; Shamsi, 2010). Appropriately appraising the workforce on the basis of their performance instils the motivation in them to excel in their work, thereby, increasing the productivity of the organisation (Shamsi, 2010). In this regard, performance management should be defined as a strategic business approach.

It is, after all, concerned with the broader issues facing the business if it is to function competitively in the dynamic markets of the 21st century (Brudan, 2010; Chabault, et al., 2012; Shamsi, 2010; Schläfke et al., 2013; Shank 1989; Tsui, et al., 2013). Performance management systems should consider the inclusion of analytical tools, especially when data are potentially available and can be converted into business signals (Aguinis, 2013). Decision makers might, however, need to acquire new skills (e.g. mathematical, statistical, econometrics, and IT) to develop the ability to use business analytics more effectively, especially in terms of managing employee performance (Schläfke et al., 2013). Performance management analytics refers to the extensive use of data and analytical methods to understand relevant business dynamics, to effectively control key performance drivers, and to actively increase organizational performance (Schläfke et al., 2013).
The increased use of technology in the execution of personnel functions such as performance management raises concerns regarding the human–machine interaction and seems to call for more research into the whole area of ergonomics (Schreuder & Coetzee, 2010) and the collection and analysis of accurate criterion data (Cascio, & Aguinis, 2010).

2.7 CHAPTER SUMMARY

In this chapter, the focus was on the concepts of psychomotor ability, intelligence and learning potential, as well as the measurement thereof. Reference was made to the evolving influences of cognitive ability, and the various theories were presented. The arena of psychometric testing was discussed and the traditional versus dynamic assessment methods explored. The main criticism of traditional psychometric tests is that they assume that the opportunities for learning have been similar across cultures, which in South Africa is clearly not the case. It was therefore essential for dynamic tools to be developed that measure learning potential rather than products of learning.

Because organisations are faced with ever-increasing uncertainty, turbulence and changes in the external environment, accurate assessment for the purpose of selecting high performing employees is becoming more and more important. Both learning potential and psychomotor ability, the two predictor variables of this study, were discussed in depth. Job performance was explained and the importance of performance measurement was emphasised. Different performance managements system approaches were explored and future trends investigated.
CHAPTER 3: RESEARCH ARTICLE

THE PREDICTORS OF PERFORMANCE AND THE PERFORMANCE OF PREDICTORS - PSYCHOMOTOR ABILITY AND LEARNING POTENTIAL AS PREDICTORS OF JOB PERFORMANCE

Department of Industrial and Organisational Psychology, UNISA

Louis Olivier
Marié de Beer
ABSTRACT

The principal aim of the study was to examine the predictive validity of psychomotor ability and learning potential with regard to the work performance of drivers and machine operators in a construction company. A secondary aim of the study was to determine the learning potential and psychomotor ability of drivers and machine operators. The participants' work performance was measured and the relationship between the two independent variables and job performance was determined. A sample of 95 dedicated drivers and machine operators voluntarily participated in the study. The Vienna Test System (VTS) and TRAM (Transfer, Automisation and Memory) assessments were administered on the sample group. No statistically significant relationships were found between the VTS or TRAM and work performance in this study.

Key Terms

Psychomotor ability, learning potential, work performance, intelligence, dynamic assessment, cognitive ability, cognitive assessment, Vienna Test System (VTS), Transfer-Automisation-and-Memory (TRAM), psychometric test
INTRODUCTION

Key Focus of the Study

The 21st century knowledge economy creates unique challenges for talent attraction and the retention of high performing individuals. Managing the human factor as a strategic asset in organisations will remain the primary challenge in securing a competitive advantage. The road construction industry in South Africa is no different. There is growing competition between civil engineering contractors to secure tenders and to maximise profitability. This is only possible with a sufficient and sustainable labour force. Valid selection processes are therefore required to ensure that the most productive individuals are selected for the most suitable jobs. Reliable and valid performance predictors will assist employers in making appropriate selection decisions. Selecting high performing individuals will support and enhance overall organisational performance.

Background to the Study

Methods of attracting and securing future talent should have higher frequencies of good work performance as one of its primary objectives. It follows then that appropriate selection based on effective assessment methods can greatly enhance the quality and productivity of an organisation’s workforce (Murphy & Maree, 2006; Pulakos, 2005). In this study the investigation was focused on whether psychomotor ability and learning potential are statistically and practically significant predictors of work performance.

Trends from the Research Literature

Psychomotor abilities, skills, and constructs are measured for a number of practical purposes, including the understanding of neurological deficits, appropriate perception-stimuli interactions, safety, intelligence, emotional wellbeing, stress tolerance (Vorster, 2012) and employment selection (Keyser, 2012; Pelser, 2002).
Research has indicated that psychomotor ability is an important indicator of performance, especially for pilots, drivers and machine operators (Bergh et al., 2005; Carretta & Ree, 1996b; Fernández-Ballesteros, 2003; Johnston & Catano, 2002; Keyser, 2012; Mohan, Srivastava & Srivastava, 1984; Reynolds & Adams, 1953; Vesia, Esposito, Prime & Klavora, 2008). Psychomotor tests were used in the armed forces from as far back as the early 1900’s (Bergh et al., 2005). Research in this area increased during the First World War, as part of a general drive to improve pilot selection techniques in the United States of America (Carretta & Ree, 1996b; De Kock & Schlechter, 2009; Flotman, 2003; Griffin & Koonce, 1996; Martinussen, 1996; Mnuguni, 2011; Pelser, 2002; Ree & Carretta, 1996a). Human error, concern for safety, high pilot training costs and a growing interest in scientific selection methods justified the use of psychomotor tests, especially in aviation (De Kock & Schlechter, 2009; Duke & Ree, 1996; Griffin & Koonce, 1996; Martinussen, 1996; Ree & Carretta, 1998). By far the greatest amount of research in the use of psychomotor ability as predictor in selection exercises within the last three decades has centred on pilot selection in the military context (Bergh et al., 2005; De Kock & Schlechter, 2009). In the South African context a few research studies have focused on the predictive validity of psychomotor ability on job performance of drivers and machine operators (Bouwer, 1984; De Kock & Schlechter, 2009; Keyser, 2012; Pelser, 2002; Schoeman, 1995).

Learning potential, on the contrary, has been widely researched in South Africa (De Beer, 2000; 2006; Gilmore, 2008; Pelser 2002. Schoeman, De Beer & Visser, 2008; Taylor, 1994; 1999). The concept refers to an overall cognitive capacity and includes both present and potential or projected improved future performance. This implies that cognitive ability is dynamic and changeable and can therefore be increased (De Beer, 2006; Taylor, 1999). South African research results have indicated that learning potential is an appropriate measure to predict job performance (Gilmore, 2008; Keyser, 2012; Pelser, 2002). The dynamic assessment methodology can be utilised to effectively and scientifically measure learning potential (De Beer, 2006; Taylor 1994; 1999). Dynamic assessment is a more modern approach to assessing intelligence and learning potential in particular (Deadrick & Madigan, 1990; Murphy,
It is an interactive approach to conducting assessments and focuses on the learner's ability to respond to intervention (Haywood & Lidz, 2007).

In essence, dynamic assessment is a theory of learning and a theory about how individuals acquire and express knowledge (Benjamin & Lomofsky, 2002; Murphy, 2011). There are variations on several dimensions of the dynamic assessment model. However, there are several salient characteristics and assumptions that can be identified (Caffrey, Fuchs & Fuchs, 2008; Deadrick & Madigan, 1990; De Beer, 2006; Frisby & Braden, 1992; Grigorenko, 2009; Haywood, 2008; Haywood & Lidz, 2007; Lidz, 2009; Murphy, 2006; Zollezzi, 1995):

- Some abilities that are important for learning (in particular) are not assessed by normative, standardised intelligence tests.
- The assessor actively intervenes during the course of the assessment of the learner with the goal of intentionally inducing changes in the learner's current level of independent functioning.
- Most people typically function at less than their intellectual capacity.
- The assessment focuses on the learner's processes of problem solving, including those that promote, as well as obstruct successful learning.
- The most unique information from the assessment is information about the learner's responsiveness to intervention.
- Teaching within the test provides a useful way of assessing potential as opposed to current performance.

Learning potential scores derived from dynamic assessments can then be measured statistically against performance outcomes to determine the statistical significance of relationships between learning potential and job performance. There is a recent and growing body of literature focussing on the importance of employee performance at work (Aarabi, Subramaniam & Akeel, 2013; Brudan, 2010; Chabault, Hulin & Soparnot, 2012); Meister & Willyard, 2010; Parry & Tyson, 2011; Schläfke, Silvi, & Möller, 2013; Schat & Frone, 2011). Performance of employees is significant for organisations (Aarabi et al., 2013) because of its direct impact on overall
organisational productivity (Aarabi et al., 2013; Chabault et al., 2012; Gilmore, 2008). Job performance is central to the success of an organisation (Brudan, 2010; Chabault et al., 2012; Meister & Willyard, 2010; Parry & Tyson, 2011; Schläfke et al., 2013). Employee job performance represents the primary contribution of individuals to organisational effectiveness and the primary reason individuals are employed by organisations in the first place (Schat & Frone, 2011).

The importance of selecting employees based on their ability to perform cannot be underestimated (Pelser, 2002). The focus of this study was to investigate the relationship between both psychomotor ability and learning potential and the job performance of drivers and machine operators in a road construction company. Only two recent studies have focused on the combination of these two predictors and their relationship to job performance behaviour (Keyser, 2012; Pelser, 2002). These studies were not applied to machine operators in the road construction industries - which indicated a gap in the literature which this research addressed.

In view of the above, the core research problem and specific research objectives that were addressed in this study are founded on the following question: Do psychomotor ability and learning potential statistically significantly predict the work performance of drivers and machine operators? The following research hypotheses were posed and tested empirically: 1. Psychomotor ability and learning potential are statistically significantly and positively related to work performance. 2. Psychomotor ability statistically significantly predicts work performance. 3. Learning potential statistically significantly predicts work performance. 4. Psychomotor ability and learning potential jointly statistically significantly predict work performance.

The general aim of the study was therefore to investigate whether psychomotor ability and learning potential statistically significantly predict work performance. The specific literature aims of the research were to conceptualise psychomotor ability and learning potential, to conceptualise job performance and to ultimately investigate the theoretical relationship between psychomotor ability and learning potential in
predicting job performance, with detailed reference to drivers and machine operators.

The specific empirical aims of the research were to:

- Investigate the statistical and practical significance of the relationships between psychomotor ability, learning potential and job performance
- Investigate whether psychomotor ability and learning potential statistically significantly predict work performance amongst driver and machine operators in a road construction company
- Formulate recommendations towards optimising the selection, training and development of drivers and machine operators as well as for future research.

The Potential Value-Add of the Research

The study’s academic substance, practical merit and importance are threefold. Firstly it addresses an existing gap in the research literature in that psychomotor ability and learning potential are used in combination to predict the job performance of machine operators in the road construction industry. Secondly, the research adds to a growing body of knowledge which scientifically supports the importance of employee selection based on valid and reliable psychometric assessment practices (Fleischman 1988). According to Pelser (2002), the rationale for scientific personnel selection is to be found in two common sense realities:

- Individuals differ in terms of their abilities, knowledge, interests and personalities
- Jobs differ in terms of the skills and human qualities required to get the job done

Scientific personnel selection is based on the assumption that these differences between individuals, on the one hand, and jobs, on the other, can be measured in some way. The organisation can then capitalise on individual differences by
selecting those candidates who possess the greatest number of qualities judged to be important for success in any particular job. This should lead to a more productive organisation. Thirdly, and perhaps most importantly, a valuable consequence of the research is that it highlights the importance of generating accurate and reliable job performance criterion data for personnel selection and performance management (Ackerman, Cianciolo & Bowen, 1999; Aguinis, 2013; Bertua, Anderson & Salgado, 2005; Cascio, 1998; Cascio, & Aguinis, 2010; De Kock, & Schlechter, 2009; Fernández-Ballesteros, 2003; Gardner & Deadrick, 2012; Gregory, 2007; Hackett, 2002; Johnston & Catano, 2002; Laher & Cockcroft, 2013; Murphy & Maree, 2006; Pelser, 2002; Ployhart & Hakel, 1998; Roth et al., 2014; Russell, Colella & Bobko, 1993; Tracey, Sturman, Shao, & Tews, 2010).

