

**THE EMPIRICAL VALIDITY OF AN ASSESSMENT BATTERY FOR  
APPRENTICE ELECTRICIAN STUDENTS**

M.A. van Stelten, Hons. B.A.

Mini-dissertation submitted in partial fulfillment of the requirements for the degree  
Magister Artium in Industrial Psychology at the North-West University, Potchefstroom  
Campus

**Supervisor:** Prof. G. v/d M. Sieberhagen

**May 2008**

**Potchefstroom**

**Dedicated to my husband, Bram, and my children Marco and Sulinka.  
Thank you for your unconditional love, patience and support. I am blessed  
to have you in my life.**

## ACKNOWLEDGEMENTS

This mini-dissertation has been made possible by the input and support of many people. Acknowledgements are due to the following persons:

- Prof. G. v/d M Sieberhagen, my supervisor, for his guidance, patience, time and effort to support me through this process.
- Tina Joubert, Dr. Liezel Korf and Dr. Marie de Beer for their availability, motivation, processing of the empirical results and well-informed feedback.
- Neels de Jager, André Mannel, Adryn Maritz and Nelis Faist of the Technical College in the North-Cape Province, for their professional help and contributions to this study.
- Idette Noome for the professional language editing of the mini-dissertation.

And especially to:

- My precious husband, Bram, for his loving support, consistent encouragement, patience and belief in me and my abilities.

## REMARKS

The reader is reminded of the following:

- The way of referencing, as well as the editorial style as prescribed by the Publication Manual (5<sup>th</sup> edition) of the American Psychological Association (APA) are followed in this mini-dissertation. This practice is in line with the policy of the Programme in Industrial Psychology of the School of Human Resource Sciences at the Potchefstroom Campus of the North-West University.
- The mini-dissertation is submitted in the form of three chapters, with the second chapter being a research article.

## ABSTRACT

**Title: The empirical validity of an assessment battery for apprentice electrician students**

**Key words:** Academic performance; empirical validity; predictor; ability tests; learning potential; Technical Test Battery; Learning Potential Computerised Adaptive Test; selection; psychometric tests; apprentice electricians.

Selection and selection procedures play a key role in the ability of organisations to compete successfully in the complex global and local environment. South Africa is experiencing a serious artisanal and technical skill shortage. Given the unemployment issues and the quest for people to fill the skill shortage gap it is important for organizations to find possible solutions to ensure that they stay competitive and effective in the labour market. Research suggests that at least 12 500 artisans need to be produced each year from 2006 to 2010 to meet the demand for skilled workers. To address above mentioned challenge organizations have to develop strategies to assist them to select individuals with the best chance to succeed in training. Unfortunately, the solution is not simple as organisations have a multitude of influences and pressures which affect their decisions regarding selection procedures. Selection of individuals is becoming an increasingly complex science as organisations have to select a capable and representative workforce. They must thus select candidates that are most likely to benefit from what is offered educationally, meet the requirements stipulated in South Africa's Labour legislation, and will perform most successfully in the specific trade. Bad practice can lead to costly litigation.

The objectives of this mini-dissertation were to assess the empirical validity of the Technical Test Battery (TTB), as ability test and the Learning Potential Computerised Adaptive Test (LPCAT) as learning potential test as predictors of academic success of first year apprentice electrician students at a South African technical college serving a mining community.

The study explored the current local employment issues that affect selection for training in the technical fields. The difference between the measurement of cognitive ability and learning potential were examined and the nature of the constructs of cognitive ability and that of learning potential were discussed.

In the empirical study one of the objectives was to determine whether there was a relationship between the TTB and the LPCAT as two different predictors of the academic success of first year apprentice electrician students. The difference in the empirical (predictive) validity of the two psychometric selection instruments, if only one of the tests as opposed to if both were used in combination, were examined. Another objective was to determine if there were any differences regarding the scores on the TTB and LPCAT of students from the designated group as opposed to that of students from non-designated groups. Finally this study explored whether the TTB and LPCAT were valid predictors to be used as selection instruments for apprentice electrician students in the South African context

The research method consisted of a literature review and an empirical study. The empirical validity of the two predictor tests was validated in terms of the accuracy with which the selection instruments predicted the students' future performance. This research can be categorized as *descriptive quantitative research*.

The TTB and LPCAT scores of a sample of 89 selected apprentice electricians were compared with the number of attempts they used to pass a phase test. Data was analysed by means of descriptive statistics. Pearsons Product Moment Correlation Coefficient, t-tests, ANOVAs as well as discriminant analysis were also used to reach the research objective. Statistically significant relationships were found between the predictor and criterion variables. The results confirm that the TTB and LPCAT are indeed empirical valid tests that can be used in the selection of apprentice electricians.

# OPSOMMING

**Titel:** Die empiriese geldigheid van 'n keuringsbattery vir vakleerling-studente in die elektriese ambag.

**Sleutelterme:** Akademiese prestasie; empiriese geldigheid; voorspeller; vermoëtoets; leerpotensiaal; Tegniese Toetsbattery (“Technical Test Battery”) (TTB), Gerekenariseerde Leerpotensiaal-aanpasbaarheidstoets (“Learning Potential Computerised Adaptive Test”) (LPCAT); keuring; psigometriese toets; vakleerlinge in die elektriese ambag

Keuring en keuringsprosedures speel 'n sleutelrol in die vermoë van organisasies om suksesvol te kompeteer in die uiters mededingende arbeidsmark. Suid-Afrika beleef tans 'n kritieke tekort in mense met ambags – en tegniese vaardighede. Werkloosheid en die aanvraag na bevoegde mense in sekere tegniese vaardighede forseer organisasies om oplossings te vind vir die aanvraag na mense wat oor hierdie vaardighede beskik. Volgens navorsing moet daar 12 500 ambagsmanne per jaar opgelei word vanaf 2006 tot 2010. Om aan hierdie behoefte te voorsien moet organisasies strategieë ontwikkel om hulle behulpsaam te wees om effektiewe keuringsbesluite ten opsigte van individue met die beste kans om suksesvol opgelei te word. Ongelukkig is die oplossing nie so eenvoudig nie. Organisasies het verskeie faktore wat 'n invloed uitoefen wanneer besluite rakende keuring van kandidate geneem word. Kandidate wat geselekteer word moet in staat wees om suksesvol te wees tydens hul opleiding, organisasies moet voldoen aan al die wetlike vereistes rakende seleksie van kandidate, en kandidate moet so goed moontlik presteer in hul spesifieke ambag. Indien bogenoemde nie korrek toegepas word nie kan organisasies 'n duur prys betaal.

Die doelwit van hierdie mini-verhandeling was om die empiriese geldigheid van die Tegniese Toetsbattery (TTB), as vermoëtoets sowel as die “Learning Potential Computerised Adaptive Test” (LPCAT), as leerpotensiaaltoets as voorspellers van die akademiese sukses van eerstejaar vakleerling studente in die elektriese amba by 'n Suid-Afrikaanse tegniese kollege in 'n myngemeenskap, te bepaal.

Die navorsing het die huidige indiensnemings vraagstukke wat in die tegniese veld voorkom ondersoek. Die verskil tussen kognitiewe vermoë en leerpotensiaal is ondersoek.

In die empiriese navorsing is die verhouding tussen die TTB en LPCAT as twee verskillende voorspellers van die akademiese sukses van eerstejaar vakleerling studente in die elektriese ambag ondersoek. Die verskil in die empiriese (voorspellings) geldigheid van die twee psigometriese keurings instrumente, indien slegs een of albei instrumente in kombinasie, gebruik word is bepaal. Nog 'n doelwit was om te bepaal of daar enige verskille ten opsigte van die toetstellings van die TTB en LPCAT voorkom tussen voorheen benadeelde studente teenoor studente wat wetlik beskou word as nie-benadeel nie. Die eiddoel van die studie het die toepaslikheid van die TTB en LPCAT as geldige voorspellers wat gebruik kan word tydens die seleksie van eerstejaarvakleerlinge in die elektriese ambag binne die Suid-Afrikaanse konteks te ondersoek.

Die navorsingsmetode het bestaan uit 'n literatuurstudie sowel as empiriese navorsing. Die empiriese geldigheid van die twee voorspeller toetse is bepaal deur middel van die akkuraatheid waardeur die keuringsinstrumente die studente se toekomstige prestasie voorspel het. Die navorsing kan as beskrywende kwantitatiewe navorsing gekategoriseer word.

Die TTB en LPCAT resultate van 89 geselekteerde studente in die elektriese ambag is vergelyk met die hoeveelheid pogings wat hulle gebruik het om 'n fase toets te slaag. Die data is geanaliseer deur die gebruik van beskrywende statistiek. Pearsons Product Moment Correlation Coefficient, t-toetse, ANOVAs en 'n diskriminant analise is ook gebruik om in die statistiese analise. 'n Statisties betekenisvolle verband tussen die voorspeller – en kriterium veranderlikes is gevind. Die resultate bevestig die aanname dat die TTB en LPCAT inderdaad empiries, geldige toetse is wat in die keuring van vakstudente in die elektriese ambag gebruik kan word.

## TABLE OF CONTENTS

### CHAPTER 1: INTRODUCTION ERROR! BOOKMARK NOT DEFINED.

1.	PROBLEM STATEMENT .....	Error! Bookmark not defined.
1.1	<i>South Africa's skills needs in a global environment</i> .....	Error! Bookmark not defined.
1.2	<i>Selection processes</i> .....	Error! Bookmark not defined.
1.3	<i>Selection procedures in South Africa</i> .....	Error! Bookmark not defined.
1.4	<i>Selection procedures at the Technical College in this study</i> .....	Error! Bookmark not defined.
1.5	<i>The test instruments examined in this study</i> .....	Error! Bookmark not defined.
1.6	<i>Problem statement</i> .....	Error! Bookmark not defined.
2.	RESEARCH OBJECTIVES .....	Error! Bookmark not defined.
2.1	<i>General objective</i> .....	Error! Bookmark not defined.
2.2	<i>Specific objectives</i> .....	Error! Bookmark not defined.
2.2.1	<b>Specific objectives of the literature review</b> .....	Error! Bookmark not defined.
2.2.2	<b>Specific objectives of the empirical study</b> .....	Error! Bookmark not defined.
3.	RESEARCH METHOD .....	Error! Bookmark not defined.
3.1	<i>Literature review</i> .....	Error! Bookmark not defined.
3.2	<i>Empirical study</i> .....	Error! Bookmark not defined.
3.2.1	<b>Research design</b> .....	Error! Bookmark not defined.
3.2.2	<b>Study population</b> .....	Error! Bookmark not defined.
3.2.3	<b>Measuring battery</b> .....	Error! Bookmark not defined.
3.2.3.1	<i>The TTB</i> .....	Error! Bookmark not defined.
3.2.3.2	<i>The LPCAT</i> .....	Error! Bookmark not defined.
3.2.4	<b>Measurement of the criterion (academic success)</b> .....	Error! Bookmark not defined.
3.2.5	<b>Statistical analysis</b> .....	Error! Bookmark not defined.
4.	RESEARCH PROCEDURE.....	Error! Bookmark not defined.
5.	DIVISION OF CHAPTERS .....	Error! Bookmark not defined.
6.	CHAPTER SUMMARY.....	Error! Bookmark not defined.
7.	LIST OF REFERENCES.....	Error! Bookmark not defined.

### CHAPTER 2: RESEARCH ARTICLE.....3Error! Bookmark not defined.

### CHAPTER 3: CONCLUSIONS, RECOMMENDATIONS AND LIMITATIONS..... ERROR! BOOKMARK NOT DEFINED.

3.1	CONCLUSIONS .....	Error! Bookmark not defined.
3.2	LIMITATIONS.....	Error! Bookmark not defined.
3.3	RECOMMENDATIONS.....	Error! Bookmark not defined.
3.4	LIST OF REFERENCES.....	Error! Bookmark not defined.