What will follow is an explanation of the research design, the results obtained from the study and an overview discussion on the conducted research study.

**RESEARCH DESIGN**

**Research Approach**

A quantitative approach (Howell, 2004; Terre Blanche, Durrheim & Painter, 2006; Welman, Kruger & Mitchell, 2005) was followed and a cross-sectional field survey design (Howell, 2004; Terre Blanche, et al., 2006; Welman et al., 2005) used. This approach was preferred because a quantitative research approach allowed the researcher to measure and analyse the data statistically - especially the relationships between the independent and dependent variables and to study these relationships (Terre Blanche, et al., 2006). In the current study, this is of particular importance because the research aims were constructed and based on the investigation of the relationships between variables, or then the prediction of one dependent variable (driver and machine operator performance) by means of two independent variables (psychomotor ability and learning potential). In addition, the quantitative approach is advantageous because the researcher can be more objective about the findings of
the research (Terre Blanche, et al., 2006). Quantitative data was used to test the hypotheses of the current study using statistics.

A cross-sectional field survey design involves the observation of all of a population, or a representative subset, at one specific point in time (Welman et al., 2005). The aim is to provide data on the population under study. In a cross-sectional survey data is collected to make inferences about a population of interest at one particular point in time (Howell, 2004; Terre Blanche, et al., 2006; Welman et al., 2005). Cross-sectional surveys have been described as snapshots of the populations about which they gather data (Welman et al., 2005). Cross-sectional studies are observational in nature and are known as descriptive research, not causal. Researchers typically record the information that is present in a population, but they do not manipulate variables. This type of research can be used to describe characteristics that exist in a population, but not to determine cause-and-effect relationships between different variables (Welman et al., 2005). However, this method is often used to make inferences about possible relationships between variables - which was the case in the current study.

Research Method

In adhering to the quantitative research approach, numerical data was obtained from the research sample. All drivers and machine operators working in the company were assessed on the Vienna Test System (Schuhfried, 1996) and the TRAM assessment battery (Taylor, 1999) with primary data collected on psychomotor ability (VTS) and learning potential (TRAM) assessment results as well as work performance (performance appraisals and supervisor rankings). The data was statistically analysed by making use of correlation coefficients and regression analyses to test the stated hypotheses and answer the research question.

Research Participants

The research organisation employs approximately 650 employees, depending on contract availability. One-hundred-and-thirty of the company’s permanent members
of staff are dedicated drivers and machine operators. This group of employees constituted the target sample in terms of the current study. The research population constitute the drivers and machine operators in the wider road construction industry of South Africa. In this particular case, therefore, the target sample is a convenience sample.

At the onset of the study, the researcher was cognisant of the fact that the target sample may not be equal to the realised sample. Although a hundred percent response rate was favourable, some individuals were absent at the time of the assessments and a few others resigned from the company before the research was concluded or before the performance data was gathered. As a result, not all the participants had a complete data set. Due to unforeseen circumstances and the company’s operational requirements, not all the drivers and operators managed to complete both the VTS and TRAM assessments. In addition, participants’ machines went for irregular maintenance services which meant that the particular driver did not receive a performance score for that month. This was an unfortunate outcome of the study and contributed to the predicament of working with a very small sample in the first place. Complete data sets for a total of only 95 drivers and operators were available from the dependent variable (driver and machine operator performance) and two independent variables (psychomotor ability and learning potential).

**Measuring Instruments**

The measuring instruments used in the research were the Vienna Test System (VTS) (Schuhfried, 1996) to measure psychomotor ability and the TRAM learning potential test battery (Taylor, 1999). The research company’s performance appraisals for drivers and operators and supervisor rankings were used to measure the performance of drivers and machine operators.

*The Vienna test system*. The Vienna Test System (VTS) is a computer-assisted application of a large number of highly diverse psycho-diagnostic tests, measuring reaction times in tasks that require choosing among complex stimuli (Schuhfried, 1996). The following specific subtests of the Vienna Test System were used.
**Cognitrone** - This subtest of the Vienna Test System assesses the candidate’s ability to concentrate and to adjust his/her work tempo to different stimuli patterns (Schuhfried, 1996; 2000a). It was included because of its logical conceptual link with road construction drivers and machine operator performance. The Coefficient Alpha reliability for this subtest is generally very high and are mostly above .95 (Schuhfried, 1996; 2000a). A great number of studies on different validity concepts are available and they all show that the test is valid (Schuhfried, 1996; 2000a).

**Determination Unit** - This subtest assesses a candidate’s reaction speed, reactive stress tolerance and ability to demonstrate sustained multiple-choice reactions to rapidly changing stimuli (Schuhfried, 1996). Like the Cognitrone, this was administrated because of its conceptual links to road construction driver and operator requirements (Schuhfried, 1996; 2000a). Its focus is on the operators’ appropriate and fast responses in rapidly changing environments that may involve various stressors and stimuli (Schuhfried, 1996). Schuhfried (1996) reported an internal consistency reliability of 0.99 for the Determination Unit. In various criterion-related validity studies significant correlations between results on the Determination Unit and driving performance criteria were obtained (Karner, 2000; Karner & Neuwirth; 2000) One study, for example, showed significant correlations between the determination unit and the construct related RST3 test which measures Reactive Stress Tolerances, or RST (Karner & Neuwirth, 2000).

**Two-hand Coordination Speed and Accuracy** - This subtest assesses hand-eye and hand-hand coordination (Schuhfried, 1996; 2000b; 2000c; 2000d). It was included as a predictor in this study because of the two hand coordination requirements of driving and operating activities. Reported internal consistency reliabilities of the measures varied from 0.85 to 0.97 (Schuhfried, 1996). Karner and Neuwirth (2000) were able to show, that the performance in Two-hand Coordination tests with r=.50 is significantly associated with the assessment of driving performance. Furthermore these authors were able to demonstrate that persons with lower scores achieved significantly lower results in a standardized driving test.
The Time Movement/Anticipation test (ZBA) assesses an individual’s ability to imagine the effect of a movement and correctly estimate the movement of objects in space (Schuhfried, 1996; 2000c). The reliabilities at hand for the long form (inner consistency) show positive results above all for the time anticipation (Schuhfried, 1996; 2000c). In terms of the time anticipation section the median deviation time (total) Coefficient alpha reliabilities are respectively recorded as (0.98), median deviation time during a linear progression (0.92), median deviation time during a complex progression (.98), median deviation time during a sine-wave progression (0.92) (Schuhfried, 1996; 2000c). In terms of the motion anticipation section, the following Coefficient alpha reliability is recorded: median direction deviation (total) (0.76), median direction deviation during a linear progression (0.69), median direction deviation during a complex progression (0.72), median direction deviation during a sine-wave progression (0.62) (Schuhfried, 1996; 2000c). Currently there are validity studies available for a precursor of this test. What becomes clear from the results of an evaluation study using a driving test is that overestimating distances is more problematic than underestimating them. An evaluation of the test is ongoing (Karner & Neuwirth, 2000; Schuhfried, 1996; 2000c).

TRAM assessment battery. The TRAM assessment battery is a cognitive measure of the respondents’ overall learning potential (Taylor, 1999). The TRAM Learning Potential Test Battery (Taylor, 1999) was selected as a predictor in this study and was developed in South Africa by a South African, which enhances the overall face validity of the instrument. Essentially, the TRAM-I is a learning potential assessment instrument for candidates who fall in the illiterate and semi-literate ranges or who have had formal schooling and up to Grade 7 (Taylor, 1999). The TRAM-II is intended for application to testees with education ranging from Grade 8 to Grade 12.

The TRAM assessment battery was included in this study as a culture-fair measure of learning potential, which also portrays an indication of fluid intelligence (g_f) and general cognitive ability, or “g” (Taylor, 1994). The test requires candidates to translate symbols into other symbols, using a dictionary. The symbols are pictorial or quasi-geometric (Taylor, 1999). The symbols are translated using some underlying
rule (such as opposites – sun/moon; or symbols being used together – such as cultural artefacts). In Phase A1 of the test, candidates first complete the translation process by themselves. Thereafter they are given a lesson to explain the underlying rules, followed by the completion of Phase A2. Then they are given another test book and another dictionary to assess the transfer of skills. The final step is the completion of a memory test (Taylor, 1999). The total testing time is in the vicinity of two hours and forty-five minutes.

Taylor (1999) explained that scores are provided on the following TRAM dimensions: conceptual reasoning, automatisation, transfer, memory and understanding, speed and accuracy. Composite scores of respondents’ overall performance are also generated. This score incorporates scores of all six dimensions and each dimension is given an equal weight. Initially the overall assessment rating was the only TRAM predictor score used in this study. However, in search of any statistically significant positive relationship, TRAM subscales were investigated separately.

Considering the TRAM assessment, the **Speed** sub-scale refers to the rate at which the person does work of a routine nature and which imposes moderate intellectual demands. **Accuracy** is the proportion of the person’s work which is done correctly. **Learning rate or Automatization** is the rate at which the person learns a new task given practice and instruction. Two scores are calculated for automatization: the person’s after-the-lesson score in comparison to others who performed at a similar level on the initial session and the person’s improvement score in comparison to all norm group testees. The TRAM’s **Transfer** sub-scale scores the degree to which the person can apply and transfer existing knowledge to new challenges which differ somewhat from that he/she has encountered before. Two scores are calculated, analogous to the two learning rate scores: phase B performance relative to others who performed at a similar initial (phase A part 1) level and the person’s improvement in phase B relative to phase A part 1, as compared with all norm group testees. **Memory and Understanding** is the extent to which the person is able to grasp and understand tasks which place some intellectual demand on him/her.
Taylor (1999) reported reliability coefficients ranging from 0.62 to 0.95 for the various dimensions. In terms of validity, Taylor (1999) found that composite scores on the TRAM correlated significantly \( r=0.59; \ p=0.01 \) with academic performance in an ABET course and also with academic performance \( r=0.51; \ p=0.01 \) in N1 studies (NQF level 2, or grade 10) (Taylor, 1999).

**Performance appraisals.** Work performance was measured using the company’s performance appraisal system and the results of a separate supervisor ranking exercise. The company in which the research was conducted uses a top-down performance appraisal approach and all appraising managers received unit-standard aligned training on performance management, as part of the organisation’s management development programme. Appraisals are done on an on-going monthly basis and the scores are linked to production bonuses which are monthly. Appelbaum, Gilliland and Roy (2011) noted that adequate training must be provided to both the appraiser and the appraisee in order to avoid the many rating errors that are common in performance appraisals. Managers must also be given the opportunity to build the required relationship with these employees (Appelbaum, *et al*., 2011).

The company’s performance appraisal criteria included productivity, care for resources and a combined criterion called attitude and safety. It is a customised system designed specifically for the company - taking into account the particular needs within the industry. This customisation contributes to the system’s user-friendly interface (Sillup & Klimberg, 2010) and face validity among the company’s managers. Senior site management facilitated the driver and operator performance evaluations - which are an ongoing monthly process at the company.

**Supervisor rankings.** To further enhance the accuracy of work performance information, the researcher obtained data based on the facilitation of supervisor rankings (in addition to the performance appraisals). Forced ranking (FR) is a performance intervention, which can be defined as an evaluation method of forced distribution, where managers are required to distribute ratings for those being
evaluated, into a pre-specified performance distribution ranking (Cooper & Argyris, 1998). These rankings were done by the plant operations manager who oversees all drivers and operators, as opposed to the performance evaluations done by senior site management. The plant operations manager ranked each driver and operator, using the paired comparisons method (Cascio, 1998). He was asked to decide which operator in every pair of operators would be selected if the working conditions were particularly difficult (due to congestion in the loading areas, wet road conditions, project completion time-constraints etc.) with number 1 being the best operator, number 2 the second best operator and so forth. The groups were then refined to present the best-, worst- and other drivers for the empirical investigation.

Research Procedure

This study consisted of two distinct phases, namely an explorative literature review and an empirical study.

Phase 1: Literature Review. In the explorative literature review, the researcher endeavoured to determine the following:

1. The theory of psychomotor ability and the conceptualisation of psychomotor ability as a measurement construct.
2. The theory of learning potential and the conceptualisation of learning potential as a measurement construct.
3. The exploration of job performance and the measurement thereof.

Phase 2: Empirical Study. The empirical study involved a quantitative investigation into the statistical relationship between psychomotor ability and learning potential assessment results and job performance.

In step 1 of the empirical study, all drivers and machine operators were assessed on the Vienna Test System (Schuhfried, 1996) and the TRAM assessment battery (Taylor, 1999). In step 2, the researcher gathered the psychomotor ability and learning potential data necessary to conduct the empirical study. In step 3, data from
the driver and machine operator’s performance appraisal forms and supervisor rankings was collected from official company documentation. Notably, assessment data was gathered first and, six months thereafter (in the month of September), the criterion data was collected – as the study was concerned specifically with predictive validity and not with concurrent validity. In phase 4, the collected data was statistically analysed with computer software (SPSS ver. 20) and the results interpreted.

**Statistical Analysis**

The study followed a descriptive approach (Terre Blanche *et al.*, 2006) and was aimed at determining and describing the relationship between the selected variables. Correlations (Terre Blanche *et al.*, 2006) were used to report on whether, and to what extent, a statistically significantly (and positive) relationship between variables exist. Regression analysis (Terre Blanche *et al.*, 2006) was to be used to report on whether the independent variables (psychomotor ability and learning potential) statistically significantly predicted the dependent variable (work performance). The correlation coefficient was used as a measure to interpret the results and to specifically evaluate the statistical and practical significance of the relationship (Cohen, 1992; Howell, 2004; Terre Blanche *et al.*, 2006) between psychomotor ability, learning potential, and work performance. Subsequent to determining the correlation coefficients, the validity of psychomotor ability and learning potential as predictors of work performance was anticipated to be evaluated by means of statistical regression. Regression refers to the prediction of one dependent variable based on knowledge of the levels of one or more independent variable(s) (Howell, 2004). Singular regression would be used to respectively examine the predictive validity of psychomotor ability and learning potential in predicting work performance. Multiple regression should be used to analyse the predictive validity of both psychomotor ability and learning potential for predicting the work performance. Multiple regression is a method of studying the separate and collective contributions of several independent variables (psychomotor ability and learning potential) to the variation of a dependent variable (work performance) (Howell, 2004; Terre Blanche *et al.*, 2006).
RESULTS

Overview of the Descriptive Statistics

Descriptive statistics in respect of the predictor variables and dependent variables are presented in Table 1. Usually the mean is the best measure for describing a set of data with a single value (Hubbard, 2004; Terre Blanche et al., 2006). The aim of descriptive statistics is only to describe or analyse data, and not to draw conclusions or make inferences about the larger group (Foxcroft & Roodt, 2009; Terre Blanche et al., 2006). In Table 1, the research results are described in terms minimum, maximum, median, mean and standard deviation values.

Descriptive statistics usually include measures of central tendency (e.g. means, medians) and variance (e.g. standard deviation) (Terre Blanche, et al., 2006; Morgan, Reichert & Harrison, 2002). The standard deviation is a measure that is used to quantify the amount of variation or dispersion of a set of data values (Howell, 2004; Terre Blanche, et al., 2006). A standard deviation close to 0 indicates that the data points tend to be very close to the mean (also called the expected value) of the set, while a high standard deviation indicates that the data points are spread out over a wider range of values (Terre Blanche, et al., 2006). Statistical skewness is defined in terms of this relationship: positive/right skew means the mean is greater than (to the right of) the median, while negative/left skew means the mean is less than (to the left of) the median (Howell, 2004; Terre Blanche, et al., 2006).

In many datasets the values deviate from the mean value due to chance only and such datasets are said to display a normal distribution (Howell, 2004; Terre Blanche, et al., 2006). In a dataset with a normal distribution most of the values are clustered around the mean while relatively few values tend to be extremely high or extremely low. Many natural phenomena display a normal distribution (Babbie, 2007; Hubbard, 2004). In Table 1, the TRAM assessment scores, the various VTS subscales (independent variables) and the collected performance management data for each month (dependent variable) are summarised. Initially, criterion data for only one
month (September) was used in the statistics, but because of unexpected results the researcher investigated criterion data over nine months.

The month of September was exactly six months after the assessments were administered and it was used by the researcher to investigate predictive validity in particular. In Table 1, this data is represented by the item September under the subheading Performance Scores. However, since the performance data are captured on a monthly basis it was decided to include data from a total of nine months (June to February) and to use an average performance score (represented by Average Score in table 1) across all nine months. This was done so that the researcher could obtain a broader and more comprehensive view into any possible statistically significant relationships between the independent- and dependent variables of the research. It was anticipated that a Total Score would bring more variance to the criterion score which may improve the statistical significance of the correlations. Similarly, the researcher looked at all the sub-scales of the VTS assessments in addition to using a single VTS (total) score.

The researcher opted for a broader investigation into the respective VTS assessment scores and separately investigated all the subscales. The VTS sub-scales and dimensions are described below. The VTS Cognitrone score constitutes Concentration Speed and Concentration Accuracy (Schuhfried, 1996). The VTS Determination score constitutes the Low Stress Speed, Low Stress Accuracy, High Stress Speed, High Stress Accuracy, Medium Stress Speed and Medium Stress Accuracy sub-scales (Schuhfried, 1996). The Two-hand Coordination Speed and Accuracy score was investigated by analysing both speed and accuracy and is presented separately in Table 1 as 2-Hand Accuracy Coordination and 2-Hand Speed Coordination (Schuhfried, 1996). Similarly the Time Movement/Anticipation Test (ZBA) of the VTS was separated into Time Anticipation and Motion Anticipation (Schuhfried, 1996). The researcher then investigated the VTS’s Total Speed, Total Accuracy and a combination of the two, VTS Total Speed and Accuracy (Schuhfried, 1996) in Table 1.
<table>
<thead>
<tr>
<th>Sub-Scale</th>
<th>Min</th>
<th>Max</th>
<th>Median</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Total Score</th>
<th>Coefficient Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Concentration Speed</strong></td>
<td>29</td>
<td>80</td>
<td>55</td>
<td>54</td>
<td>12</td>
<td>5140</td>
<td>0.873</td>
</tr>
<tr>
<td><strong>Concentration Accuracy</strong></td>
<td>27</td>
<td>66</td>
<td>51</td>
<td>51</td>
<td>9</td>
<td>4854</td>
<td>0.880</td>
</tr>
<tr>
<td><strong>Low Stress Speed</strong></td>
<td>32</td>
<td>69</td>
<td>52</td>
<td>51</td>
<td>8</td>
<td>4839</td>
<td>0.852</td>
</tr>
<tr>
<td><strong>Low Stress Accuracy</strong></td>
<td>31</td>
<td>66</td>
<td>51</td>
<td>50</td>
<td>9</td>
<td>4758</td>
<td>0.852</td>
</tr>
<tr>
<td><strong>Low Stress Speed</strong></td>
<td>27</td>
<td>60</td>
<td>51</td>
<td>50</td>
<td>6</td>
<td>4713</td>
<td>0.860</td>
</tr>
<tr>
<td><strong>Low Stress Accuracy</strong></td>
<td>34</td>
<td>60</td>
<td>49</td>
<td>48</td>
<td>6</td>
<td>4586</td>
<td>0.859</td>
</tr>
<tr>
<td><strong>Medium Stress Speed</strong></td>
<td>32</td>
<td>60</td>
<td>51</td>
<td>50</td>
<td>6</td>
<td>4734</td>
<td>0.859</td>
</tr>
<tr>
<td><strong>Medium Stress Accuracy</strong></td>
<td>34</td>
<td>71</td>
<td>49</td>
<td>49</td>
<td>7</td>
<td>4630</td>
<td>0.854</td>
</tr>
<tr>
<td><strong>2-Hand Speed Coordination</strong></td>
<td>22</td>
<td>73</td>
<td>49</td>
<td>50</td>
<td>8</td>
<td>4749</td>
<td>0.885</td>
</tr>
<tr>
<td><strong>2-Hand Accuracy Coordination</strong></td>
<td>29</td>
<td>80</td>
<td>47</td>
<td>49</td>
<td>11</td>
<td>4646</td>
<td>0.874</td>
</tr>
<tr>
<td><strong>Time</strong></td>
<td>20</td>
<td>69</td>
<td>50</td>
<td>49</td>
<td>10</td>
<td>4691</td>
<td>0.882</td>
</tr>
<tr>
<td><strong>Motion Anticipation</strong></td>
<td>20</td>
<td>80</td>
<td>49</td>
<td>50</td>
<td>11</td>
<td>4722</td>
<td>0.883</td>
</tr>
<tr>
<td><strong>Total Speed</strong></td>
<td>33</td>
<td>62</td>
<td>51</td>
<td>50</td>
<td>6</td>
<td>4762</td>
<td>0.855</td>
</tr>
<tr>
<td><strong>Total Accuracy</strong></td>
<td>37</td>
<td>63</td>
<td>50</td>
<td>49</td>
<td>5</td>
<td>4698</td>
<td>0.855</td>
</tr>
<tr>
<td><strong>Total Speed &amp; Accuracy</strong></td>
<td>38</td>
<td>62</td>
<td>50</td>
<td>50</td>
<td>5</td>
<td>4714</td>
<td>0.854</td>
</tr>
</tbody>
</table>

**Learning Potential (TRAM)**

<table>
<thead>
<tr>
<th>Sub-Scale</th>
<th>Min</th>
<th>Max</th>
<th>Median</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Total Score</th>
<th>Coefficient Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Speed</strong></td>
<td>10</td>
<td>90</td>
<td>50</td>
<td>49</td>
<td>16</td>
<td>4690</td>
<td>0.778</td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td>10</td>
<td>90</td>
<td>50</td>
<td>54</td>
<td>17</td>
<td>5110</td>
<td>0.830</td>
</tr>
<tr>
<td><strong>Learn Rate</strong></td>
<td>20</td>
<td>90</td>
<td>60</td>
<td>57</td>
<td>19</td>
<td>5420</td>
<td>0.799</td>
</tr>
<tr>
<td><strong>Transfer</strong></td>
<td>10</td>
<td>90</td>
<td>50</td>
<td>55</td>
<td>20</td>
<td>5220</td>
<td>0.767</td>
</tr>
<tr>
<td><strong>Memory</strong></td>
<td>0</td>
<td>90</td>
<td>40</td>
<td>33</td>
<td>29</td>
<td>3140</td>
<td>0.691</td>
</tr>
<tr>
<td><strong>OAR</strong></td>
<td>15</td>
<td>84</td>
<td>51</td>
<td>53</td>
<td>12</td>
<td>5062</td>
<td></td>
</tr>
</tbody>
</table>