# APPENDICES

## LIST OF TABLES

	<b>Page</b>
Table 1: Compilation of the study population	46
Table 2: Differences between the mean scores of the selected candidates And the non-selected candidates on the subtests of the TTB and the LPCAT	57
Table 3: Differences between the scores of the selected candidates and the non-selected candidates on the predictors	58
Table 4: Descriptive statistics in respect of the predictor variables	59
Table 5: Matrix of correlations between the scores on the predictor variables	60
Table 6: Mean number of attempts needed to pass the different phase tests (criterion variables)	61
Table 7: Matrix of correlations between scores on the predictors (the TTB and the LPCAT) and the criterion variables (mean number of attempts needed to pass the different phase tests)	63
Table 8: Matrix of correlations between scores on the predictors (the TTB and the LPCAT) and the criterion variables (mean number of attempts needed to pass the different phase tests)(after adjustments for restriction of range)	64
Table 9: Comparison of the mean test scores on the predictors (the TTB and the LPCAT) of the candidates who needed one, two or three attempt (s) to pass the different phase tests	66
Table 10: Differences between the means of the designated and non-designated groups regarding their scores on the predictors (independent sample t-tests)	72
Table 11: Difference between the designated and non-designated groups regarding their scores on the predictors	73

## LIST OF TABLES (continue)

	<b>Page</b>
Table 12: Significance of the differences between non-designated and designated groups regarding their scores on the criterion	74
Table 13: Differences between the non-designated and the designated groups regarding their scores on the dependent criterion variant (number of attempts needed to pass the different phase tests)	75
Table 14: Level of education of the selected candidates and their performance on the predictor variables	76
Table 15: Classification results of the discriminant analysis for Phase Test 1	78
Table 16: Classification results of the discriminant analysis for Phase Test 2	79
Table 17: Classification results of the discriminant analysis for Phase Test 3	79
Table 18: Classification results of the discriminant analysis for Phase Test 4	80
Table 19: Classification results of the discriminant analysis for Phase Test 5	81
Table 20: Classification results of the discriminant analysis for Phase Test 6	81
Table 21: Classification results of the discriminant analysis for Phase Test 7	82
Table 22: Classification results of the discriminant analysis for Phase Test 8	83

## LIST OF FIGURES

	<b>Page</b>
Figure 1: The selection process	54
Figure 2: Number of attempts needed to successfully complete the phase tests	15 and 62
Figure 3: Parabolic regression curve regarding the frequency of attempts at passing the different phase tests	77

## **LIST OF APPENDIXES**

	<b>Page</b>
Appendix A: Reliability and validity of the TTB	103
Appendix B: Reliability and validity of the LPCAT	117
Appendix C: The contents of the different phase tests	121



# **CHAPTER 1**

The empirical validity of an assessment battery for apprentice electrician students.

## **INTRODUCTION**

The objective of this mini-dissertation is to evaluate the empirical validity of the Technical Test Battery (TTB) and the Learning Potential Computerised Adaptive Test (LPCAT) as instruments that can be used to select apprentices for training as electricians at a Technical College in the Northern Cape, Republic of South Africa. The College is currently sponsored by an iron ore mine in the same region. (To maintain confidentiality, the College and the mine are not named in this study.)

Chapter 1 focuses on the problem statement, the aim of the study and the research method used in the study. The research procedure is explained and an outline of the subsequent chapters is given.

### **1. PROBLEM STATEMENT**

#### ***1.1 South Africa's skills needs in a global environment***

The World Competitiveness Report (World Economic Forum, 2001/2002) indicates that at present South Africa is experiencing a serious employment crisis. In the last decade or so, local organisations have experienced some of the greatest changes in their history (Brummer, Badenhorst, & Neuland, 2005). Commodity price changes and exchange rate fluctuations, global opportunities, global competition, mergers, takeovers and restructuring are only a few of the critical issues that organisations currently have to face on an ever-increasing scale (Brummer et al., 2005). Aside from such complex and competitive macro-environmental influences, organisations must also meet the challenges posed by local factors such as the country's new Constitution (South Africa, 1996), a changed and changing political dispensation, the promulgation and implementation of a series of new Acts of Parliament to regulate matters pertaining to labour (Lopes, Roodt, & Mauer, 2001) and the severe skills shortages experienced throughout the country (Mlambo-Ngcuka, 2007). There are thus a multitude of influences and pressures which affect organisations' decisions regarding their management of global and local demands and forces. To create a sustainable competitive advantage in this complex environment, organisations must analyse

financial factors, their competitors and, perhaps most of all, their employees. Such factors can have both a short- and a long-term impact.

In South Africa there is a serious shortage of skilled, qualified and experienced people to fill particular roles and take up particular professions, occupations or specialisations in the labour market (Mlambo-Ngcuka, 2007). A significant proportion of the South African population lacks the basic competencies or skills needed to meet either the new or the existing challenges in organisations (Davies & Farquharson, 2004). In 2006, five high-profile priority skills areas were identified for immediate attention; amongst them are the skills needed by artisans and technical skills (Mlambo-Ngcuka, 2007). The economy currently produces only approximately 5000 artisans per year, but research suggests that at least 12 500 artisans must be produced each year over the next four years to meet the demand (Mlambo-Ngcuka, 2007). This is a projected increase of 30 000 artisans between 2007 and 2010. A number of priority trades and the requirements for each have also been identified. The work done by electricians is one of the identified trades (Mlambo-Ngcuka, 2007).

An analysis of workplace skills plans (WSPs) for the period 1 April 2006 to 31 March 2007 has revealed that employers in the Mining and Mineral Sector could not fill 12 422 positions, the equivalent of 2.6% of the total employment for that sector (Anon, 2005/2006). Most of the hard-to-fill vacancies were in the occupational categories for technicians and trades workers, where there were 5 159 vacant positions (Anon, 2005/2006).

The Skills Development Act of 1998 (South Africa, 1998b) and the Skills Development Levies Act 9 of 1999 (South Africa, 1999) are already designed to address such shortages (Davies & Farquharson, 2004). Since these Acts have been implemented, urgent attention is being given to the alignment and articulation of the various training pathways that are available. Learnerships are one of the key components of the Act (Anon, 2005/2006; South Africa, 1998b). A learnership is defined by De Jaeger, Hattingh and Huster (2002, p. 21) as “a route to a nationally recognised qualification that relates to an occupation and consists of a structured learning component and practical work experience”. The Mining and Mineral Sector has recommended promoting learnerships and apprenticeships in Further Education

and training colleges by expanding incentives to enrol more learners and developing more skills programmes (Mlambo-Ngcuka, 2007). Employers benefit from getting employees who have been well trained from both a theoretical and practical point of view, and they can claim grants each time a learnership agreement is entered into, which helps them cover some of the training costs (Anon, 2005/2006). Although the government incentives offered to promote employer involvement seem enticing to many organisations, many grants are only awarded once the candidate completes the learnership. If an apprentice leaves before completing the learnership, the organisation must re-hire and re-train candidates and must cover expenses on these incentives (Kearney, 2005). This suggests that organisations need to select candidates very carefully.

Every Sector Education and Training Authority (SETA) must have a skills plan. Such a sector skills plan must describe the labour market conditions and skills needs for the sector concerned and the supply of skills to that sector (South Africa, 2007). To meet their skills shortage demands, organisations must select employees that can perform successfully. Hence effective selection procedures are of cardinal importance to any organisation. A dynamic selection process can assist an organisation in measuring and predicting the behaviour of the individuals that it chooses to meet its needs (Van der Merwe, 2005) and in helping these individuals to meet the challenges of the environment they are part of. It can also reduce the expense to an organisation arising when candidates leave before they have completed their learnerships.

## ***1.2 Selection processes***

Selecting employees is a complex decision-making process that requires a comprehensive evaluation of each applicant's performance (López, 2004). To make an objective decision, multiple selection criteria must be considered simultaneously (López, 2004). Every selection technique has its own merits. In this regard, research has confirmed that using standardised tests is a psychometrically sound assessment method (Baron & Butterworth, 1998; Friedenberg, 1995).

Psychometric testing is the most common procedure to predict human behaviour used in selection (Van der Merwe, 2005). Standardised tests are commonly employed as a selection aid. Research has shown that well-constructed psychometric tests

- predict job performance better than any other single selection method;
- provide objective information about a candidate;
- avoid the problems of subjective bias associated with other techniques such as interviews or references; and
- have been shown in general to lead to better and fairer employment decisions (SHL, 2006).

Psychometric tests have been used for over a century. According to Taylor (1994), in 1882, Galton established a laboratory in London where people could be tested on visual acuity and reaction-time tasks. In 1905, a French psychologist, Binet, and his colleague Simon introduced the first intelligence tests (De Beer, 2005). Outtz (2002) indicates that tests designed to measure general mental ability have been used in employment selection for more than 80 years.

The quest for fair and valid assessment tools has long been pursued at advanced levels in societies such as those in countries like Canada, Australia, Britain and the United States of America (Cummins, 1984; Samuda, Kong, Cummins, Pascual-Leone, & Lewis, 1989). Both international and local research has demonstrated that psychometric assessment can play an important role in significantly improving the selection process for new entrants (Anastasi & Urbina, 1997; Friedenberg, 1995; Grussendorff, Liebenberg, & Houston, 2004; Lopes et al., 2001; Nunns & Ortlepp, 1994). The standardisation of tests has also been widely researched. Standardisation refers to the uniformity of the test material, the administration of and the scoring procedure for the instruments being used (Huysamen, 1998).

### ***1.3 Selection procedures in South Africa***

Several psychological tests have already been standardised in South Africa (Owen & Taljaard, 1996), but many were originally developed with white test takers in mind (Huysamen, 2002). Selection testing in South Africa has ranged from testing for prior learning (that is, for core competencies and acquired knowledge) to testing for learning potential (Grussendorff et al., 2004).

In the South African context, selection issues are unique and complex. The country has 11 official languages; many social and educational problems; and there are huge disparities in the socio-economic and educational backgrounds of individuals (De Beer, 2005). The labour situation has changed dramatically since 1994 (Lópes et al., 2001), and fair selection practices are critical (Huysamen, 2002; Snelgar & Potgieter, 2003). Any discriminatory practices left over from the apartheid system have to be eradicated, and unsuitable instruments have to be replaced by measures designed to redress past injustices. Recent political and legislative changes require selection to be fair and transparent (Zaaiman, 1998). In order to address the issue of group representivity in employment outcomes and to ensure equal representation in all occupational categories and levels in the workplace, Section 15 of the Employment Equity Act 55 of 1998 (South Africa, 1998a) includes affirmative action measures. Institutions therefore face a multiple challenge when they make selection decisions. They have to select the best applicants available, resolve the skills shortage problem and comply with legal requirements without compromising standards or fairness (Nunns & Ortlepp, 1994). The challenge then is to adhere to all the guidelines for fair selection and still to find well-qualified, competent and experienced employees in the identified areas of scarce and critical skills.

Against the above background, it is fair to state that there is a vast need to establish the extent to which selection procedures can predict the future performance and potential of applicants in the identified scarce occupations in South Africa (De Jong & Visser, 2000). To address this need for responsible and accurate selection decisions and decision-making tools, researchers must determine which tests will give the best results in particular circumstances. Results must also be interpreted responsibly (Van der Merwe, 2002). There are many matters that need to be considered in order for this balance to be achieved. This highlights the need for the responsible use of psychometric tests and other selection procedures.

Section 8 of the Employment Equity Act (South Africa, 1998a) emphasises the need for the reliability, validity and fairness of assessment procedures and states that such procedures may not be biased. It states that “Psychological testing and other similar assessments of an employee are prohibited unless the test or assessment being used:

- has been scientifically proven to be valid and reliable;

- can be applied fairly to all employees; and
- is not biased against any employee or group.”

Users of psychological tests and similar assessment instruments throughout South Africa have become increasingly concerned about the legitimacy of the assessment procedures they use. Organisations currently have no option but to validate the tests and other assessment instruments that they use. This is the only way to prove that these instruments comply with the requirements set by the Employment Equity Act (Pelser, Bergh, & Visser, 2005). According to Lopes et al. (2001) and Boolsen and Theron (1996), an employer should be able to prove scientifically that

- the information gathered during the selection process is valid and job-relevant;
- the use of the selection technique makes it a business necessity;
- the information is combined in a fair manner for decision-making; and
- the selection efficiency justifies the use of the selection technique.

Pelser et al. (2005) have identified two key reasons for investing time and effort in researching selection procedures in South Africa, namely:

- to remain within the ambit of the law (this refers to industrial relations, ethical and cultural issues, and the social utility of fairness considerations); and
- to maximize the probability of selecting the potentially most productive candidate(s) (this relates to issues such as shortages of skills, the impact of globalisation as experienced in the South African economy and the need for increased competitiveness).