**Performance Scores**

<table>
<thead>
<tr>
<th>Sub-Scale</th>
<th>Min</th>
<th>Max</th>
<th>Median</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Total Score</th>
<th>Coefficient Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>June</strong></td>
<td>24</td>
<td>100</td>
<td>61</td>
<td>61</td>
<td>13</td>
<td>5834</td>
<td>0.715</td>
</tr>
<tr>
<td><strong>July</strong></td>
<td>29</td>
<td>100</td>
<td>62</td>
<td>66</td>
<td>19</td>
<td>6295</td>
<td>0.690</td>
</tr>
<tr>
<td><strong>August</strong></td>
<td>39</td>
<td>100</td>
<td>65</td>
<td>69</td>
<td>19</td>
<td>6535</td>
<td>0.700</td>
</tr>
<tr>
<td><strong>September</strong></td>
<td>38</td>
<td>100</td>
<td>68</td>
<td>71</td>
<td>19</td>
<td>6767</td>
<td>0.750</td>
</tr>
<tr>
<td><strong>October</strong></td>
<td>32</td>
<td>100</td>
<td>68</td>
<td>68</td>
<td>15</td>
<td>6425</td>
<td>0.714</td>
</tr>
<tr>
<td><strong>November</strong></td>
<td>5</td>
<td>100</td>
<td>52</td>
<td>50</td>
<td>25</td>
<td>4728</td>
<td>0.698</td>
</tr>
<tr>
<td><strong>December</strong></td>
<td>23</td>
<td>100</td>
<td>62</td>
<td>65</td>
<td>19</td>
<td>6183</td>
<td>0.687</td>
</tr>
<tr>
<td><strong>January</strong></td>
<td>42</td>
<td>100</td>
<td>76</td>
<td>76</td>
<td>16</td>
<td>7202</td>
<td>0.734</td>
</tr>
<tr>
<td><strong>February</strong></td>
<td>13</td>
<td>100</td>
<td>73</td>
<td>70</td>
<td>21</td>
<td>6657</td>
<td>0.759</td>
</tr>
<tr>
<td><strong>Average Score</strong></td>
<td>48</td>
<td>86</td>
<td>65</td>
<td>66</td>
<td>10</td>
<td>6292</td>
<td>0.670</td>
</tr>
<tr>
<td><strong>Total Score</strong></td>
<td>436</td>
<td>778</td>
<td>586</td>
<td>596</td>
<td>89</td>
<td>56626</td>
<td></td>
</tr>
</tbody>
</table>
Table 1 also includes the coefficient alpha values for each of the sub-dimensions. Coefficient alpha values are usually positive and close to 1. Values lower than 0.70 show inadequate reliability – which means the test results cannot be used (Terre Blanche et al., 2006). For the VTS test, an overall coefficient alpha value of 0.878 was obtained, while a value of 0.847 was obtained for the TRAM test. These values indicate a high reliability. Individual coefficient alpha values shown represent what the overall value would be should the sub-test be removed. This aids in deciding which sub-tests may be removed in future research to obtain a better result.

Opting for a broader approach meant the researcher had to investigate the variance between the dimensions of each particular test. In statistics and research, internal consistency is typically a measure based on the correlations between different items on the same test (or the same sub-scale on a larger test) (Gregory, 2007; Terre Blanche, et al., 2006). It measures whether several items that propose to measure the same general construct produce similar scores (Terre Blanche, et al., 2006). Techniques that involve analysis of item variances are more appropriately termed measures of internal consistency, since they indicate the degree to which the various items on a test are inter-correlated (Cascio & Aguinis, 2005; Terre Blanche, et al., 2006). Tables 2, 3 and 4 below show the intercorrelations for each test and for the sample’s performance scores, along with the average correlation for each.
Table 2

**Intercorrelations and P-values for VTS Sub-Scores**

<table>
<thead>
<tr>
<th></th>
<th>CS</th>
<th>CA</th>
<th>LSS</th>
<th>LSA</th>
<th>HSS</th>
<th>HSA</th>
<th>MSS</th>
<th>MSA</th>
<th>2HS</th>
<th>2HA</th>
<th>TIM</th>
<th>MOT</th>
<th>TS</th>
<th>TA</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>-0.04</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSS</td>
<td>0.52</td>
<td>0.17</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSA</td>
<td>0.47</td>
<td>0.23</td>
<td>0.90</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSS</td>
<td>0.49</td>
<td>0.08</td>
<td>0.71</td>
<td>0.56</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSA</td>
<td>0.40</td>
<td>0.21</td>
<td>0.76</td>
<td>0.78</td>
<td>0.67</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSS</td>
<td>0.51</td>
<td>0.22</td>
<td>0.78</td>
<td>0.71</td>
<td>0.83</td>
<td>0.71</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSA</td>
<td>0.42</td>
<td>0.22</td>
<td>0.84</td>
<td>0.83</td>
<td>0.69</td>
<td>0.85</td>
<td>0.83</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2HS</td>
<td>0.11</td>
<td>-0.06</td>
<td>0.09</td>
<td>0.08</td>
<td>0.20</td>
<td>0.07</td>
<td>0.15</td>
<td>0.07</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2HA</td>
<td>0.18</td>
<td>0.14</td>
<td>0.36</td>
<td>0.38</td>
<td>0.27</td>
<td>0.31</td>
<td>0.30</td>
<td>0.38</td>
<td>-0.10</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TIM</td>
<td>0.07</td>
<td>0.06</td>
<td>0.11</td>
<td>0.15</td>
<td>0.09</td>
<td>0.04</td>
<td>0.08</td>
<td>0.12</td>
<td>0.25</td>
<td>0.28</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOT</td>
<td>0.13</td>
<td>0.22</td>
<td>0.25</td>
<td>0.17</td>
<td>0.17</td>
<td>0.13</td>
<td>0.17</td>
<td>0.19</td>
<td>-0.23</td>
<td>0.15</td>
<td>0.05</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TS</td>
<td>0.10</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.06</td>
<td>0.05</td>
<td>0.05</td>
<td>0.06</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>TA</td>
<td>0.35</td>
<td>0.49</td>
<td>0.74</td>
<td>0.77</td>
<td>0.54</td>
<td>0.70</td>
<td>0.65</td>
<td>0.77</td>
<td>0.00</td>
<td>0.65</td>
<td>0.43</td>
<td>0.49</td>
<td>0.70</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Average correlation = 0.470
Std. deviation = 0.349

Note. CS=Concentration Speed. CA=Concentration Accuracy. LSS=Low Stress Speed. LSA=Low Stress Accuracy. HSS=High Stress Speed. HSA=High Stress Accuracy. MSS=Medium Stress Speed. MSA=Medium Stress Accuracy. 2HS=Two Hand Coordination Speed. 2HA=Two Hand Coordination Accuracy. TIM=Time Anticipation. MOT=Motion Anticipation. TS=Total Speed. TA=Total Accuracy.

---

Table 3

**Intercorrelations and P-values for TRAM Sub-Scores**

<table>
<thead>
<tr>
<th></th>
<th>Speed</th>
<th>Accuracy</th>
<th>Learn Rate</th>
<th>Transfer</th>
<th>Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.31</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learn Rate</td>
<td>0.49</td>
<td>0.28</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transfer</td>
<td>0.53</td>
<td>0.24</td>
<td>0.60</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Memory</td>
<td>0.39</td>
<td>0.26</td>
<td>0.23</td>
<td>0.45</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Average Correlation = 0.581
Std. deviation = 0.292

Note. *p < 0.05
### Table 4

**Intercorrelations and P-values for Monthly Performance Scores**

<table>
<thead>
<tr>
<th></th>
<th>June</th>
<th>July</th>
<th>Aug</th>
<th>Sept</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
<th>Jan</th>
<th>Feb</th>
</tr>
</thead>
<tbody>
<tr>
<td>June</td>
<td>1.00</td>
<td>0.48</td>
<td>0.21</td>
<td>0.11</td>
<td>0.25</td>
<td>0.25</td>
<td>0.31</td>
<td>0.03</td>
<td>0.09</td>
</tr>
<tr>
<td>July</td>
<td>1.00</td>
<td>0.37</td>
<td>1.00</td>
<td>0.34</td>
<td>0.15</td>
<td>0.51</td>
<td>0.41</td>
<td>0.22</td>
<td>0.26</td>
</tr>
<tr>
<td>August</td>
<td>0.21</td>
<td>1.00</td>
<td>0.21</td>
<td>1.00</td>
<td>0.25</td>
<td>0.34</td>
<td>0.41</td>
<td>0.15</td>
<td>0.13</td>
</tr>
<tr>
<td>September</td>
<td>0.11</td>
<td>0.15</td>
<td>-0.03</td>
<td>0.10</td>
<td>0.07</td>
<td>0.34</td>
<td>0.15</td>
<td>0.15</td>
<td>0.13</td>
</tr>
<tr>
<td>October</td>
<td>0.25</td>
<td>0.31</td>
<td>0.51</td>
<td>0.43</td>
<td>1.00</td>
<td>0.34</td>
<td>0.41</td>
<td>0.15</td>
<td>0.13</td>
</tr>
<tr>
<td>November</td>
<td>0.25</td>
<td>0.34</td>
<td>0.41</td>
<td>0.49</td>
<td>0.53</td>
<td>1.00</td>
<td>0.34</td>
<td>0.15</td>
<td>0.13</td>
</tr>
<tr>
<td>December</td>
<td>0.03</td>
<td>0.22</td>
<td>0.15</td>
<td>0.11</td>
<td>0.08</td>
<td>0.08</td>
<td>0.13</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>January</td>
<td>0.09</td>
<td>0.26</td>
<td>0.00</td>
<td>0.14</td>
<td>0.00</td>
<td>0.01</td>
<td>0.18</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Average correlation = 0.360  
Std. deviation = 0.359

Note. *p < 0.05

Intercorrelations indicate a connection between the scores of the different sub-scales for each person undergoing the test. For example, one might expect a person who scores very high in some of the tests to maintain this high-scoring trend. If the correlation is low (<0.75), there is an indication that such trends are not maintained. This may indicate that the various sub-tests aim to test a large range of skills.

From Tables 2, 3 and 4, it is evident that the average correlation for the performance management scores is very low for the months over which the group was assessed. An obvious explanation for the absence of a statistically significant relationship is the subjectivity of the company’s performance appraisal system, which again contributes to the criterion problem. For example, based on test scores employees may perform very well while their actual job performance may be lacking, or the line manager fails to assign an accurate performance score. There is always a possibility of bias on the side of the rating supervisor. This may be due to both short- and long term factors such as isolated incidents of abnormal performance (good or bad) or simply a clash of personalities between an operator and his or her supervisor.
The TRAM mean scores of the sample group (Group 1) (Table 1) were subsequently compared to a similar group (Group 2). Most of the results obtained from Group 2, however, were considerably higher than the TRAM mean scores of Group 1. The TRAM mean scores for Group 2 are shown in Table 5. Group 2 was based on a sample of 208 employees of large organisations operating in the Gauteng province (Taylor, 1999), with a mean of 10.9 years of education.

Table 5

TRAM Mean Scores of the Comparison Group (Group 2) (n = 208)

<table>
<thead>
<tr>
<th>TRAM Dimensions</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed</td>
<td>454.26</td>
<td>272.95</td>
</tr>
<tr>
<td>Accuracy</td>
<td>60.93</td>
<td>18.67</td>
</tr>
<tr>
<td>Learn Rate</td>
<td>100.00</td>
<td>10.83</td>
</tr>
<tr>
<td>Transfer</td>
<td>100.00</td>
<td>12.18</td>
</tr>
<tr>
<td>Memory</td>
<td>16.40</td>
<td>8.36</td>
</tr>
</tbody>
</table>

Note. Std Dev=Standard Deviation

The researcher investigated the educational level of the sample group to effectively compare the two groups. This information is given in Table 6. The statistics for the entire comparison group are shown alongside the split for comparison purposes.
Table 6

TRAM Scores According to Level of Education for the Sample Group (n=95)

<table>
<thead>
<tr>
<th>Sample Group</th>
<th>n</th>
<th>Min</th>
<th>Max</th>
<th>Median</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRAM</td>
<td>95</td>
<td>15</td>
<td>84</td>
<td>51</td>
<td>53</td>
<td>12</td>
<td>5062</td>
</tr>
<tr>
<td>VTS</td>
<td>95</td>
<td>38</td>
<td>62</td>
<td>50</td>
<td>50</td>
<td>5</td>
<td>4714</td>
</tr>
<tr>
<td>PM</td>
<td>95</td>
<td>48</td>
<td>86</td>
<td>65</td>
<td>66</td>
<td>10</td>
<td>6292</td>
</tr>
</tbody>
</table>

Employees in the Sample Group with Less than 7 Years Education

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Min</th>
<th>Max</th>
<th>Median</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRAM</td>
<td>20</td>
<td>15</td>
<td>69</td>
<td>45</td>
<td>44</td>
<td>13</td>
<td>888</td>
</tr>
<tr>
<td>VTS</td>
<td>20</td>
<td>40</td>
<td>52</td>
<td>46</td>
<td>46</td>
<td>4</td>
<td>921</td>
</tr>
<tr>
<td>PM</td>
<td>20</td>
<td>52</td>
<td>83</td>
<td>68</td>
<td>68</td>
<td>9</td>
<td>1363</td>
</tr>
</tbody>
</table>

Employees in the Sample Group with 7-9 Years Education

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Min</th>
<th>Max</th>
<th>Median</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRAM</td>
<td>20</td>
<td>35</td>
<td>73</td>
<td>50</td>
<td>52</td>
<td>10</td>
<td>1045</td>
</tr>
<tr>
<td>VTS</td>
<td>20</td>
<td>40</td>
<td>62</td>
<td>46</td>
<td>47</td>
<td>6</td>
<td>945</td>
</tr>
<tr>
<td>PM</td>
<td>20</td>
<td>48</td>
<td>83</td>
<td>72</td>
<td>69</td>
<td>12</td>
<td>1372</td>
</tr>
</tbody>
</table>

Employees in the Sample Group with 10-12 Years Education

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Min</th>
<th>Max</th>
<th>Median</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRAM</td>
<td>55</td>
<td>38</td>
<td>84</td>
<td>56</td>
<td>57</td>
<td>11</td>
<td>3129</td>
</tr>
<tr>
<td>VTS</td>
<td>55</td>
<td>38</td>
<td>62</td>
<td>52</td>
<td>52</td>
<td>5</td>
<td>2848</td>
</tr>
<tr>
<td>PM</td>
<td>55</td>
<td>50</td>
<td>86</td>
<td>64</td>
<td>65</td>
<td>9</td>
<td>3558</td>
</tr>
</tbody>
</table>

Note. Std Dev=Standard Deviation

The average years of education were only slightly lower for the sample group in comparison to Group 2. The sample group’s average years of education was 8.9 years and the other group’s average years of education was 10.9 years. It is not probable that this difference in years of education could contribute to the large dissimilarities in the mean scores of the two groups. Especially because the TRAM learning potential assessment was designed in the South African context with a specific aim to level the playing field and not to discriminate unfairly in terms of an individual’s learning background (Taylor, 1999). Learning potential is a measure of what the person could achieve in the future with regards to mastering new or cognitively challenging material (De Beer, 2006; Taylor, 1999). Tests of specific
abilities, on the other hand, assess the person’s current mastery of certain cognitive domains (Gregory, 2007), which reflects the opportunities to learn that the person had in the past (Claassen, 1997; Pretorius et al., 2009; Taylor, 1999).

The intercorrelations between educational level and the respective test scores were TRAM (0.482; \( p = 0.000 \)), VTS (0.494; \( p = 0.000 \)) and PM (-0.153; \( p = 0.134 \)). This means that the correlations between the level of education and the TRAM and VTS scores tend towards a positive relationship, while there is a tendency towards a negative relationship, albeit small, between the level of education and performance scores.

South African education has not reached a level where all individuals have an opportunity to deploy their full potential (Coetzer, Battisti, Jurado, & Massey, 2011). Unfortunately, even learning potential is somewhat crimped by poor education (Claassen, 1997; Prinsloo, 2013). Poor schools tend to emphasize mechanical or habitual repetition in learning whereas better schools encourage the learner to think for him or herself and use learning strategies (Coetzee, Botha, Kiley, Truman & Tshilongamulenzhe, 2012; Strauss & du Toit, 2010). Nevertheless, the assessment of learning potential is still a preferable way to evaluate a person’s likely effectiveness in cognitively challenging environments (De Beer, 2005; Taylor, 1999). It is therefore unlikely that the two years educational difference between the two groups could contribute to the large difference in mean scores.

The TRAM mean scores of the sample group were compared against a second comparison group (Group 3). Group 3 consists of 195 apprentice applicants at the Johannesburg Municipality (Taylor, 1999). The average education of the municipality was 9.3 years of education and the average age was 22.74. Very similar results were derived from the second comparison group (Group 3) – the mean scores on most of the TRAM sub-scales were considerably lower for the sample group of this study. Group 3’s mean scores are summarised in Table 7 below.
The sample group (compared to the two other groups) scored considerably higher on one particular TRAM sub-scale – Memory. The understanding and memorization of novel material is obviously important in any kind of learning environment or even new job contexts. Candidates who try to understand the material rather than simply use it in a rote fashion tend to obtain higher scores (Taylor, 1999). High scores in this section of the test are attained by those who have succeeded in identifying the underlying principles in the material and have then made an effort to commit it to memory. Low scores are typically attained by those who worked mechanically with the material and did not cognise it (Taylor, 1994).

Memory declines with age (Barlow & Durand, 2005) and one would expect older individuals to have lower scores on Memory compared to younger individuals. Scientists now believe that memory loss could start as early as 45 years of age, as opposed to 65 years mentioned in previous studies (Barlow & Durand, 2005; Nevid, Rathus & Greene, 2008). The sample group of drivers and operators was, on average, almost double the age (42) of the second comparison group (23). Yet, the older test-takers scored much higher on the TRAM Memory sub-scale than their younger counterparts nineteen years their junior. Clearly it is highly unlikely that age had any influence on the difference in Memory scores. Table 8 below depicts the age profile of the drivers and operators in the study.
Table 8

**TRAM and VTS Test Results Split According to Age Group (n=95)**

<table>
<thead>
<tr>
<th>Sample Group</th>
<th>n</th>
<th>Min</th>
<th>Max</th>
<th>Median</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRAM</td>
<td>95</td>
<td>15</td>
<td>84</td>
<td>51</td>
<td>53</td>
<td>12</td>
<td>5062</td>
</tr>
<tr>
<td>VTS</td>
<td>95</td>
<td>38</td>
<td>62</td>
<td>50</td>
<td>50</td>
<td>5</td>
<td>4714</td>
</tr>
<tr>
<td>PM</td>
<td>95</td>
<td>48</td>
<td>86</td>
<td>65</td>
<td>66</td>
<td>10</td>
<td>6292</td>
</tr>
</tbody>
</table>

**Employees in the sample group younger than 42 years**

<table>
<thead>
<tr>
<th>Sample Group</th>
<th>n</th>
<th>Min</th>
<th>Max</th>
<th>Median</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRAM</td>
<td>44</td>
<td>38</td>
<td>84</td>
<td>57</td>
<td>58</td>
<td>11</td>
<td>2533</td>
</tr>
<tr>
<td>VTS</td>
<td>44</td>
<td>40</td>
<td>62</td>
<td>52</td>
<td>52</td>
<td>5</td>
<td>2275</td>
</tr>
<tr>
<td>PM</td>
<td>44</td>
<td>50</td>
<td>86</td>
<td>63</td>
<td>65</td>
<td>10</td>
<td>2849</td>
</tr>
</tbody>
</table>

**Employees in the sample group older than 42 years**

<table>
<thead>
<tr>
<th>Sample Group</th>
<th>n</th>
<th>Min</th>
<th>Max</th>
<th>Median</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRAM</td>
<td>51</td>
<td>15</td>
<td>80</td>
<td>47</td>
<td>50</td>
<td>12</td>
<td>2529</td>
</tr>
<tr>
<td>VTS</td>
<td>51</td>
<td>38</td>
<td>62</td>
<td>48</td>
<td>48</td>
<td>5</td>
<td>2439</td>
</tr>
<tr>
<td>PM</td>
<td>51</td>
<td>48</td>
<td>84</td>
<td>68</td>
<td>68</td>
<td>10</td>
<td>3443</td>
</tr>
</tbody>
</table>

Apart from memory, the sample group had lower mean scores (Table 1) compared to both comparison groups (Table 5 & 7). As the name suggests, a person’s score on the **speed** dimension indicates the rate at which he or she is likely to work when doing tasks of moderate intellectual difficulty. The tasks are repetitive, and each one involved applying a number of relatively straightforward steps. The speed score reflects work processing speed, irrespective of whether the answer obtained is correct or not. The sample group is therefore likely to work slower than the two comparison groups. Lower scores on the **accuracy** dimensions means the sample group probably made many more errors compared to the comparison groups due to carelessness and oversights. The sample group’s lower **learning rate** (automatization) scores indicate that they were more sluggish to become quicker and more efficient at executing a novel task – they made insufficient use of their exposure to the test material, compared to the two other groups. The sample group’s lower **transfer** dimension indicates that they tend to see each problem as completely new and have to refer to others (such as supervisors) for help and guidance. The capacity to transfer is essential to effective performance in a changing...
environment, because almost all challenges that face an employee every day are not identical to previous challenges and problems that he or she may have encountered and already mastered (Taylor, 1999).

**Test of Normality**

Normality tests are used to determine if a data set is well-modelled by a normal distribution and to compute how likely it is for a random variable underlying the data set to be normally distributed (Babbie, 2007). More precisely, these tests are a form of model selection, and can be interpreted in several ways, depending on one’s interpretations of probability (Morgan *et al.*., 2002). Normality tests compare the shape of a sample distribution to the shape of the normal curve (Morgan *et al.*., 2002). It therefore assumes, if the sample is normally shaped, the population from which it came can be assumed to be normally distributed for the particular variable. For this purpose the Shapiro-Wilk test was used (Morgan *et al.*., 2002). The null-hypothesis of this test is that the population is normally distributed. Thus, if the $p$-value is less than the chosen significant level ($p = 0.05$), then the null hypothesis is rejected and there is evidence that the data tested are not from a normally distributed population (Babbie, 2007; Hubbard, 2004; Morgan *et al.*., 2002).

**Table 9**

*Results Obtained from the Shapiro-Wilk Normality Test (n=95)*

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Statistic</th>
<th>Sig. (p)</th>
<th>Normal Distr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concentration Speed</td>
<td>0.987</td>
<td>0.442</td>
<td>Yes</td>
</tr>
<tr>
<td>Concentration Accuracy</td>
<td>0.960</td>
<td>0.005</td>
<td>No</td>
</tr>
<tr>
<td>Low Stress Speed</td>
<td>0.990</td>
<td>0.716</td>
<td>Yes</td>
</tr>
<tr>
<td>Low Stress Accuracy</td>
<td>0.968</td>
<td>0.020</td>
<td>No</td>
</tr>
<tr>
<td>Low Stress Speed</td>
<td>0.918</td>
<td>0.000</td>
<td>No</td>
</tr>
<tr>
<td>Low Stress Accuracy</td>
<td>0.963</td>
<td>0.009</td>
<td>No</td>
</tr>
<tr>
<td>Medium Stress Speed</td>
<td>0.964</td>
<td>0.010</td>
<td>No</td>
</tr>
<tr>
<td>Medium Stress Accuracy</td>
<td>0.970</td>
<td>0.029</td>
<td>No</td>
</tr>
<tr>
<td>2-Hand Speed Coordination</td>
<td>0.984</td>
<td>0.281</td>
<td>Yes</td>
</tr>
<tr>
<td>2-Hand Accuracy Coordination</td>
<td>0.962</td>
<td>0.007</td>
<td>No</td>
</tr>
<tr>
<td>Time Anticipation</td>
<td>0.951</td>
<td>0.001</td>
<td>No</td>
</tr>
<tr>
<td>Motion Anticipation</td>
<td>0.964</td>
<td>0.011</td>
<td>No</td>
</tr>
<tr>
<td>Total Speed</td>
<td>0.982</td>
<td>0.223</td>
<td>Yes</td>
</tr>
<tr>
<td>Total Accuracy</td>
<td>0.990</td>
<td>0.735</td>
<td>Yes</td>
</tr>
<tr>
<td>Total Speed &amp; Accuracy</td>
<td>0.989</td>
<td>0.597</td>
<td>Yes</td>
</tr>
</tbody>
</table>
The main conclusion drawn from Table 9 is that the data is generally not normally distributed - a result which directed the use of nonparametric statistics (Morgan et al., 2002) for further statistical analyses. Only 8 of a possible 32 sub-dimensions across all the assessment results were normally distributed. These were concentration speed, low stress speed, 2-hand speed coordination, total speed, total accuracy and total speed & accuracy, June’s performance score and October’s performance scores. Nonparametric statistics are frequently used in the social sciences to examine the differences or associations for nominal and ordinal level data (Hubbard, 2004; Morgan et al., 2002).