Local and international research has found that people from disadvantaged backgrounds tend to score poorly when they are assessed by means of standard psychometric instruments (De Beer, 2005; Ployhart, Ziegert, & McFarland, 2003; Ryan, 2001; Shochet, 1994). These poor scores often reflect a lack of educational opportunities and not necessarily a lack of potential (De Beer, 2005; Visser & Hanslo, 2005). Tests of cognitive abilities generally measure an individual’s current level of functioning and they rely heavily on the assumption that all examinees have had comparable opportunities to acquire the skills and abilities being measured (De Beer, 2005; Murphy & Maree, 2006; Shochet, 1994). However, cognitive ability tests

produce racial differences that are three to five times larger than those noted for other predictors (Hunter & Hunter, 1984; Outtz, 2002).

Given the socio-economic and educational disadvantages that many citizens experienced in the past and the requirements of the Skills Development Act (South Africa, 1998b), local researchers have indicated that there is a clear need for assessment strategies to focus on future potential, rather than on current ability (Andrews, 1996; De Beer, 2000; De Villiers, 1999). Some researchers suggest that cognitive ability tests should be combined with other predictors in order to reduce any adverse impact, while increasing overall validity (Outtz, 2002; Shochet, 1994; Taylor, 1994).

It is important to note the difference between standard selection processes and those performed in the context of a learnership (Davies & Farquharson, 2004). Standard approaches tend to focus on the “skills held” by a candidate and generally seek to select the “best candidate” for a particular position, based on existing qualifications and a set of criteria. In the case of learnerships, the process generally focuses on a developmental context. This means that the focus is on candidates with the greatest potential to develop their skills in order to meet the objectives of the learnership (Davies & Farquharson, 2004).

According to Kearney (2005), only 60% of candidates complete their learnerships. Studies by DesJardins (2002) and Kearney (2005) indicate that the retention, dropout and throughput of students are vital to academic planning and central to issues of the implied costs that are imposed on society, the educational institution and the individual. DesJardins (2002) argues that the dropout of “disadvantaged students” may result in a social cost because it may lead to further racial and socio-economic disparities in future generations. For the institution, there is a cost because imbalances between the intake and the completion of training have considerable budgetary and financial implications. Employers will not find suitably qualified employees. For an individual, the cost of leaving the educational institution without successfully completing his or her training is a potential loss in earnings and being sidelined when he or she applies for a job; not to mention the personal and emotional issues related to dropping out for academic reasons (Visser & Hanslo, 2005). However, Kearney

(2005) maintains that the problem can be addressed through rigorous selection processes, as he argues that accurate selection procedures can predict the future performance of candidates. Doing so may help the government and training facilities to reach the annual target of 12 500 trained artisans mentioned at the start of the study, and may save on training costs by choosing only those candidates that have a reasonable chance of succeeding in their training.

#### ***1.4 Selection procedures at the Technical College in this study***

The Technical College that was reviewed in this investigation has replaced a traditional apprenticeship system with learnerships which supply students with a work-based route to a nationally recognised qualification within the National Qualifications Framework (NQF). The College provides a structured learning programme and combines this programme with workplace experience to contribute to the supply of artisans who have the necessary skills to meet the increased demand. Based on the local mines' occupational needs, particular mines provide learnerships. The mine concerned thus carries all training costs and gives each student a monthly allowance. The mine does not have any payback agreement with the student if he or she does not complete the training. It is therefore of the utmost importance for the College to select those candidates who

- are most likely to benefit from what is offered educationally;
- meet the requirements stipulated in South Africa's Labour legislation; and
- will perform most successfully in the specific trade.

After some market research, as part of the pre-selection criteria, the Technical College set a matriculation certificate or the completion of an N2 qualification in the relevant trade as its minimum requirements for entering the selection process. In the selection process itself, candidates then have to complete the Technical Test Battery (TTB) (SHL, 1996), the Learning Potential Computerized Adaptive Test (LPCAT) (De Beer, 2000), a structured interview and a medical examination. This forms a part of the selection battery for artisan students. No job analysis was done. The same selection criteria are used for all five trades (electrician, fitter and turner, boilermaker, diesel mechanic and millwright) that are presented at the College.

The duration of the training for artisans is two years. The first year's training is done at the Training College. The second year consists of practical training at the mine. Students that are doing different courses complete different modules. The number of attempts required for a student to pass a phase test during his or her first year of training at the college is used as the criterion variable in this study.

### ***1.5 The test instruments examined in this study***

The two instruments used at the Technical College under review are the Technical Test Battery (TTB) and the Learning Potential Computerised Adaptive Test (LPCAT).

The Technical Test Battery (TTB) is an ability test battery. It consists of a series of tests designed to test the suitability of individuals for selection and development in technically or practically oriented jobs (SHL, 1996).

The Learning Potential Computerised Adaptive Test (LPCAT) was developed in South Africa as a non-verbal, culture-fair measure of learning potential in the domain of cognitive or general non-verbal reasoning ability (De Beer, 2000). It takes into account the fact that examinees come to the testing situation from diverse and often disadvantaged backgrounds (De Beer, 2000).

### ***1.6 Problem statement***

The purpose of this study is to do an empirical validity study of the TTB and LPCAT as predictors in the selection for apprentice electricians at a Technical College in the Northern Cape. The College was selected for this study because the Technical College expressed a specific need to determine the empirical validity of its selection battery, particularly the TTB and the LPCAT. Although the College offers courses in five fields, it was decided to focus on the selection of apprentice electricians only in order to limit the scope of the study to an in-depth study of one area.

It is important for the Technical College to know to what extent the two selected tests predict the academic performance of first year electrician students and to support specific interpretations with evidence and theory. The College wants the selection methods it follows to be fair in allowing opportunities for access, and effective in

predicting success in the learnerships they provide. The following research questions arise:

- What are the current local employment issues that affect selection for training in the technical fields?
- What is the nature of the constructs of cognitive ability and learning potential?
- What is the difference between the measurement of cognitive ability and the measurement of learning potential?
- Is there any relationship between the TTB and the LPCAT as two different predictors of the academic success of first-year apprentice electricians?
- Is there a difference between the empirical (predictive) validity of the psychometric selection method if only the TTB or only the LPCAT is used, as opposed to if both are used in combination?
- Are there any differences regarding the scores on the TTB and the LPCAT of students from designated groups as opposed to the scores of students from non-designated groups?
- Can the results of the TTB as an ability test and of the LPCAT as a learning potential test be used as predictors of the academic success of first year apprentice electricians?

## **2. RESEARCH OBJECTIVES**

### **2.1 *General objective***

Given the problem statement set out in Section 1.6 above, the general objective of this study is to do an empirical validity study of the TTB and the LPCAT as selection instruments for first year apprentice electricians at a Technical College in the Northern Cape in the Republic of South Africa.

### **2.2 *Specific objectives***

#### **2.2.1 *Specific objectives of the literature review***

The objectives of the literature review are to investigate

- the current local employment issues that affect selection for training in technical fields;

- the nature of the constructs of cognitive ability and learning potential; and
- the difference between the measurement of cognitive ability and the measurement of learning potential.

### **2.2.2 Specific objectives of the empirical study**

The objectives of the empirical study are to investigate

- whether there is a relationship between the TTB and the LPCAT as two different predictors of the academic success of first-year apprentice electricians;
- whether there is a difference in the empirical (predictive) validity of the psychometric selection method, if only one of the tests is used as opposed to if both are used in combination;
- whether there are any differences regarding the scores on the TTB and the LPCAT of students from designated groups as opposed to the scores of students from non-designated groups; and
- whether the TTB and the LPCAT are valid predictors to be used as selection instruments for predicting the success of apprentice electricians in South Africa.

## **3. RESEARCH METHOD**

The research method consists of a literature review and an empirical study. The article option is followed in this study.

### **3.1 Literature review**

The literature review explores current employment issues with a focus on the technical fields. Different selection challenges in the South African environment are discussed. Different selection methods are explored. Existing literature concerning two different selection instruments, namely the TTB and the LPCAT, are investigated. Differences regarding the scores on the TTB and the LPCAT of students from designated groups as opposed to those of students from non-designated groups are examined. The measurement and structure of ability and learning potential are

investigated. The relevance of the research findings concerning selection in the South African context are the main focus.

## **3.2 Empirical study**

### **3.2.1 Research design**

Kerlinger (1986, p. 279) defines a research design as “the plan and structure of investigation so conceived as to obtain answers to research questions”. The plan is the overall programme of the research, and it outlines what the researcher will do. The structure is a paradigm or model of the relations between the variables of a study. The purpose of a research design is thus

- to provide answers to research questions (Kerlinger, 1986);
- to control variance (Kerlinger, 1986); and
- to enhance the internal and external validity of the research (Mouton & Marais, 1992).

According to De Jong and Visser (2000), the empirical validity of research can be approached in the traditional way or via contemporary decision theory. If the traditional approach is used, a selection technique is validated in terms of the accuracy with which the technique predicts future performance (Cascio, 1997). Contemporary decision theory emphasises the utility of a certain selection decision; this implies that the measurement and prediction of performance is viewed as using technical data that has to be evaluated in terms of its utility for the organisation (Robertson, Iles, Gratton, & Sharpley, 1991). In this study, the traditional approach is applied.

This research can be categorised as descriptive quantitative research. Descriptive research identifies the characteristics of a phenomenon or explores possible correlations between two or more phenomena (Leedy & Ormrod, 2005). It examines a situation as it is. Descriptive research does not involve changing or modifying the situation under investigation, nor is it intended to determine cause-and-effect relationships (Leedy & Ormrod, 2005). Quantitative research is a form of conclusive research in which the data collection procedures are fairly structured (Struwig & Stead, 2001). A correlation study is used to reach the objectives of this study. Leedy and Ormrod (2005) explain that a correlation study determines the extent to which

differences in one characteristic or variable are related to differences in one or more other characteristics or variables. A correlation exists if, when one variable increases, another variable either increases or decreases in a somewhat predictable fashion.

The aim of this study is to determine how well two psychological tests (the TTB and the LPCAT) predict an individual's future performance. The predictor variables are the results of the scores on the TTB and the LPCAT. To define the criterion of success in this study, the success of an individual's performance is expressed as the number of attempts a student needs to pass a phase test.

### **3.2.2 Study population**

The sample consists of selected apprentice electricians registered at the Technical College from 2003 to 2006. To comply with the Employment Equity Act (South Africa, 1998a), 75% of the total number of selected students have to be members of the designated groups (Africans, coloureds, Asians, people with disabilities and women) and 25% may be members of the non-designated group (white males without disabilities). At least 10% have to be women.

### **3.2.3 Measuring battery**

A predictive criterion-related validity study is done to reach the research objectives. The criterion-related validity of a test may be defined as the extent to which the scores on the test predict the scores on the criterion concerned (Huysamen, 1998). In predictive designs, data on the selection procedure are collected at or at about the time when individuals are selected. After a specified period, criterion data are collected (Society for Industrial & Organisational Psychology of South Africa, 2005). According to Huysamen (1998), the predictive validity of a test refers to the accuracy with which a test predicts or forecasts some future behaviour or the future status of individuals.

The Technical College uses a standardised selection procedure. The selection procedure includes a combination of measuring instruments. For pre-selection purposes, an applicant must send his or her Curriculum Vitae to the Bursary Administrator. The Bursary Administrator then selects the candidates who conform to the requirements. The prerequisites are the following:

- a matriculation certificate, or a minimum of an N2 qualification; preferably an N3 qualification in the particular trade; and as part of the pre-selection criteria, the Technical College set a matriculation certificate or the completion of an N2 qualification in the relevant trade as its minimum requirements for entering the selection process)
- the completion of the four relevant subjects for the particular trade (in this study, the electrical trade) – maths, electronics, electro trade theory and engineering science.

All the applicants who conform to the requirements then do the TTB (SHL, 1996), and the LPCAT (De Beer, 2000). Candidates who measure up to the pre-determined norms attend a structured interview. Successful candidates must also pass a medical examination for final selection. Equal weights are assigned to all measuring instruments.

The objective of this study is to determine the empirical validity of the TTB and the LPCAT test results as predictors of the number of attempts a student will make to pass the various phase tests during the first year of the electrician learnership.

The selection process can be summarised as a successive-hurdle technique. This means that in order to be selected, candidates must successfully pass various screening steps. At each step or hurdle, some candidates are rejected (Van der Merwe, 2002). The selection process followed at the College is set out in Figure 1.

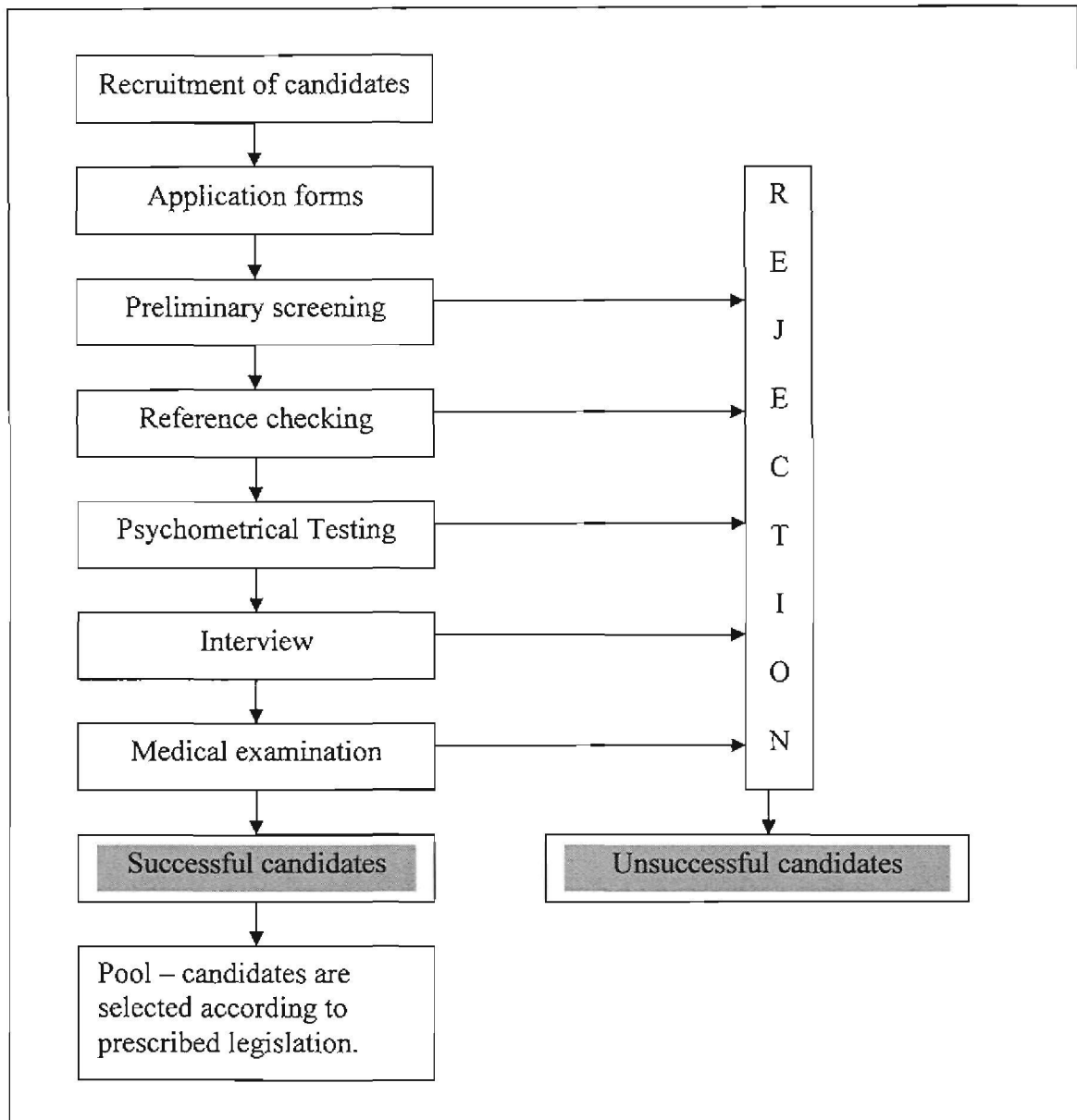


Figure 1. The selection process (adapted from Van der Merwe, 2002, p. 78).

One of the advantages of the TTB and the LPCAT is that they can be administered to a group. First, the subtests of the TTB are administered. The administration procedures and sample questions are on separate administration cards. The tests are all in a multiple-choice, paper and pencil format and they can be scored by hand or machine. The multiple-choice format and customised answer sheets permit quick scoring using plastic scoring key overlays (SHL, 2000). The TTB can also be scored using the NCS “Opscan5®” optical mark reader, which is connected to a personal computer.

The TTB is norm-referenced. In norm-referenced tests, a person's performance is compared to that of a group of comparable individuals (Huysamen, 1998). Norm tables (percentiles and T-scores) are provided for all the TTB tests. According to De Jager (personal communication, February 7, 2006), the norm line used by this particular training institution was developed by providing SHL with information (the raw scores of testees) on applicants tested between 2003 and 2006.

It is important to note that the TTB is supported by SHL's ongoing research and development programme. Accredited test users are regularly updated with recent norms and validation studies from a range of organisations (SHL, 2000). The tests are not freely available. They are distributed by SHL under strict control to suitably trained and qualified test users (SHL, 2000).

After a break of half an hour, the LPCAT is administered by means of computers. The LPCAT results consist of four different scores (De Beer, 2000):

- the pre-test score (T-score, stanine, percentile score);
- the post-test score (T-score, stanine, percentile score);
- the difference score (T-score); and
- the composite score (single score on T-score scale).

The composite score consists of a justifiable and reasoned combination of the first three scores. This allows for easier comparison of the cognitive developmental levels of the different candidates (De Beer, 2000). The composite score is used by the College.

The results of the TTB and the LPCAT are then interpreted by an industrial psychologist. The industrial psychologist presents a written recommendation to the Bursary Administrator. The Bursary Administrator and the Senior Training Officer for the particular trade interview the successful candidates. After the interview, the selected candidates are placed in a pool. They are selected from this pool into the learnership training programme, in line with the requirements of the Employment Equity Act (South Africa, 1998a) and the Skills Development Act (South Africa, 1998b).

### 3.2.3.1 *The TTB*

The TTB is used because of the following benefits:

- it tests relevant skills and reduces the risk of selecting poor performers;
- it saves time by ensuring that time and resources are only invested in the best candidates throughout the remainder of the selection process;
- it reduces costs by taking forward only the best candidates;
- it identifies key talent more quickly and helps to make selection decisions faster;
- it is an objective method of assessment which keeps the selection process fair and equitable for all candidates; and
- it is a psychometrically sound and efficient selection instrument.

See **Appendix A** for information on the reliability and validity of the test battery. A number of international validation studies have been carried out on it, including that of Henry, Hawkey, and Baron (2000). In a certain sense the TTB can thus be viewed as a so-called “old generation” test, but it is still in active use today.

The emphasis in the TTB has always been on the assessment of relevant work-related skills rather than on abstract conceptions of human intelligence (Henry et al., 2000). It is a series of ability tests designed specifically for the selection and development of individuals in technically or practically oriented jobs (SHL, 2000). The test battery consists of eight different tests – Verbal Comprehension, Numerical Computation, Visual Estimation, Mechanical Comprehension, Technical Understanding, Numerical Reasoning, Fault Diagnosis and Spatial Recognition (SHL, 2000). This specific training college only uses four of these tests, namely Numerical Reasoning, Mechanical Comprehension, Fault Diagnosis and Spatial Recognition. The TTB tests are designed to be used separately or in any combination. The selected tests were identified by the resident industrial psychologist using the guidance provided in the TTB manual and by the SHL consultants as most applicable for the selection of artisans for this College. More detailed descriptions of the four selected TTB tests used in the selection of apprentices in the electrician trade are set out below.

- Numerical Reasoning

Simple reasoning skills with numbers, an emphasis on understanding, reasoning and recognising shortcuts to reach solutions are examined. The test consists of written problems, expressed within the context of technical work. Percentages, fractions, decimals and diagrams, as well as knowledge of the four basic rules of arithmetic form part of the test.

Candidates must spend more time on generating and choosing strategies to solve numerical problems than on calculating their answers. They may use calculators in the test. The Numerical Reasoning Test consists of twenty five (25) items. The time allowed to complete the test is twelve (12) minutes (SHL, 2000).

- Spatial Recognition

The candidate's ability to recognise and match shapes in various orientations is tested. This skill is crucial in the design, manufacture and assembly of individual components and machinery.

Candidates must choose the one figure from a set of five that is identical to a given shape. The choices are often rotated relative to the given pattern and the test is organised in such a way that progressively finer discriminations are required to select the identical figure (SHL, 2000). This ability is important in the design, production and maintenance of machine tools, engineering components, mechanical sub-assemblies and other activities involving designing, fitting and inspection (SHL, 2000).

There are thirty six (36) items in this test. The time allowed to complete the test is ten (10) minutes.

- **Mechanical Comprehension**

This test focuses on spatially oriented items. The candidate must visualise movement, rather than recall situations obviously based on “textbook” laws of physics. This test identifies aptitude and potential rather than current knowledge. It measures the understanding of basic mechanical principles and their applications to a number of devices, including pulleys, gears and levers – a core skill relevant in many technical jobs (SHL, 2000).

The test consists of thirty six (36) questions, each relating to a three-dimensional drawing. The time allowed to complete the test is eighteen (18) minutes.

SHL (2000) has identified Mechanical Comprehension as an area where group differences often occur. SHL (2000) believes that different levels of interest, different levels of involvement by men and women and by people of different ethnic backgrounds have a definite influence on differences in the test scores of applicants.

- **Fault Diagnosis**

This test assesses the ability to identify faults in logical systems. Equipment, especially that associated with electronics and computing, brings with it a lot of challenges in terms of maintenance and servicing (SHL, 2000). The ability to isolate faults and problems in a system is therefore important.

The test requires a candidate to locate which element in an arrangement of coded symbols is not working as specified. No specialised knowledge of faultfinding or electronics is required; in this way applicants with little or no technical background can also be assessed (SHL, 2000). The ability to find and locate faults is applicable in machine maintenance and electronics, software, process control systems and communications.

The test consists of thirty six (36) items. The time allowed to complete the test is eighteen (18) minutes.

### 3.2.3.2 *The LPCAT*

The Learning Potential Computerised Adaptive Test (LPCAT) (De Beer, 2000) is included in the selection process. The LPCAT is intended to serve as a screening instrument that can be used in the context of South Africa's past educational inequities (De Beer, 2000). The following features are specifically built into it (De Beer, 2000):

- It uses only figural non-verbal items. This method minimises the verbal-content or other school-related prior learning on the test performance.
- It provides a learning experience which allows applicants from disadvantaged backgrounds to indicate at what level they may be able to perform if relevant learning opportunities are provided. A test-train-test approach is followed. The applicant's initial level of performance and his or her improvement when interpreting his or her overall level of cognitive development are taken into account. A standardised training method is used to improve the comparability of their test scores.
- Item Response Theory-based (IRT) scoring ensures the psychometric characteristics of the LPCAT. The difference score between the pre-test and post-test performance is improved by the IRT-based scoring.
- Computerized adaptive testing saves administration time spent on interpreting the results.
- This test is developed in South Africa. Multicultural groups were used in the item analysis, standardisation and validation of the test. This helps to support the LPCAT's validity in multicultural cognitive assessment.

Murphy and Maree (2006) claim that the socio-economic and educational disadvantages that many students have experienced in the past highlight the need to focus on future potential rather than on current ability. The non-verbal, diagrammatical format of the LPCAT limits the cultural bias that would be present if candidates were to be required to respond to items in a second or third language (Lópes et al., 2001; Taylor, 1999). This test is therefore included in the selection process in an attempt to negate the effect of language ability, past educational influences and competency on test scores (Van Eeden, De Beer, & Coetzee, 2001).

Based on the results of a study by Hugo and Claassen (1991) and general international consensus on culture-fair test content, De Beer (2000) decided to use non-verbal items of the figural type only. Verbal items and number series were excluded on purpose in an attempt to negate the effects of prior learning (De Beer, 2000).

The item types chosen for the LPCAT were figure series, figure analogies and pattern completion. The three item types used in the LPCAT can be described as follows (De Beer, 2000):

- **Figure series**  
A series consisting of four figures, each in a square, is presented. A fifth square is empty. In the Figure series the applicant must be able to make inferences from a rule from the part given of the series and to complete the series accordingly.
- **Figure analogies**  
Two figures that correspond in a certain way are given in a block. In a second block, a third figure is given, and a fourth one must be selected so that it corresponds with the third figure in the same way as the second figure corresponds with the first. The first and third figures also correspond in a certain way (for example, size) – hence the second and fourth figures should correspond in the same way as the first and the third.
- **Pattern completion**  
A block consisting of nine squares (in three rows and three columns) is presented. In each row and column, three figures form a pattern. The candidate must find the right figure in the last row and column. The applicant must be able to make inferences from a rule from the given part of the series and to complete the pattern.