A significance value of 0.05 was used for all of the statistical analyses. This particular value was chosen to make an event very unlikely to occur just by random sampling variation. P-values evaluate how well the sample data support the null hypotheses (Babbie, 2007; Hubbard, 2004; Morgan et al., 2002), that is, they measure how compatible the data are with the null hypotheses (Hubbard, 2004).
• A small \( p \)-value (\( p \leq 0.05 \)) indicates strong evidence against the null hypothesis, so that the researcher may reject the null hypothesis (Morgan et al., 2002).

• A large \( p \)-value (\( p > 0.05 \)) indicates weak evidence against the null hypothesis, so that the researcher cannot reject the null hypothesis (Morgan et al., 2002).

• \( p \)-values very close to the cut-off (0.05) are considered to be marginal (could go either way) (Morgan et al., 2002).

The majority of the results shown in table 8 are statistically non-significant and therefore fail to reject the null hypothesis regarding the normality of the distribution of the values of the different variables

**Correlations**

Correlations were used in an attempt to represent the relationships between variables by means of a correlation coefficient (Hubbard, 2004; Terre Blanche, et al., 2006). The correlation coefficient signifies the strength of co-variation between two variables (Terre Blanche, et al., 2006). The aim of correlation statistics is therefore to describe the strength and direction of the linear relationship that exists between two measured variables (Babbie, 2007; Cooper & Schindler, 2001; Terre Blanche, et al., 2006). Both the Pearson- and Spearman's Rho correlation coefficients (Morgan et al., 2002) were calculated to determine the degree of relationship amongst the predictor and criterion variables. The statistical significance of all the correlations was determined.

Correlations can vary in magnitude from -1 to +1, with -1 indicating a perfect negative linear relationship, (as one variable increases/decreases, the other decreases/increases respectively), +1 indicating a perfect positive linear relationship (as one variable increases/decreases, the other increases/decreases respectively) and 0 indicating no linear relation between two variables (Huysamen, 1987). As statistical significance of this value is largely influenced by sample size, effect sizes
should be interpreted based on the magnitude of the correlations (Hubbard, 2004; Terre Blanche, et al., 2006). In Table 10, the Spearman's Rho correlation coefficient was used as a nonparametric measure of statistical relationships between the variables (Morgan et al., 2002). The results clearly show that there are no statistically significant relationships between the variables – from both the Pearson and Spearman tests.

Table 10

**Correlation Coefficients and P-values**

<table>
<thead>
<tr>
<th></th>
<th>TRAM Score</th>
<th>TRAM Aggregate</th>
<th>VTS Score</th>
<th>VTS Aggregate</th>
</tr>
</thead>
<tbody>
<tr>
<td>June</td>
<td>-0.134</td>
<td>-0.128</td>
<td>-0.121</td>
<td>-0.123</td>
</tr>
<tr>
<td></td>
<td>(P=0.19)</td>
<td>(P=0.21)</td>
<td>(P=0.24)</td>
<td>(P=0.23)</td>
</tr>
<tr>
<td>July</td>
<td>-0.061</td>
<td>-0.006</td>
<td>-0.099</td>
<td>-0.106</td>
</tr>
<tr>
<td></td>
<td>(P=0.58)</td>
<td>(P=0.95)</td>
<td>(P=0.34)</td>
<td>(P=0.30)</td>
</tr>
<tr>
<td>August</td>
<td>-0.093</td>
<td>0.163</td>
<td>-0.143</td>
<td>-0.167</td>
</tr>
<tr>
<td></td>
<td>(P=0.37)</td>
<td>(P=0.11)</td>
<td>(P=0.16)</td>
<td>(P=0.10)</td>
</tr>
<tr>
<td>September</td>
<td>0.043</td>
<td>-0.031</td>
<td>-0.015</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>(P=0.68)</td>
<td>(P=0.77)</td>
<td>(P=0.69)</td>
<td>(P=0.88)</td>
</tr>
<tr>
<td>October</td>
<td>-0.174</td>
<td>0.137</td>
<td>-0.216</td>
<td>-0.212</td>
</tr>
<tr>
<td></td>
<td>(P=0.09)</td>
<td>(P=0.18)</td>
<td>(P=0.03)</td>
<td>(P=0.03)</td>
</tr>
<tr>
<td>November</td>
<td>-0.027</td>
<td>0.012</td>
<td>-0.116</td>
<td>-0.126</td>
</tr>
<tr>
<td></td>
<td>(P=0.50)</td>
<td>(P=0.91)</td>
<td>(P=0.26)</td>
<td>(P=0.22)</td>
</tr>
<tr>
<td>December</td>
<td>0.044</td>
<td>0.063</td>
<td>-0.145</td>
<td>-0.139</td>
</tr>
<tr>
<td></td>
<td>(P=0.67)</td>
<td>(P=0.44)</td>
<td>(P=0.18)</td>
<td>(P=0.17)</td>
</tr>
<tr>
<td>January</td>
<td>0.051</td>
<td>0.199</td>
<td>0.052</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td>(P=0.62)</td>
<td>(P=0.77)</td>
<td>(P=0.69)</td>
<td>(P=0.66)</td>
</tr>
<tr>
<td>February</td>
<td>0.191</td>
<td>-0.119</td>
<td>0.107</td>
<td>0.097</td>
</tr>
<tr>
<td></td>
<td>(P=0.09)</td>
<td>(P=0.23)</td>
<td>(P=0.36)</td>
<td>(P=0.35)</td>
</tr>
<tr>
<td>Total Score</td>
<td>-0.048</td>
<td>0.075</td>
<td>-0.161</td>
<td>-0.163</td>
</tr>
<tr>
<td></td>
<td>(P=0.84)</td>
<td>(P=0.47)</td>
<td>(P=0.11)</td>
<td>(P=0.11)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>TRAM Score</th>
<th>TRAM Aggregate</th>
<th>VTS Score</th>
<th>VTS Aggregate</th>
</tr>
</thead>
<tbody>
<tr>
<td>June</td>
<td>-0.125</td>
<td>-0.127</td>
<td>-0.076</td>
<td>-0.075</td>
</tr>
<tr>
<td></td>
<td>(P=0.22)</td>
<td>(P=0.22)</td>
<td>(P=0.46)</td>
<td>(P=0.47)</td>
</tr>
<tr>
<td>July</td>
<td>-0.030</td>
<td>-0.051</td>
<td>-0.107</td>
<td>-0.107</td>
</tr>
<tr>
<td></td>
<td>(P=0.77)</td>
<td>(P=0.95)</td>
<td>(P=0.30)</td>
<td>(P=0.39)</td>
</tr>
<tr>
<td>August</td>
<td>-0.060</td>
<td>0.123</td>
<td>-0.160</td>
<td>-0.160</td>
</tr>
<tr>
<td></td>
<td>(P=0.86)</td>
<td>(P=0.95)</td>
<td>(P=0.12)</td>
<td>(P=0.15)</td>
</tr>
<tr>
<td>September</td>
<td>0.052</td>
<td>-0.020</td>
<td>-0.006</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td>(P=0.32)</td>
<td>(P=0.36)</td>
<td>(P=0.95)</td>
<td>(P=0.86)</td>
</tr>
<tr>
<td>October</td>
<td>-0.171</td>
<td>0.143</td>
<td>-0.226</td>
<td>-0.225</td>
</tr>
<tr>
<td></td>
<td>(P=0.09)</td>
<td>(P=0.18)</td>
<td>(P=0.03)</td>
<td>(P=0.03)</td>
</tr>
<tr>
<td>November</td>
<td>0.002</td>
<td>0.038</td>
<td>-0.086</td>
<td>-0.097</td>
</tr>
<tr>
<td></td>
<td>(P=0.98)</td>
<td>(P=0.71)</td>
<td>(P=0.41)</td>
<td>(P=0.35)</td>
</tr>
<tr>
<td>December</td>
<td>0.073</td>
<td>0.058</td>
<td>-0.116</td>
<td>-0.130</td>
</tr>
<tr>
<td></td>
<td>(P=0.48)</td>
<td>(P=0.58)</td>
<td>(P=0.28)</td>
<td>(P=0.24)</td>
</tr>
<tr>
<td>January</td>
<td>0.064</td>
<td>0.174</td>
<td>0.071</td>
<td>0.061</td>
</tr>
<tr>
<td></td>
<td>(P=0.54)</td>
<td>(P=0.59)</td>
<td>(P=0.49)</td>
<td>(P=0.56)</td>
</tr>
<tr>
<td>February</td>
<td>0.207</td>
<td>-0.172</td>
<td>0.106</td>
<td>0.091</td>
</tr>
<tr>
<td></td>
<td>(P=0.94)</td>
<td>(P=0.99)</td>
<td>(P=0.33)</td>
<td>(P=0.36)</td>
</tr>
<tr>
<td>Total Score</td>
<td>0.001</td>
<td>0.032</td>
<td>-0.136</td>
<td>-0.143</td>
</tr>
<tr>
<td></td>
<td>(P=0.96)</td>
<td>(P=0.76)</td>
<td>(P=0.18)</td>
<td>(P=0.16)</td>
</tr>
</tbody>
</table>
In Table 10, a correlation coefficient was determined between the performance management scores (for each month along with the total score), the TRAM assessment scores and the total (aggregated) TRAM score. Similarly, correlation coefficients were calculated for the VTS score, total (aggregated) VTS score and for a combination of the two tests. No statistically significant relationship was identified. The researcher then looked at the effect of grouping the sample group into different categories.

No statistically significant correlations were found and as a result regression analyses were not performed.

**Supervisor Rankings**

As an additional independent performance rating, the supervisor of the drivers and operators ranked their performance as an alternative to using performance management data exclusively. The rankings were done by the plant-operations manager. This person is solely responsible for the recruitment, well-being and dispersion of all the company’s drivers and operators. Because the company utilises a vast range of different vehicles and equipment it was difficult to group the drivers and operators per machine or vehicle category. This was especially the case because many drivers and operators are qualified to operate multiple machines. Through an elimination process the plant operations manager (supervisor) selected the best operator or driver for a specific machine or vehicle that the participant is mostly responsible for. Some groups only had 3 members which made it statistically impossible to accurately measure the supervisor’s score against that of the company’s performance management scores. Instead, the supervisor elected the best and worst driver or operator in a machine category through an elimination process. He selected the best and worst drivers and operators based on his interpretation of their performance. Drivers and operators were then grouped, based on this selection, into best-, worst- and other groups.
A point of departure was to first rank the drivers and operators into groups (best, worst and other) based on the mean scores obtained from the assessment results. For this purpose, the sample was divided into three distinct groups: participants who scored from 0 – 60 were categorised as worst; 61 – 74 as other; and 75 – 100 as best. These results are summarised in Table 11 with the number of participants falling into each group shown.

Table 11  
*Sample Group Rankings According to Mean Scores (n=95)*

<table>
<thead>
<tr>
<th>Ranking in terms of test mean scores</th>
<th>Best</th>
<th>Worst</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRAM</td>
<td>6</td>
<td>69</td>
<td>20</td>
</tr>
<tr>
<td>VTS</td>
<td>0</td>
<td>93</td>
<td>2</td>
</tr>
<tr>
<td>TRAM &amp; VTS</td>
<td>0</td>
<td>86</td>
<td>9</td>
</tr>
<tr>
<td>Ranking in terms of Performance Score</td>
<td>14</td>
<td>12</td>
<td>69</td>
</tr>
<tr>
<td>Ranking in terms of Supervisor Ranking</td>
<td>23</td>
<td>31</td>
<td>41</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>9.89</td>
<td>35.27</td>
<td>27.16</td>
</tr>
</tbody>
</table>

Note. Best=75 – 100. Worst=0 - 60. Other= 61 - 74

The supervisor’s rankings were then compared to the mean scores of the TRAM, VTS and performance scores to investigate any correlation between the supervisor’s groups and the groups based on the assessment mean scores. No statistically significant relationship was found between the supervisor’s rankings and the performance scores. The standard deviation is shown to show the large variation between the different tests’ rankings.

The researcher was also interested to see the quantitative difference between the supervisor’s rankings and the rankings obtained from the various scores. Movements are labelled as regress, improve or no change. For example, if the supervisor had...
ranked a candidate as best and according to his test score was ranked as worst, the movement would be a decrease from best to worst. Table 12 shows the number total movements in each case.

Table 12

*Ranking Movement According to Mean Scores (n=95)*

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>TRAM</th>
<th>VTS</th>
<th>TRAM &amp; VTS</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decreased - Best to Other</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Decreased - Other to Worst</td>
<td>50</td>
<td>67</td>
<td>63</td>
<td>23</td>
</tr>
<tr>
<td>Decreased - Best to Worst</td>
<td>11</td>
<td>14</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>Improved - Worst to Other</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Improved - Other to Best</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>Improved - Worst to Best</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>No Change</td>
<td>22</td>
<td>14</td>
<td>15</td>
<td>46</td>
</tr>
</tbody>
</table>

| Correlation Coefficient            | -0.02846 | 0.089805 | -0.058531487 | 0.09258245 |

At the on-set of the research statistically significant relationships between predictor and criterion data were expected. The conclusion of the research study, however, did not yield any of the expected results.

Hypothesis 1 stated that psychomotor ability and learning potential are statistically significantly and positively related to work performance. Hypothesis 2 stated that psychomotor ability statistically significantly predicts work performance. Hypothesis 3 stated that learning potential statistically significantly predicts work performance. Hypothesis 4 predicted that psychomotor ability and learning potential jointly statistically significantly predict work performance. Based on the results above it is evident that none of the hypotheses can be accepted and that the null hypotheses cannot be rejected.
DISCUSSION

The main objective of the study was to investigate the theoretical and statistical relationship between psychomotor ability and learning potential in predicting job performance, with specific reference to drivers and operators.

The study is of importance because it contributes to a larger body of knowledge on the effective selection of high performing individuals. The study also contributes toward emphasising the importance of developing accurate performance management tools and recording accurate performance measurement scores.

The results of this study were however unexpected. No statistically significant relationship was found between the variables which is surprising considering recent research literature. At least four recent studies did report statistically significant relationships either on the same or very similar tests (Aguilera-Vanderheyden, 2013; Gilmore, 2008; Keyser, 2012; Pelser, 2002). Aguilera-Vanderheyden (2013) found that specific psychomotor assessment metrics were identified as being predictive of injuries in the South African mining environment. Gilmore (2008) found that there is a statistically significant relationship between learning potential and job performance. Results from research conducted by Keyser (2012) indicated that a learning potential test battery comprising LPCAT and the VTS is valid for predicting which candidates would be the higher scorers of safety behaviour and who would be lower scorers of safety behaviour in a mining company. Pelser (2002) found that a learning potential test and psychomotor ability tests were valid predictors of the job performance of haul truck operators.

From this study, however, the results showed no statistically significant relationship between psychomotor ability, learning potential and job performance. Not one of the formulated hypotheses could therefore be supported from the results. The hypotheses were:
H1: Psychomotor ability and learning potential are statistically significantly and positively related to work performance.

H2: Psychomotor ability statistically significantly predicts work performance

H3: Learning potential statistically significantly predicts work performance

H4: Psychomotor ability and learning potential jointly statistically significantly predict work performance

The main reason why none of these hypotheses could be supported is because the research results showed that there are no statistically significant relationships between the variables.

Limitations

It came to light during later interviews with participants after the study was completed that the appraisers were only remotely acquainted with the individual driver and operator being assessed. This was mainly the case due to drivers and operators being rotated to different construction sites with different construction managers scoring their performance. Furthermore, the drivers and operators were oblivious to how they are being assessed, by whom they are being assessed and what the monthly assessment results were. In contrast to verifiable and substantiated predictor data, the criterion data cannot therefore be considered as being reliable or valid.

The small sample of drivers and operators should also be noted as an additional limitation to the study. According to Cohen (1992), statistical power refers to the probability of detecting a relationship between predictor and criterion in a sample if such a relationship exists in the population. Power is probably more easily understood as the probability of rejecting a null hypothesis when it is, in fact, false. Borenstein, Rothstein and Cohen (2001) maintained that power together with the sample size, the significance level and the effect size of the research project form a closed system. In other words, once any three values are known, the fourth can be calculated. The aim of a power analysis is to find a balance between these factors,
taking cognisance of the resources available and the purpose of the research. The power of a study increases as sample size (n) increases, as alpha (significance) increases and as the magnitude of the effect in the population (effect size) increases (Cascio, 1998). According to Wilcox (2003), the higher the number of candidates in the sample pool (i.e. the bigger the sample), the higher the power of the study will be.

Cohen (1992) used 2 independent variables, testing for correlation and using an alpha of 0.05. Cohen (1992) found that, for a power of 0.80, a medium effect size would require a sample size of 67 and for a small effect size require a sample of 482 research participants. According to Cohen (1992) then, the sample size of this study fell between small and medium. This means that extreme outliers may influence the distribution.

**Recommendations**

In feeding back the results to the organisation, the researcher recommended that the performance appraisal system be redesigned. Further research with quality criterion data was proposed to effectively address the research question and to establish whether the company’s assessment instruments accurately predict the work performance of its drivers and operators. A new, redesigned performance appraisal system for these drivers and operators should take specific technical competencies into account. Furthermore, employees should be properly informed on the criteria they will be assessed on, mangers (appraisers) should be thoroughly trained in appraising performance and evaluations should be more task-specific to exclude project allowables and site profitability.

Whether an organization uses sophisticated software or a simple paper-based method for performance appraisals, there are elements that are common to almost all performance appraisal systems (Law, 2007; Brown, Hyatt & Benson, 2010; Prowse & Prowse, 2009; Sillup & Klimberg, 2010):
1. An individual’s performance or behaviors (which can also include traits) are rated, evaluated, or judged by someone else.
2. Evaluations are scheduled (annually or quarterly), not tied to the completion of projects.
3. All employees are evaluated using the same system.
4. The process is usually mandatory, and tied to a reward system (pay raise or promotion).
5. Information from evaluation is kept in the employee’s file.

Additionally, many organizations use a similar process for implementing a performance appraisal system (Aamodt, 2012; Aarabi et al., 2013; Brudan, 2010; Meister & Willyard, 2010; Sillup & Klimberg, 2010; Steward et al., 2010) to:

1. Determine the reason for evaluating employee performance; for example, for increasing salaries, providing training and feedback, and promoting or firing employees.
2. Consider environmental and cultural limitations; for example, cultural differences in receptivity to feedback, the amount of money available for pay increases, the receptivity of management to the process (if already overworked), etc.
3. Determine who will evaluate performance; for example, the immediate manager only, the manager and the employee only, the manager and peers, etc.
4. Select the best appraisal criteria and methods for the organization’s goals; for example, the dimensions to be rated, such as competencies (KSAs), tasks, goals/objectives, or traits; how the dimensions should be weighted (some more important than others or all equal); and whether to compare employees to each other, use objective measures, or have managers rate performance.
5. Train raters on how to evaluate performance using the criteria and methods chosen, and train employees on how the performance appraisal system works.
6. Advise managers on how performance should be documented; for example,
using an ongoing journal or record of performance throughout the year, updating objectives as they change, mid-year check-ins, etc.

7. Make recommendations for evaluating and documenting performance; for example, obtaining and reviewing objective data (days absent, customer feedback, etc.), reviewing critical incident logs, and avoiding common errors (leniency error, strictness error, halo error, etc.).

8. Determine how to communicate appraisal results to employees; for example, sending ratings to employees via email to review prior to meeting, meeting face-to-face to discuss, and separating appraisal feedback from discussions of pay increases or promotions.

9. Guide managers on making decisions for corrective action planning or terminating an employee; for example, considering “at will” doctrine, understanding legal/illegal reasons for terminating an employee, etc.

10. Monitor the legality and fairness of the appraisal system; for example, providing training to managers, avoiding discriminatory practices, etc.

There are also numerous ways that managers can improve the performance appraisal process:

- Performance improves most when specific, challenging goals are established (Appelbaum, et al., 2011; DeNisi & Pritchard, 2006; Schat & Frone, 2011)
- Including an employee in the goal-setting process helps increase motivation (Appelbaum, et al., 2011; Guralnik & Wardi 2003; O'Sullivan, 2009; Steward, 2010)
- Coaching should be a frequent, not a once-a-year, activity (Aamodt, 2012; Aarabi et al., 2013; Brudan, 2010).

Performance discussions should not be conducted with salary or promotion in the balance (Aamodt, 2012; Brown, et al., 2010; Prowse & Prowse, 2009; Steward, 2010). To improve supervisor ratings one may limit the rating task to those features of the job that supervisors can observe and quantify. Another important consideration is to ensure that supervisors spend time with employees, observing and discussing
their work. This will allow raters to familiarise themselves with the employee, his/her daily routine and how the tasks contribute toward overall objectives. If a specific task performance is more difficult to assess objectively via observation, supervisors should talk to employees to develop an understanding of the quality of their decision-making and reasoning. Supervisors could also do quality sampling by discussing specific task aspects and associated decisions.

It is critically important that human resources managers (or those implementing a performance appraisal process) train managers effectively in performance management, which is more than mere appraisal (Appelbaum et al. (2011); O’Sullivan, 2009; Sillup & Klimberg, 2010; Steward et al., 2010). Performance management is a comprehensive system that relies on a manager's role as a coach (DeNisi & Pritchard, 2006). Performance management involves frequent developmental meetings between a manager and employee, with a focus on developing employee strengths, providing learning and growth opportunities, and continually setting goals together and providing mutual feedback (Ahmed, Hussain, Ahmed, & Akbar, 2010; DeNisi & Pritchard, 2006). The employer and employee have an active role to play in effectively managing performance and to ultimately increase overall productivity, profitability and competitiveness (Meister & Willyard, 2010). The importance of successfully predicting performance is undisputed. However, quality and reliable criterion data is key to investigating the statistical significance of performance predictors.
REFERENCES


CHAPTER 4

CONCLUSIONS, LIMITATION AND RECOMMENDATIONS

4.1 INTRODUCTION

In this chapter the study is concluded, the findings and limitations are discussed and recommendations are made. The aim of the study was to determine whether learning potential and psychomotor performance can predict job performance of drivers and machine operators in a road construction company. Accordingly, the study determined the learning potential and psychomotor ability of drivers and operators. The respondents’ work performance was measured and the relationship between the two measures of the independent variables (psychomotor ability and learning potential) and the dependent variable (work performance) determined.

4.2 THE HYPOTHESES

In order to achieve the aim, the researcher formulated the following hypotheses:

H1: Psychomotor ability and learning potential are statistically significantly and positively related to work performance.
H2: Psychomotor ability statistically significantly predicts work performance
H3: Learning potential statistically significantly predicts work performance
H4: Psychomotor ability and learning potential jointly statistically significantly predict work performance

4.3 CONCLUSIONS

Due to various practical considerations, all the data in terms of predictor and criterion variables were not available for the whole population. Bearing these reduced frequencies in mind, sample size in certain of the research instances may limit the extent to which adequate statistical power can be achieved to provide a meaningful
test of the hypotheses (Cohen, 1992). This is especially the case in the current study because the research involves a relatively complex multivariate design in which criterion unreliability and range restriction may be present (Cascio, 1998). The fact that there is no significant correlation between the independent variables (psychomotor ability and learning potential) and the dependent variable (job performance) could indicate that the criterion was not as reliable as the researcher had hoped.