The test consists of 270 items (there are 90 items in each of the three item types). The items vary in terms of their level of difficulty. Item difficulty and ability are measured on the same scale and the level of difficulty of each item is determined beforehand (De Beer, 2000). Items are selected to match the estimated ability level of each individual. A unique set of items can be administered to every person. By taking the

ramifications of IRT into consideration, test scores are comparable, even when different sets of items are used to obtain the scores (Van Eeden, et al., 2001).

To standardise the test administration, all the test administrators receive relevant training. The difference between the present level of performance and the potential future level of performance can be interpreted as the applicant's undeveloped learning potential (De Beer, 2000). The focus is thus on the potential to improve performance and not only on current abilities.

The composite score of the LPCAT is used to make recommendations on the most suitable students. The students must have the potential to function at college level. (The results of the reliability and validity of the LPCAT are reported on in **Appendix B.**)

#### **3.2.4 Measurement of the criterion (academic success)**

The academic success of the first year students is used as the criterion variable (the theoretical part of the training).

The electrician's trade training which is the focus of this study consists of three disciplines: Electrical, Electronics and Programmable Logic Controllers. Each of the three disciplines has its own modules. The training is done by an accredited assessor in the particular trade and it starts with the Electrical training modules.

The Electrical modules consist of basic hand skills like filing, sawing, taping and drilling. After each module, there is a criterion test which is evaluated by the accredited assessor who did the training. Thematic modules are clustered together and a phase test is completed after all the collective modules have been completed. A phase test can be theoretical, or practical, or both. If the test is theoretical, it consists of the description of a practical process: a student must achieve 100% for the theoretical portion of the test, because one fault in a process can cost his or her "life". The test is scored by the assessor who did the training. A student must pass a prescribed module before he or she may continue to the next module. A student can make three attempts to pass a test. If the student fails the first time, a personal

performance appraisal is done. If he or she fails on the second attempt a personal appraisal is again done and he or she receives a written warning. If the student fails a third time, disciplinary action takes place and his or her contract is discontinued.

The Electronics and Programmable Logic Controller parts follow the same process, but with a different accredited assessor – one who is a specialist in the particular field – as the trainer and assessor of the students.

After all the modules and the phase tests have been completed, the student goes to a mine for one year's practical training. At the mine the student has to complete on-the-job modules that have to be signed off by his or her supervisor (who is a tradesperson in that particular field). After the practical year the student does a trade test that is scored by the Senior Training Administrator of the Technical College. After passing the final test, the student qualifies as an electrician.

### **3.2.5 Statistical analysis**

The statistical analysis in this study is done using Statistica version 7 (Statsoft Inc, 2007).

The data analysis is done using descriptive statistics (means, standard deviations, minimums, maximums, skewness, kurtosis, percentages, and frequencies). To investigate the criterion-related validity of the test battery, the correlation between the test data and the criterion data is determined by means of Pearson's Product Moment Correlation Coefficient. The criterion variable is rather small in range of scores (1 to 3), therefore the criterion variable is treated as a categorical variable. Analysis involves a discriminant analysis, t-tests and one way analyses of variance (ANOVAs).

The analyses also provides information about effect sizes and the statistical significance associated with predictor-criterion relationships. To detect the effect size, a number independent of the original measurement unit is needed (Cohen, 1988). This procedure is performed by standardising the raw effect size as expressed in the measurement unit of the criterion variable by dividing it by the standard deviation of the measures in the respective populations (the standard deviation is also in the

original measurement unit). The effect size of the difference between the means is referred to as the d-statistic. The d-statistic is the difference between the means of the two variables, divided by the pooled standard deviation. To supply a common conventional frame of reference, the terms “small”, “medium” and “large” effect size is used for the magnitude of the differences between means. A small effect size is  $d \sim 0,20$ . This implies that the phenomena under investigation are either not under good experimental control or not under good measurement control, or both. A medium effect size is  $d \sim 0,50$ . A large effect size is  $d \sim 0,80$  or larger. Such a large effect size implies that two populations are so different that almost half of their areas do not overlap (Cohen, 1988). Effect sizes (based upon differences between means) can assist the researcher in making professional judgements about the strength of predictor-criterion relationships (Cohen, 1988). These judgements are based on the premise that a difference of at least a medium effect between two means is indicative of a difference that is of practical significance. Ideally, a researcher should perhaps only feel satisfied when the difference between two means indicates a large effect.

Relationships vary in strength and direction. Strength is concerned with how closely the two sets of data are related. The direction of relationships refers to whether one variable increases when the other increases (positive) or whether one variable decreases when the other increases (negative) (SHL, 2004). To provide a researcher with a frame of reference in which to appraise differences in the degree of correlation, specific values relating to a “small” effect size ( $\sim 0,10$ ), a “medium” effect size ( $\sim 0,30$ ) and a “large” effect size ( $0,50+$ ) serve as definitions and express relationships (Cohen, 1988).

To determine the degree of relationship between the scores obtained on the TTB and/or the LPCAT and the students’ passing or failing the various modules and phase tests, the Pearson’s Product Moment Correlation coefficient is used. The correlation coefficient obtained in this way is known as the validity coefficient. The higher such a (positive) correlation, the better the criterion-related validity of the instrument. (Huysamen, 1998). T-tests, ANOVAs and discriminant analysis are used to determine which independent variable(s) used independently (the TTB or the LPCAT), or in combination (the TTB and the LPCAT) best predict(s), the dependent variable (the success rate of first year apprentice electrician students). To measure the relationships

referred to above, the mean test scores of the groups who needed one, two or three attempts respectively are compared to investigate whether there are any significant differences. This is done by means of one-way analyses of variance (ANOVAs) (where the sample sizes in the various criterion groups are sufficient) and t-tests (where the sample sizes are not sufficient). A discriminant analysis is performed to establish whether the predictor test results will in fact predict the academic success of first year apprentices in the electrical trade. The criterion variable is considered to be a categorical ordinal for the purpose of these analyses.

In this study, the criterion is restricted in range. Guilford (1965:342) stated that the coefficient of validity in a restricted group is “almost invariably smaller than what it would be in an unrestricted group”. In such cases, the validity coefficients are therefore an underestimate of their true value (SHL, 2004). The formula used in this study to correct for the restricted range is the following (SHL, 2004, p.12):

$$r_{xy} = \frac{SD_y}{SD_x}$$

$$r_{cr} = \sqrt{1 - r_{XY}^2 + r_{XY}^2 \left( \frac{SD_y^2}{SD_x^2} \right)}$$

Where

$r_{cr}$  = Validity (correlation) coefficient corrected for restricted range

$x$  = TTB and LPCAT scores of all applicants

$X$  = TTB and LPCAT scores of selected applicants

$Y$  = Criterion scores (academic success) of selected applicants

$r_{XY}$  = Validity coefficient ,

$SD_y$  = Standard deviation of the TTB and LPCAT scores of all applicants

$SD_x$  = Standard deviation of the TTB and LPCAT scores of selected applicants

#### **4. RESEARCH PROCEDURE**

The Technical College where this study is done has four intakes per year for the electrician trade, namely in January, May, June and October. Twelve (12) successful candidates are selected per intake. The students have to complete six modules per year. The Bursary Administrator receives a progress and evaluation report on each student after the completion of each module.

The researcher linked the TTB and LPCAT scores with the number of attempts of each of the selected apprentice students to pass a phase test.

#### **5. DIVISION OF CHAPTERS**

Chapter 1: Introduction, problem statement and objectives

Chapter 2: Article

Chapter 3: Conclusions, limitations and recommendations

#### **6. CHAPTER SUMMARY**

In this chapter, the problem statement and research objectives have been discussed. The researcher has described the research method and specific assessment instruments in detail. The different data-analysis methods used and the research procedure have also been described.

Chapter 2 consists of a research article. The current employment situation in South Africa, issues relating to student selection in the technical field and various challenges in the selection process are discussed. The focus of the article is on the empirical validity of the TTB and the LPCAT as predictive selection instruments used with apprentice electricians selected for training at a Technical College in the Northern Cape, Republic of South Africa.

## 7. LIST OF REFERENCES

- Anon. (2005/2006). Services Sector Education and Training Authority (SETA): Producing learnerships that are fit for the purpose. *South Africa Business Guidebook*, 9, 134-135.
- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice-Hall.
- Andrews, S. G. (1996). *A small-scale investigation of the group administration of Feuerstein's learning potential assessment device*. Unpublished M.Ed. thesis, University of Natal (KwaZulu-Natal), Pietermaritzburg.
- Baron, H., & Butterworth, A. (1998, 1 June). *Good practice for fair selection*. Paper presented at the BPS Test User Conference, Thames Ditton, UK.
- Boolsen, M., & Theron, C. C. (1996). Kennisbasiswa, gebruikspatrone en gebruikshindernisse ten opsigte van die riglyne vir die ontwikkeling en validering van personeel-keurings-prosedures. *Tydskrif vir Bedryfsielkunde*, 22(2), 7-14.
- Brummer, H. L., Badenhorst, J. A., & Neuland, E. W. (2005, December). An evaluation of the most important competitive analysis methods applied by global mining firms to determine the future intent of a competitive force. *Southern African Business Review*, 10(3), 19-29.
- Cascio, W. F. (1997). *Applied psychology in human resource management*. Englewood Cliffs, NJ: Prentice-Hall.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum & Associates.
- Cummins, J. (1984). *Bilingualism and special education: Issues in assessment and pedagogy*. Clevedon, UK: Multilingual Matters.
- Davies, T., & Farquharson, F. (2004). Implementing learnerships: Learner recruitment and selection – lessons learnt from the KwaZulu-Natal pilot projects. *South African Journal of Higher Education*, 18(1) (sic), 336-352.
- De Beer, M. (2000). *Learning Potential Computerised Adaptive Test (LPCAT): Technical Manual*. Pretoria: Unisa Press.
- De Beer, M. (2005). Development of the Learning Potential Computerised Adaptive Test (LPCAT). *South African Journal of Psychology*, 35(4) (sic), 717-747.

- De Jaeger, S., Hattingh, S., & Huster, E. (2002). *How to implement a learnership: A guide for employers and training providers*. Commissioned by the German Technical Co-operation in partnership with MERSETA. Pretoria: GTZ.
- De Jong, A., & Visser, D. (2000). Black and White employees' fairness perceptions of personnel selection techniques. *South African Journal of Psychology, 30*(4), 17-24.
- De Villiers, A. B. (1999). *Disadvantaged student's academic performance: Analysing the zone of proximal development*. Unpublished D.Phil. thesis, University of Cape Town, Cape Town.
- DesJardins, S. (2002, June). *Studying the timing of student departure from college*. Paper presented at the Annual Forum of the Association for Institutional Research (AIR), Cape Town.
- Friedenberg, L. (1995). *Psychological testing: Design, analysis, and use*. Amherst, MA: Allyn & Bacon.
- Grussendorff, S., Liebenberg, M., & Houston, J. (2004). Selection for the Science Foundation Programme (University of Natal): The development of a selection instrument. *South African Journal of Higher Education, 18*(1) (sic), 265-272.
- Guilford, J. P. (1965). *Fundamental statistics in psychology and education* (4th ed.). New York: McGraw-Hill.
- Henry, E., Hawkey, D., & Baron, H. (2000). *Technical Test Battery Manual and User's Guide*. Thames Ditton, UK: SHL Group.
- Hugo, H. L. E., & Claassen, N. C. W. (1991). *The functioning of the GSAT Senior for students of the Department of Education and Training*. Pretoria: Human Sciences Research Council.
- Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96*, 72-98.
- Huysamen, G. K. (1998). *Psychological measurement: An introduction with South African examples*. Pretoria: Van Schaik.
- Huysamen, G. K. (2002). The relevance of the new APA standards for educational and psychological testing for employment testing in South Africa. *South African Journal of Psychology, 32*(2), 26-33.
- Kearney, S. (2005, Sept). Dropouts: How you can arrest the slide. *Process and Control Engineering, 58*(8), 30.