None of the hypotheses above could be supported by the findings of the study because no statistically significant relationship between the variables could be established. It is desirable for criterion measures to be highly reliable, as low reliability of the criterion measure places a ceiling on the validity coefficients that are attainable. This may lead to a type 2 error, namely in this case, missing a significant validity coefficient that was present (Ackerman et al., Aguinis, 2013; Cascio, 1998; Cascio, & Aguinis, 2010; De Kock, & Schlechter, 2009; Fernández-Ballesteros, 2003; Gardner & Deadrick, 2012; Gregory, 2007).

Bouwer’s (1984) research also concluded with no significant correlations for psychomotor measures with job performance for heavy duty truck drivers in a South African study. The fact that high correlations of Learning potential with performance were not found should probably not have been surprising, since the occupational position of drivers and operators could most likely be classified as a lower complexity job. The literature highlights that general (cognitive) ability or g is inclined to correlate better with job performance in more complex jobs (Jensen, 1986). Nonetheless, various studies have indicated that although the validity of cognitive ability varies across jobs it never approaches zero (Hunter & Hunter, 1984; Hunter, 1986; Schmidt & Hunter, 1981; Schmidt et al, 1988) – a contention that seems to be partially supported in terms of the Supervisor ranking criterion in the current study.
4.4 LIMITATIONS

A limitation to the study could be that no moderator variables have been taken into account. This is necessary in order to understand and control for the potential effects of anticipated moderator variables, such as age, years of education and years of operating experience, on the relationships between the predictors and the job performance criteria. Pelser (2002) found various significant correlations between the moderator variables and the respective predictors. This would indicate that the moderator variables did, in fact, impact on the performance of the various candidates on the predictors of this study. The moderators, could very well impact on the predictor-criterion relationship in this way.

A further important limitation was that no attempt was made to control for motivational aspects of driver and operator performance. This is a typical limitation which may have had a significant impact on the results. Driver and operator performance was assessed based on performance records captured on a routine basis by the company’s performance management system. The drivers and operators were hence not aware of their performance being assessed. It is thus conceivable that the predictor variables related to driving ability, whilst the criterion variables utilised in the study related to driving behaviour. Driving behaviour can be affected by a myriad of motivational factors that may have impacted on the correlations found in the study (Cascio, 1998). This aspect would also have had the effect of depressing the validity coefficients reported in this study thereby making the possibility of a type 2 error in this study a distinct possibility.

4.4.1 Limitations Pertaining to the Criteria

As in many validation studies, more specifically in the psychomotor field (Griffin & Koonce; 1996), the criterion problem (Cascio, 1998; Pelser, 2002) was a relevant factor in the current study. Due to the subjective nature of the performance management data criterion unreliability is suspected for the criterion. However, no attempt was made to determine the extent of such unreliability, bearing in mind that
there was insufficient data available to attempt to estimate its effect. Hence, the statistical correction of the validity coefficients could also not be attempted – a procedure referred to as correction for attenuation (Carretta & Ree, 2000). It is off course desirable for criterion measures to be highly reliable, low reliability of the criterion measure, places a ceiling on the validity coefficients that are attainable. This may also lead to a type 2 error, namely in this case, missing a significant validity coefficient that is, in fact, present (Carretta & Ree, 2000, Cascio, 1998).

A factor that has always been a problem in validation research in general, and more specifically in research in the psychomotor field, is the criterion and how it should be reliably measured (Pelser, 2002). Early validation work on various pilot selection batteries including psychomotor assessment was marred by the lack of an acceptable criterion for flight performance, due to inconsistent ratings of instructors, poor record-keeping of pilot performance and a low percentage of pilots who actually failed (Mc Farland, 1953) – the criterion problem at its best (Anastasi, 1988; Cascio, 199; Pelser, 2002). Griffin and Koonce (1996) emphasized that the criterion problem in pilot selection in the US military service, is still alive and well. Historically, a dichotomous pass/fail criterion was used in pilot selection studies. This became progressively more inappropriate as improved selection techniques made the selection pool progressively more homogenous (they are generally all good), leading to lower levels of attrition (i.e., fewer candidates fail). Due to the decreased variance in the ability of the candidates being assessed, correlations between predictors and criteria by definition cannot be very high (Griffin & Koonce, 1996; Pelser, 2002). It also limits the understanding of the relationships between predictors and criteria and limits statistical power (Duke & Ree, 1996, Hunter & Schmidt, 1990; Pelser, 2002).

This led to a move toward different criteria such as check ride or flight grades, and more objective criteria, such as the number of training flight hours required before a candidate is deemed competent to fly solo (Duke & Ree, 1996). Since these criteria are more or less normally distributed and continuous (rather than artificially dichotomous), higher uncorrected correlations have been reported in studies in this field (Carretta & Ree, 1994; Griffin & Koonce, 1996; Pelser, 2002)).
Since this study was a criterion validity study and the data collected belonged to the employees of the research organisation who were employed with the company at the time of the study, pre-selection of the candidates occurred earlier. This may have resulted in a certain amount of restriction of range within the scores. Restriction of range occurs when the variances on the variables are unrepresentatively small (Pelser, 2002). The sample was also one of convenience and therefore could not be generalised to the wider population of drivers and machine operators. The study was limited to the technical departments in the research organisation, which is only one of the many business units of the company. Historically, the recruitment procedures may have varied for the different levels and positions in the organisation and the various profiles of groups of employees may have differed. Because all the participants were employed in the organisation at that time and had therefore all successfully passed the various recruitment processes, there could again have been a restriction of range of scores. Since some information (e.g. education level) for all the incumbents was not available, this restriction could not be further investigated.

The two prominent limitations of the study are:

- The number of the units of analysis in the study (sample size) which is dictated by the type of research problem investigated. Because the sample size was so small, it was difficult to find significant relationships from the data, as statistical tests normally require a larger sample size to ensure a representative distribution of the population and to be considered representative of groups of people to whom results will be generalized or transferred.

- A serious lack of reliable criterion data proved to be a significant obstacle in finding a trend and a meaningful relationship between the variables. However, this limitation should be used as an opportunity to motivate the need for future research in this field.
5. RECOMMENDATIONS

The researcher recommends that an attempt be made to corroborate and extend on the findings of this study, by using a predictive design in cross-validation (Cascio, 1998). This would minimize the range restriction, which is typical of predictive designs and is hypothesized to be active in the current study, mostly due to the levels of experience that the candidates have, impacting on their scores on the predictors and hence confounding the relation between predictor and criterion.

In order to attempt to limit the effects of motivational levels on the objective performance criteria measures, it is recommended that all subjects be informed that there will be systematic monitoring of each individual’s operating performance over a specified period. Although this will not totally cancel out motivational effects, it may have the effect of limiting their impact.

If better predictive validity coefficients were to be found in these studies, it would be possible to better explore the hypotheses of this study in terms of the incremental validity of psychomotor ability beyond learning potential. The existence of the general (cognitive) ability or $g$ factor reported in this study as well as the existence of any other factors beyond $g$ (e.g. a possible psychomotor factor or psychomotor precision factor) can also be confirmed.

In order to minimize the effect of the criterion problem, the reliability and validity of criteria need attention. In terms of the objective criteria collected from the performance management data, an obvious necessity is the professional technical auditing of the performance management system before any cross-validation studies should be undertaken in order to ensure that the objective criteria are reliable.

In terms of the more subjective criteria obtained from the supervisor rankings, it is recommended that a behaviourally anchored rating scale (Cascio, 1998) be developed to ensure more detailed, quantifiable, reliable and valid supervisor assessments per individual driver and operator.
This study was of a limited scope. Consequently, the only criterion dealt with in the research was job performance. No attempt was made to link the predictors to either training or safety criteria, both of which hold promise for continuous research. The methodical collection of information regarding the training needs of drivers and operators, for instance, provides for a more rational approach to effective training and development. Such a data base would afford a comprehensive overview of competencies and deficiencies, both within and between individual drivers and operators. From this the content, level and focus of training could be customized to meet the reported training and development needs, thereby streamlining the needs analysis process to enhance efficiency and effectiveness.

Similarly, safety is a crucial consideration in the road construction industry, and arguably enjoys even more focus than productivity. The Occupational Health and Safety Act (1993) place the onus on the employer to ensure that safety risks are minimized. The consequences and potential cost of not adhering to the Safety Act (1993) are alarming and demands the attention of any respectable employer. It therefore makes sense that the road construction industry should be interested in identifying potential operators, who display the least risk from a safety perspective. It can further be argued that the Vienna Test System (VTS) subtests conceptually relate better to safety than to productivity measures.

- The Cognitrone (Schuhfried, 2000a) yielded data in terms of the candidate’s ability to concentrate and to adjust his or her work tempo to different stimuli patterns. Hence, all the Vienna Test System subtests have a strong conceptual link to safety.
- The Determination unit (Schuhfried, 1996) specifically focuses on the operator’s capacity to make appropriate and fast responses in rapidly changing environments that may involve stress. The test starts off slowly, gains speed to a very fast response requirement (approximating high stress situations e.g. accident or near-accident situations) and then slows down marginally (approximating the period just after the accident/near-accident).
• The Two-hand coordination subtest (Schuhfried, 2000c) is specifically focused on the candidate’s hand-eye and hand-hand coordination, which is conceptually related to safety in terms of small movements that need to be made during vehicle and machine operation.

• The Distance estimation time and motion measures (Schuhfried, 2000b), for example, attempt to identify those candidates who are least likely to underestimate distance and hence stop too late or cut in front of moving machinery when it is not safe to do so.

• The Signal detection subtest (Schuhfried, 2000d), the results of which were not used in this validation exercise, yields data on the candidate’s ability to sustain concentration levels in monotonous conditions.

Despite its importance, there are very few good validity studies relating to safety criteria. This may be due to the difficulty of obtaining safety criteria that are reliable (accidents are generally speaking, relatively infrequent events and near-misses are seldom reliably reported). In order to add to the literature in this crucial field, it is recommended that the safety variables pertinent to drivers and machine operator performance should be measured using a simulator. The predictor data from the current research can then be correlated with simulator performance focusing on safety variables, such as the number of times during the simulation exercise that the operator stopped too late; underestimated the speed of approaching vehicles; or displayed risk behaviour (e.g. driving too fast, overtaking on an incline etc.). The simulator can also be used to train and assess drivers and machine operators in terms of operating in conditions that cannot be practiced or assessed in real life, such as accident situations (break failures, tyre bursts, near-misses, slippery road conditions, operating in tight conditions) and so on. This would make a significant contribution to both the research literature and practical safety in the road construction industry.

Based on the findings the researcher makes the following recommendations for practice and further research.
5.1 Practice

In an effort to limit the effects of an average performance management score, it is recommended that all employees be informed that there will be systematic monitoring of each individual’s job performance behaviour over a twelve-month period, with quarterly intervals. This should provide a robust measure of the criterion variable of job performance.

It is recommended that a broader job performance profile should be obtained of each employee and not only a single performance score. This will also enhance the quality of the criterion measure.

5.2 Further Research

Further research should be conducted on the following topics:

- Additional performance management results that should be taken into consideration to obtain broader performance behaviour profiles of drivers and operators.
- Development of a reliable, objective performance management system to measure and obtain reliable performance scores which can then be compared to psychomotor ability and learning potential.
- Objective measurement of work performance – and specifically in the road construction industry.
- The reliability and validity of the dependent variable.

It would be of great benefit for researchers to establish different job performance models based on job level complexity to more accurately identify the core traits associated with the employees’ differing tasks. Also, due to the limited number of studies using psychomotor scores to predict driver and machine operator performance, research should be conducted to expand on and enhance the findings of the present study.
6. CHAPTER SUMMARY

This chapter briefly discussed the findings and limitations of the study and made recommendations for practice and further research.
REFERENCES


Studer, Q. (2008). *Results that last: hardwiring behaviours that will take your company to the top*. NJ: Wiley.