- Kerlinger, F. N. (1986). *Foundations of behavioral research* (3rd ed.). Fort Worth, TX: Harcourt College.
- Leedy, P. D., & Ormrod, J. E. (2005). *Practical research* (8th ed.). New York: Merrill Prentice Hall.
- Lopes, A., Roodt, G., & Mauer, R. (2001). The predictive validity of the Apil-B in a financial institution. *Journal of Industrial Psychology*, 27(1), 61-69.
- López, J.C.L. (2004). *Multicriteria decision aid application to a student selection problem*. Unpublished M.Ed. thesis, University of Occidente, Sinaloa, Mexico.
- Mlambo-Ngcuka, P. (2007, March) Progress Report on Joint Initiative On Priority Skills Acquisition (JIPSA). *Government Communications (GCIS)* <http://www.pmg.org.za/briefings/briefings.php?id=325> Date of access: 18 April 2007.
- Mouton, J., & Marais, H. C. (1992). *Basic concepts: Methodology of the behavioural sciences*. Pretoria: Human Sciences Research Council.
- Murphy, R., & Maree, D. J. F. (2006). A review of South African research in the field of dynamic assessment. *South African Journal of Psychology*, 36(1), 168-191.
- Nunns, C., & Ortlepp, K. (1994). Exploring predictors of academic success in Psychology 1 at Wits University as an important component of fair student selection. *South African Journal of Psychology*, 24(4), 201-207.
- Outtz, J. L. (2002). The role of cognitive ability tests in employment selection. *Human Performance*, 15, 161-171.
- Owen, K., & Taljaard, J. J. (1996). *Handbook for the use of psychological and scholastic tests of the HSRC*. Pretoria: Human Sciences Research Council.
- Pelser, M., Bergh, Z. C., & Visser, D. (2005). The concurrent validity of learning potential and of psychomotor ability measures for the selection of haul truck operators in an open-pit mine. *South African Journal of Industrial Psychology*, 31(2), 58-70.
- Ployhart, R. E., Ziegert, J. C., & McFarland, L. A. (2003). Understanding racial differences on cognitive ability tests in selection contexts: An integration of stereotype threat and applicant reaction research. *Human Performance*, 16, 231-259.
- Robertson, I. T., Iles, P. A., Gratton, L., & Sharpley, D. (1991). The impact of personnel selection and assessment methods on candidates. *Human Relations*, 44, 963-982.

- Ryan, A. M. (2001). Explaining the Black-White test score gap: The role of test perceptions. *Human Performance*, 14, 45-75.
- Samuda, R. J., Kong, S. L., Cummins, J., Pascual-Leone, J., & Lewis, J. (1989). *Assessment and placement of minority students*. Toronto: C.J. Hogrefe.
- Shochet, I. M. (1994). The moderator effect of cognitive modifiability on a traditional undergraduate admissions test for disadvantaged black students in South Africa. *South African Journal of Psychology*, 24(4), 208-215.
- SHL. (1996). *Assessment materials*. Thames Ditton, UK: SHL group.
- SHL. (2000). *Technical Test Battery: Manual and User's Guide*. Thames Ditton, UK: SHL group.
- SHL. (2004). *Occupational Testing: Technical Manual*. Thames Ditton, UK: SHL group.
- SHL. (2006, May). Are tests fair? *Newsline*. Pretoria: SHL People Performance.
- Snelgar, R. J., & Potgieter, T. (2003). An exploratory study of group adverse impact in a recruitment and selection strategy. *SAJEMS NS*, 6, 89-98.
- Society for Industrial & Organisational Psychology of South Africa. (2005). *Guidelines for the validation and use of assessment procedures for the workplace*. Pretoria: SIOPSA.
- Statsoft Inc. (2001). *Statistica (data analysis software system), version 6*. Tulsa, OK: Statsoft Inc.
- Struwig, F. W., & Stead, G. B. (2001). *Planning, designing and reporting research*. Cape Town: Maskew Miller Longman.
- SOUTH AFRICA. (1996). New Constitution 108 of 1996. Pretoria: Government Printer.
- SOUTH AFRICA. (1998a). Employment Equity Act 55 of 1998. Pretoria: Government Printer.
- SOUTH AFRICA. (1998b). Skills Development Act 97 of 1998. Pretoria: Government Printer.
- SOUTH AFRICA. (1999). Skills Development Levies Act 9 of 1999. Pretoria: Government Printer.
- SOUTH AFRICA. (2007). Department of Labour. Report study of scarce skills and critical skills needs in the mining and minerals sector. Final report 2007. <http://www.mqa.org.za/content.asp?subID=37> Date of access: 14 April 2007.

- Taylor, T. R. (1994). A review of three approaches to cognitive assessment, and a proposed integrated approach based on a unifying theoretical framework. *South African Journal of Psychology*, 24(4), 184-193.
- Taylor, T. R. (1999). *Administrator's manual for Tram-1 Battery: Assessment of Learning Potential for Semi-literates*. Johannesburg: Arolab.
- Van Eeden, R., De Beer, M., & Coetzee, C. H. (2001). Cognitive ability, learning potential, and personality traits as predictors of academic achievement by engineering and other science and technology students. *South African Journal of Higher Education*, 15, 171-179.
- Van der Merwe, G. (2005). *The construct validity of the first view<sup>TM</sup> for selection purposes in South Africa*. Unpublished Master's Dissertation, University of Pretoria, Pretoria.
- Van der Merwe, R. P. (2002). Psychometric testing and human resource management. *South African Journal of Industrial Psychology*, 28(2), 77-86.
- Visser, A. J., & Hanslo, M. (2005). Approaches to predictive studies: Possibilities and challenges. *South African Journal of Higher Education*, 19, 1160-1176.
- World Economic Forum. (2001/2002). *World competitiveness report*. Available at: <https://members.weforum.org/site/homeprivate.nsf/Content/Africa+Economic+Summit+2003%5C-Coompetitiveness+Profiles> Date of access: 25 Sept. 2006.

# **THE EMPIRICAL VALIDITY OF AN ASSESSMENT BATTERY FOR APPRENTICE ELECTRICIAN STUDENTS**

## **Abstract**

In South Africa there are several critical skills shortages. Artisans' and technical skills have been identified as two priority skills areas. Appropriate selection batteries can assist technical colleges in selecting candidates who will be successful in their training. To ensure fairness and accountability, the Technical Test Battery (TTB) as an ability test and the Learning Potential Computerised Adaptive Test (LPCAT) as a learning potential test were used as part of a selection battery for apprentice electricians. The objective of this study was to assess the empirical validity of the TTB and LPCAT as predictors of academic success of first-year apprentice electrician students at a South African technical college serving a mining community. The TTB and LPCAT scores of a sample of 89 selected apprentice electricians were compared to the number of attempts they needed to pass each of eight phase tests (the success criteria). The data were analysed by means of descriptive statistics. Pearsons Product Moment Correlation Coefficient, t-tests, ANOVAs and discriminant analysis were used in the statistical analysis. Statistically significant relationships were found between the predictor and criterion variables. The results confirm that the TTB and LPCAT are indeed empirically valid tests that can be used in the selection of apprentice electricians.

# **DIE EMPIRIESE GELDIGHEID VAN 'N KEURINGSBATTERY VIR EERSTEJAAR VAKLEERLINGE IN DIE ELEKTRIESE AMBAG**

## **Opsomming**

In Suid-Afrika bestaan daar verskeie kritieke vaardigheidstekorte. Ambags- en tegniese vaardighede is as twee prioriteit-bevoegdheidsvelde geïdentifiseer. Toepaslike keuringsbatterye kan tegniese kolleges help om kandidate te keur wat hulle opleiding suksesvol sal deurloop. Om regverdigheid en verantwoordbaarheid te verseker, is die Tegnieuse Toetsbattery (TTB) as 'n vermoëtoets en die Gerekenariseerde Leerpotensiaal-aanpasbaarheidstoets (LPCAT) as 'n leerpotensiaaltoets as deel van 'n keuringstoetsbattery vir vakleerlinge in die elektriese ambag ingesluit. Die doelwit van hierdie studie was om die empiriese geldigheid van die TTB en die LPCAT te bepaal as voorspellers van die akademiese sukses van eerstejaar vakleerlinge in die elektriese ambag aan 'n Suid-Afrikaanse tegniese kollege wat 'n myngemeenskap bedien. Die TTB- en LPCAT-tellings van 'n steekproef van 89 geselekteerde vakleerlinge in die elektriese ambag is vergelyk met die getal pogings wat hulle moes aanwend om elk van agt fasetoetse (die sukseskriteria) te slaag. Die data is met behulp van beskrywende statistiek ontleed. Die Pearsons-produkthemomentkorrelasiekoëffisiënt, t-toetse, ANOVA's en 'n diskriminantanalise is in die statistiese analise gebruik. Statisties betekenisvolle verbande is tussen die voorspellings- en kriteriumveranderlikes gevind. Die resultate bevestig dat die TTB en LPCAT inderdaad empiries geldige toetse is wat in die keuring van vakleerlinge in die elektriese ambag gebruik kan word.

## **PROBLEM STATEMENT**

### ***South Africa's skills needs in a global environment***

The World Competitiveness Report (World Economic Forum, 2001/2002) indicates that South Africa is currently experiencing a serious employment crisis. In the last decade or so, local organisations have experienced some of the greatest changes in their history (Brummer, Badenhorst, & Neuland, 2005). Commodity price changes and exchange rate fluctuations, global opportunities, global competition, mergers, takeovers and restructuring are only a few of the critical issues that organisations currently have to face on an ever-increasing scale (Brummer et al., 2005). Aside from such complex and competitive macro-environmental influences, organisations must also meet the challenges posed by local factors such as South Africa's new Constitution (South Africa, 1996), a changed and changing political dispensation, the promulgation and implementation of a series of new Acts of Parliament to regulate matters pertaining to labour (Lopes, Roodt, & Mauer, 2001) and the severe skills shortages experienced throughout the country (Mlambo-Ngcuka, 2007). There are thus a multitude of influences and pressures which affect organisations' decisions regarding their management of global and local demands and forces. To create a sustainable competitive advantage in this complex environment, organisations must analyse financial factors, their competitors and their employees. Such factors can have both a short- and a long-term impact.

In South Africa there is a serious shortage of the skilled, qualified and experienced people needed to fill particular roles and take up particular professions, occupations or specialisations in the labour market (Mlambo-Ngcuka, 2007). A significant proportion of the South African population lacks the basic competencies or skills needed to meet either the new or the existing challenges in organisations (Davies & Farquharson, 2004). In 2006, five high-profile priority skills areas were identified for immediate attention; amongst them are the skills needed by artisans and technical skills (Mlambo-Ngcuka, 2007). The economy currently produces only approximately 5 000 artisans per year, but research suggests that at least 12 500 artisans need to be produced each year over the next four years to meet the demand (Mlambo-Ngcuka, 2007). This is a projected increase of 30 000 artisans between 2007 and 2010. A number of priority trades and the requirements for each have also been identified. The work done by electricians is one of the identified trades (Mlambo-Ngcuka, 2007).

An analysis of the workplace skills plans (WSPs) for the period from 1 April 2006 to 31 March 2007 has revealed that employers in the Mining and Mineral Sector could not fill as many as 12 422 positions, the equivalent of 2.6% of the total employment figures (Anon, 2005/2006). Most of the hard-to-fill vacancies were in the occupational categories for technicians and trades workers, where there were 5 159 vacant positions (Anon, 2005/2006).

Since the establishment of the Mining Qualification Authority (MQA), the percentage of companies in the Mining and Mineral Sector that have been affected by skills shortages has remained more or less constant at 50%. A survey (Anon, 2005/2006) even indicated a marked increase in skills shortages in 2006 compared to 2004 and 2005. In the trade workers occupational category of the MQA study, adequately qualified electricians constituted 16% of the scarce skills that were identified (Anon, 2005/2006). It is interesting to take note of the following statement in the latest report on skill shortages that Mlambo-Ngcuka (2007) presented to the MQA: “From 2000 to 2005, the research results showed that skills shortages were strongly linked to the need for the sector to change the racial and gender profile of the professional, managerial and technicians and trades workers categories. In the current survey employers seemed less concerned with equity issues and tended more to report general shortages in the market. This indicates a worsening of the skills shortage situation”.

The Skills Development Act of 1998 (South Africa, 1998b) and the Skills Development Levies Act 9 of 1999 (South Africa, 1999) are designed to assist to address the above-mentioned challenge (Davies & Farquharson, 2004). Since these Acts have been implemented, urgent attention is being given to the alignment and articulation of the various training pathways that are available. In the final report on scarce and critical skills to the MQA, it was recommended that the mining sector should develop strategies to retain more of the identified scarce skills candidates (Anon, 2005/2006). According to this report, skills shortages in the technicians’ and trades’ workers category should be dealt with by means of the promotion of learnerships among employers.

Learnerships are one of the key components of the Act (Anon, 2005/2006; South Africa, 1998b). A learnership is defined by De Jaeger, Hattingh and Huster (2002) as “a route to a nationally recognised qualification that relates to an occupation and consists of a structured learning component and practical work experience.” Every Sector Education and Training Authority (SETA) must develop a skills plan. In that sector skills plan, the labour market

conditions and skills needs in each sector and the supply of skills to each sector must be described (Department of Labour, 2005).

The Mining and Mineral Sector recommended promoting learnerships and apprenticeships in Further Education and training colleges by expanding incentives to enrol more learners and developing more skills programmes (Mlambo-Ngcuka, 2007). Employers benefit from getting employees who have been well trained from both a theoretical and practical point of view, and they can claim grants each time a learnership agreement is entered into, which helps them cover some of the training costs (Anon, 2005/2006). However, although the government incentives offered to promote employer involvement may seem enticing to many organisations, many grants are only awarded once the candidate completes the learnership. If an apprentice therefore leaves before completing the learnership, the organisation must re-hire and re-train candidates and must cover expenses on these incentives (Kearney, 2005). This suggests that organisations need to select candidates very carefully.

According to Kearney (2005), only 60% of candidates complete their learnerships. Studies by DesJardins (2002) and Kearney (2005) indicate that the retention, dropout and throughput of students are vital to academic planning and central to issues of the implied costs that are imposed on society, the educational institution and the individual. DesJardins (2002) argues that the dropout of “disadvantaged students” may result in a social cost because it may lead to further racial and socio-economic disparities in future generations. For the institution, there is a cost because imbalances between the intake and the completion of training have considerable budgetary and financial implications. Employers will not find suitably qualified employees. For an individual, the cost of leaving the educational institution without successfully completing his or her training is a potential loss in earnings and being sidelined when he or she applies for a job, not to mention the personal and emotional issues related to dropping out for academic reasons (Visser & Hanslo, 2005). However, Kearney (2005) maintains that the problem can be addressed through rigorous selection processes, as he argues that accurate selection procedures can predict the future performance of candidates. Doing so may help the government and training facilities to reach its annual target of 12 500 trained artisans, and may save on training costs by choosing only those the candidates that have a good chance of succeeding in their training.

### ***Selection issues in South Africa***

Given the unemployment issues and the quest for people to fill the skill shortage gaps, it is important to find possibilities to solve the labour problems. Unfortunately, the solution is not simple. Organisations face multiple challenges when they make selection decisions (Pelser, Bergh, & Visser, 2005; Nunns & Ortlepp, 1994). Two key reasons for investing time and effort in selection procedures in South Africa are

- remaining within the ambit of the law (this refers to industrial relations, ethical and cultural issues, and the social utility of fairness considerations); and
- maximising the probability of selecting the potentially most productive candidate(s) (this relates to issues such as skills shortages, the impact of globalisation as experienced in the South African economy and the need for increased competitiveness).

Psychometric tests have been used for over a century. According to Taylor (1994), in 1882 Galton established a laboratory in London where people could be tested on visual acuity and reaction-time tasks. In 1905 a French psychologist, Binet, and his colleague Simon introduced the first intelligence tests (De Beer, 2005). Outtz (2002) indicates that tests designed to measure general mental ability have been used in employment selection for more than 80 years.

The quest for fair and valid assessment tools has long been pursued at advanced levels in societies such as those in countries like Canada, Australia, Britain and the United States of America (Cummins, 1984; Samuda, Kong, Cummins, Pascual-Leone, & Lewis, 1989). Both international and local research has demonstrated that psychometric assessment can play an important role in significantly improving the selection process for new entrants (Anastasi, & Urbina, 1997; Friedenber, 1995; Grussendorff, Liebenberg, & Houson, 2004; Lopes et al., 2001; Nunns & Ortlepp, 1994). Standardisation refers to the uniformity of the test material, the administration of and the scoring procedure for the instruments being used (Huysamen, 1998).

In the South African context, selection issues are unique and complex. The country has 11 official languages; many social and educational problems; and there are huge disparities in the socio-economic and educational backgrounds of individuals (De Beer, 2005). The labour situation has changed dramatically since 1994 (Lopes et al., 2001), and fair selection

practices are critical (Huysamen, 2002; Snelgar & Potgieter, 2003). Any discriminatory practices left over from the apartheid system have to be eradicated, and unsuitable instruments have to be replaced by measures designed to redress past injustices.

Every selection technique has its own merits, but research has confirmed the use of standardised tests to be the most psychometrically sound assessment method (Baron & Butterworth, 1998; Friedenberg, 1995). Psychometric testing is the most common procedure in the prediction of human behaviour during selection (Van der Merwe, 2005). Standardised tests are commonly employed as a selection aid. Research has shown that well-constructed psychometric tests

- predict job performance better than any other single selection method;
- provide objective information about a candidate;
- avoid the problems of subjective bias associated with other techniques such as interviews or references; and
- have been shown in general to lead to better and fairer employment decisions (SHL, 2006; Van der Merwe, 2002).

Several psychological tests have already been standardised in South Africa (Owen & Taljaard, 1996), but many were originally developed with white test takers in mind (Huysamen, 2002). In South Africa one of the major stumbling blocks regarding the use of psychometric tests stems from the complexity of creating fair selection practices for all culture groups (Huysamen, 1996; Huysamen, 2002; Snelgar & Potgieter, 2003).

Fair selection practices are critically important (Huysamen, 2002; Snelgar & Potgieter, 2003). Section 8 of the Employment Equity Act (South Africa, 1998a) emphasises the need for the reliability, validity and fairness of assessment procedures and states that such procedures may not be biased. It insists that “Psychological testing and other similar assessments of an employee are prohibited unless the test or assessment being used

- has been scientifically proven to be valid and reliable;
- can be applied fairly to all employees; and
- is not biased against any employee or group.”

From the above discussion, it is clear that both researchers and selection practitioners must ensure that the selection techniques that are used have a high degree of validity, fairness and reliability for a specific situation (Anastasi & Urbina, 1997). Perceptions of unfairness may lead to legal actions taken against the employer with resulting financial implications that may affect the bottom line (De Jong & Visser, 2000).

Organisations currently have no other option but to validate the tests and other assessment instruments that they use. This is the only way to prove that these instruments comply with the requirements set by the Employment Equity Act (Pelser et.al, 2005). According to Lopes, et al. (2001) and Boolsen and Theron (1996), an employer should be able to prove scientifically that

- the information gathered during the selection process is valid and job-relevant;
- the use of the selection technique makes it a business necessity;
- the information is combined in a fair manner for decision-making; and
- the selection efficiency justifies the use of the selection technique.

For the above reasons, effective selection procedures are of cardinal importance to any organisation. A dynamic selection process can assist the organisation in the measurement and prediction of the behaviour of individuals that suit its needs (Van der Merwe, 2005) and help the organisation to meet the challenges of the environment it is part of. Selecting employees is a complex decision-making process which requires a comprehensive evaluation of the applicant's performance (López, 2004). To make an objective decision, multiple selection criteria should be considered simultaneously (López, 2004).

Van der Merwe (2002) stated that the requirements of the Employment Equity Act are a definite move towards making selection decisions more scientific. Kriek (1998) already expressed a similar view four years earlier, commenting that appropriate legislation would improve assessment practices in South Africa.

Local and international research has found that people from disadvantaged backgrounds tend to score poorly when they are assessed by means of standard psychometric instruments (De Beer, 2005; Ployhart, Ziegert, & McFarland, 2003; Ryan, 2001; Shochet, 1994; Van Eeden, De Beer, & Coetzee, 2005). In order to address the issue of group representativity in

employment outcomes and to ensure equal representation in all occupational categories and levels in the workplace, Section 15 of the Employment Equity Act 55 of 1998 (South Africa, 1998a) includes affirmative action measures. DesJardins (2002) warns that the dropout of “disadvantaged students” may result in further racial and socio-economic disparity in future generations.

### ***Measurement of abilities and learning potential***

According to Taylor (1994), almost all published tests for use in industry and many of those for educational use are designed to measure constructs such as abilities. Conventional ability measurement instruments have increasingly come under fire for being biased or unfair to disadvantaged groups (Taylor, 1994). Traditional measures often used in the assessment of cognitive functions of individuals are regarded as unsuitable for two reasons. Firstly, a particular language is used in the test items and instructions of most cognitive tests, and the overall lower test scores of culturally diverse individuals are often indicators of poor verbal knowledge. Thus language proficiency can have an impact on the lower Full Scale IQ (Grey, 2000). Hugo and Claassen (1991) have established that large differences exist and may influence test scores when tests are administered in English to black pupils whose first language is an African language. Secondly, the verbal-item content in such cases may fail to provide important information regarding the potential learning and reasoning abilities of a person from a different culture (Grey, 2000).

The IQ measure is considered to reflect largely the individual’s previous learning experience and not any future potential for growth (Shochet, 1994). Conventional psychometric tests are therefore often not very suitable for the selection of students from disadvantaged backgrounds, because they are a static measure of current abilities which gives no indication of the student’s potential to learn (Van der Merwe & De Beer, 2006). Poor scores often reflect a lack of educational opportunities and not necessarily a lack of potential (De Beer, 2005; Visser & Hanslo, 2005). Tests of cognitive abilities generally measure the individual’s current level of functioning and rely heavily on the assumption that all examinees have had comparable opportunities to acquire the skills and abilities being measured (De Beer, 2005; Murphy & Maree, 2006; Shochet, 1994). It is interesting that cognitive ability tests produce racial differences that are three to five times larger than that of other predictors (Hunter & Hunter, 1984; Outtz, 2002). Notwithstanding that fact, an overview of literature identified

cognitive ability tests among the most predictive and practically efficient predictors of job performance for most occupations (Hunter & Hunter, 1984; Schmidt & Hunter, 1998).

It is now important to take note of the difference between standard selection processes and those performed in the context of a learnership (Davies & Farquharson, 2004). Standard approaches tend to focus on the “skills held” by the candidate and generally seek to select the “best candidate” for a particular position, based on existing qualifications and a set of criteria. In the case of learnerships, it is generally done in a developmental context. This means that the focus is on candidates with the greatest potential to develop their skills in order to meet the objectives of the learnership (Davies & Farquharson, 2004).

Given the socio-economic and educational disadvantages that many South African citizens have experienced in the past and the requirements of the Skills Development Act (South Africa, 1998b), local researchers have indicated that there is a clear need for assessment strategies to focus on future potential rather than on current ability (Andrews, 1996; De Beer, 2000; De Villiers, 1999). The assessment of learning potential can be more process-oriented, focusing on the learning that takes place during assessment (Hamers & Resing, 1993).

It is important to note that the assessment of learning potential does not only refer to the assessment of disadvantaged learners in South Africa (Murphy & Maree, 2006). It can be applied to any kind of learner who is struggling with one or more component(s) of cognitive functioning. Children who have not had an adequate learning experience or who find themselves in a different culture are also learners who can and do benefit from dynamic assessment measures (Kannevsky, 1995; Murphy & Maree, 2006). The learning potential theory assumes that potential may reside in learners; a potential that may remain untapped if the learner is assessed in a conventional way (Kannevsky, 1995; Murphy & Maree, 2006). Learning potential is often observed over a period of time, during which cognitive development takes place (Murphy & Maree, 2006). Murphy and Maree (2006) suggest that the measure of it should be applied in a complementary way for additional value-laden and qualitative information that cannot be obtained through the exclusive use of static measures. Testing for learning potential is not a method through which standards are lowered in order to accommodate larger numbers of disadvantaged students, it only seeks to identify those disadvantaged students who have the potential to perform well within the correct learning context and identifies learners who are likely to be overlooked if they are not assessed in this

manner (Murphy & Maree, 2006). Some researchers suggest that cognitive ability tests be combined with other predictors, like learning potential tests, in order to reduce the adverse impact while increasing overall validity (Outtz, 2002; Shochet, 1994; Taylor, 1994).

### ***Difference between ability and learning potential***

The preceding discussion brings us to the need of understanding the construct of ability and that of learning potential. According to Murphy and Maree (2006), the difference between the terms *potential*, *ability* and *aptitude* can be described as follows: the three terms represent different aspects of intellectual functioning and measurement. The term *potential* is taken from a process-oriented approach to assessment. It is included in a pre-test and post-test design. The learning or mediation in between is an indication of “*potential*” (De Beer, 2002; Murphy & Maree, 2006; Shochet, 1994). The terms *ability* and *aptitude* are more concerned with broad-based psychological constructs in the domain of intellectual functioning and take a more static or product-based approach to assessment.

A large volume of research has been conducted in the area of intelligence. In an article by Taylor (1994), he described the development and application background to the measuring of performance in the domain of intellectual ability, citing (Spearman (1927), Thurstone (1938), Burt (1949), Vernon (1962), Guilford (1967), Cattell (1971), Snow and Lohman (1984), Horn (1986) and Ackerman (1988). Each of the aforementioned psychologists produced a different structure of intellect (Taylor, 1994). In his article, Taylor (1994) mentioned that Cattell (1971) offered a higher-order theory which distinguished two forms of intelligence, namely, fluid and crystallised. Crystallised intelligence is a basic inherited capacity as developed by an interaction with environmental characteristics which are found in any society. Ability tests would be classified as measures of crystallised intelligence, thus be strongly affected by cultural influences and also schooling (Taylor, 1994). Fluid intelligence refers to specialised skills and knowledge promoted by and required in a given culture.

It is now known that intelligence is a complex and dynamic phenomenon (Levy, Robert, Gillio, Dee, Fenner, & Cornman-Levy, 1998). A study of intelligence conducted at Harvard University identified “multiple intelligences” (Levy et al., 1998). It was determined that intelligence is truly a multidimensional phenomenon and that there are at least seven types of intelligence (Levy et al., 1998). According to the aforementioned authors every individual possesses all seven types of intelligence but different people may possess different

capabilities in one or more forms of intelligence. Levy, et al. (1998) also stated that if any person is placed under proper guidance and instruction, all types of intelligence are likely to develop.

### ***Problem statement regarding the Technical College***

The Technical College that was reviewed in this investigation has replaced a traditional apprenticeship system with learnerships which supply students with a work-based route to a nationally recognised qualification within the National Qualifications Framework (NQF). The College provides a structured learning programme and combines this programme with workplace experience to contribute to the supply of artisans who have the necessary skills to meet the increased demand. Based on the local mines' occupational needs, particular mines provide learnerships. The mine concerned thus carries all training costs and gives each student a monthly allowance. The mine does not have any payback agreement with the student if he or she does not complete the training. It is therefore of the utmost importance for the College to select those candidates who

- are most likely to benefit from what is offered educationally;
- meet the requirements stipulated in South Africa's Labour legislation; and
- will perform most successfully in the specific trade.

After some market research, as part of the pre-selection criteria, the Technical College set a matriculation certificate or the completion of an N2 qualification in the relevant trade as its minimum requirements for entering the selection process. In the selection process itself, candidates then have to complete the Technical Test Battery (TTB) (SHL, 1996), the Learning Potential Computerized Adaptive Test (LPCAT) (De Beer, 2000), a structured interview and a medical examination. This forms a part of the selection battery for artisan students.

The two instruments used at the Technical College under review are the Technical Test Battery (TTB) and the Learning Potential Computerised Adaptive Test (LPCAT).

The Technical Test Battery (TTB) is an ability test battery which consists of a series of tests designed to be suitable for the selection and development of individuals in technically or practically oriented jobs (SHL, 1996). The Learning Potential Computerised Adaptive Test

(LPCAT) was developed in South Africa as a non-verbal, culture-fair measure of learning potential in the domain of cognitive or general non-verbal reasoning ability (De Beer, 2000). It takes into account the fact that examinees come to the testing situation from diverse and often disadvantaged backgrounds (De Beer, 2000).

No job analysis was done. The same selection criteria are used for all five different trades (electrician, fitter and turner, boilermaker, diesel mechanic and millwright) that are presented at the College.

The duration of the training for artisans is two years. All the selected students receive learnerships. The learnership training cost for one tradesperson is approximately R115 000. The first year's training is done at the training college and the second year consists of practical training at a mine. Students doing different courses complete different modules.

The electrical trade training consists of three disciplines: Electrical, Electronics and Programmable Logic Controllers. Each of the three disciplines has its own modules. The training is done by an accredited assessor in the particular trade and it starts with the electrical part of the training modules.

The electrical modules consist of basic hand skills like filing, sawing, taping and drilling. After each module, there is a criterion test which is evaluated by the accredited assessor who did the training. Thematic modules are clustered together and a phase test is completed after all the collective modules have been completed. A phase test can be theoretical, or practical, or both. If the test is theoretical, it consists of the description of a practical process: a student must achieve 100% for the theoretical portion of the test, because one fault in a process can cost the person his or her "life". The test is scored by the assessor who did the training. A student must pass a prescribed module before he or she may continue to the next module. A student can make three attempts to pass a test. If the student fails the first time, a personal performance appraisal is done. If he or she fails on the second attempt, a personal appraisal is again done and he or she receives a written warning. If the student fails a third time, disciplinary action takes place and his or her contract is discontinued.

The Electronics and Programmable Logic Controller parts follow the same process, but with a different accredited assessor – who is a specialist in the particular field – as the trainer and assessor of the students.

After all the modules and the phase tests have been completed, the student goes to a mine for one year's practical training. At the mine the student has to complete on-the-job modules that have to be signed off by his or her supervisor (who is a tradesman in that particular field). After the practical year the student does a trade test that is scored by the senior training administrator of the Technical College. After passing the final test the student qualifies as an electrician.

### ***Problem statement for this study***

The objective of this study was to determine the empirical validity of the two psychometric instruments (TTB and LPCAT) which form part of the investigated technical college's selection battery. Before the implementation of these specific tests, the pass rate of apprentices was about 60%; after the tests were implemented, the pass rate increased to more than 90%. The Technical College expressed a need to determine the empirical validity of its selection battery, particularly the TTB and the LPCAT. It is important for the Technical College to know to what extent the two selected tests predict the academic performance of first year electrician students and to support specific interpretations with evidence and theory. They also wanted to establish whether both tests are necessary in the selection process. The College wants the selection methods it follows to be fair in allowing opportunities for access and effective in predicting success in the learnerships it provides.

The information obtained in this study can be of value when training colleges select students for the electrician trade. Colleges will know if the TTB and LPCAT as selection instruments are worthy aids in the problem of selecting candidates who will be successful in their training. The costs of poor selection are enormous. If a training college can predict the success of a student's performance more accurately it will have a big impact on training costs and the utilisation of the training college's resources and it will make a positive contribution toward providing skilled people to meet the shortage the market is experiencing.

This study focuses on the predictive validity of the TTB and the LPCAT but, as far as the "selected candidates" are concerned, cognisance has to be taken of the role that the structured

interview and the medical examination play in this regard. Thus, the TTB and LPCAT scores attained by the “selected candidates” are those of a rather select group.

## **RESEARCH METHOD**

### ***Research design***

Kerlinger (1986) defines a research design as “the plan and structure of investigation so conceived as to obtain answers to research questions”. The plan is the overall programme of the research which outlines what the researcher will do. The structure is a paradigm or model of the relations among the variables of a study. The purpose of a research design is thus

- to provide answers to research questions (Kerlinger 1986);
- to control variance (Kerlinger, 1986); and
- to enhance the internal and external validity of the research (Mouton & Marais, 1992).

According to De Jong and Visser (2000), the empirical validity of research can be approached in the traditional way or via contemporary decision theory. If the traditional approach is used, a selection technique is validated in terms of the accuracy with which the technique predicts future performance (Cascio, 1997). Contemporary decision theory emphasises the utility of a certain selection decision; this implies that the measurement and prediction of performance is viewed as using technical data that has to be evaluated in terms of its utility for the organisation (Robertson, Iles, Gratton, & Sharpley, 1991). In this study the traditional approach is applied.

This research can be categorised as descriptive quantitative research. Descriptive research identifies the characteristics of a phenomenon or explores possible correlations between two or more phenomena (Leedy & Ormrod, 2005). It examines a situation as it is. Descriptive research does not involve changing or modifying the situation under investigation, nor is it intended to determine cause-and-effect relationships (Leedy & Ormrod, 2005). Quantitative research is a form of conclusive research in which the data collection procedures are fairly structured (Struwig & Stead, 2001). A correlation study is used to reach the objectives of this study. Leedy and Ormrod (2005) explain that a correlation study determines the extent to which differences in one characteristic or variable are related to differences in one or more other characteristics or variables. A correlation exists if, when one variable increases, another variable either increases or decreases in a somewhat predictable fashion.

The aim of this study was to determine how well two psychological tests (the TTB and the LPCAT) predict an individual's future performance. The predictor variables are the results of the scores on the TTB and the LPCAT. To define the criterion of success in this study, the success of an individual's performance was expressed as the number of attempts a student had to make to pass a phase test.

## STUDY POPULATION

Table 1

*Compilation of the study population (N = 89)*

Item	Category	% of all cases
Language	Afrikaans	65,91%
	English	5,68%
	African	28,41%
Gender	Male	93,26%
	Female	6,74%
Ethnicity	Black (Designated)	61,80%
	White (Non-designated)	38,20%
Qualification	Grade 11 & below (N2)	19,10%
	Grade 12	19,10%
	Post Matric Certificate (N4-N6)	61,80%

In Table 1 the compilation of the study population is demonstrated. The sample consisted of 89 selected apprentice electricians from 2003 to 2006. The total number of selected students associated with the designated group – which according to Schedule 4 of the Employment Equity Act (Act 55 of 1998) (South Africa, 1998a) includes (Africans, coloureds and Asians, people with disabilities, and women) – were 61,80%. The women selected formed only 6,74% of the sample. There were 38,20% selected students from the non-designated group – which includes white males without disabilities. Most of the sample were in possession of a Post Matric certificate (61,80%).

## **MEASURING BATTERY**

### ***Measuring instruments***

The TTB is used because of the following benefits:

- it tests relevant skills and reduces the risk of selecting poor performers;
- it saves time by ensuring that time and resources are only invested in the best candidates throughout the remainder of the selection process;
- it reduces costs by taking forward only the best candidates;
- it identifies key talent more quickly and helps to make selection decisions faster;
- it is an objective method of assessment which keeps the selection process fair and equitable for all candidates; and
- it is a psychometrically sound and efficient selection instrument.

(See **Appendix A** for information on the reliability and validity of the test battery.) A number of international validation studies have been carried out on it, including that of Henry, Hawkey and Baron (2000).

The emphasis in the TTB has always been on the assessment of relevant work-related skills rather than on abstract conceptions of human intelligence (Henry et al., 2000). It is a series of ability tests designed specifically for the selection and development of individuals in technically or practically oriented jobs (SHL, 2000). The test battery consists of eight different tests – Verbal Comprehension, Numerical Computation, Visual Estimation, Mechanical Comprehension, Technical Understanding, Numerical Reasoning, Fault Diagnosis and Spatial Recognition (SHL, 2000). This specific training college only uses four of the tests, namely Numerical Reasoning, Mechanical Comprehension, Fault Diagnosis and Spatial Recognition. The TTB tests are designed to be used separately or in any combination. The selected tests were identified by the industrial psychologist, employed by the mine, using the guidance provided in the TTB manual and by the SHL consultants, as most applicable for the selection of artisans for this College. More detailed descriptions of the four selected TTB tests used in the selection of apprentices in the electrician trade are set out below.

- **Numerical Reasoning**

Simple reasoning skills with numbers, an emphasis on understanding, reasoning and recognising shortcuts to reach solutions are examined. The test consists of

written problems, expressed within the context of technical work. Percentages, fractions, decimals and diagrams, as well as knowledge of the four basic rules of arithmetic form part of the test. Candidates must spend more time on generating and choosing strategies to solve numerical problems than on calculating their answers. They may use calculators in the test. The Numerical Reasoning Test consists of twenty five (25) items. The time allowed to complete the test is twelve (12) minutes (SHL, 2000).

- **Spatial Recognition**

The candidate's ability to recognise and match shapes in various orientations is tested. This skill is crucial in the design, manufacture and assembly of individual components and machinery.

Candidates must choose the one figure from a set of five that is identical to a given shape. The choices are often rotated relative to the given pattern and the test is organised in such a way that progressively finer discriminations are required to select the identical figure (SHL, 2000). This ability is important in the design, production and maintenance of machine tools, engineering components, mechanical sub-assemblies and other activities involving designing, fitting and inspection (SHL, 2000).

There are thirty six (36) items in this test. The time allowed to complete the test is ten (10) minutes.

- **Mechanical Comprehension**

This test focuses on spatially oriented items. The candidate must visualise movement, rather than recall situations obviously based on "textbook" laws of physics. This test identifies aptitude and potential rather than current knowledge. It measures the understanding of basic mechanical principles and their applications to a number of devices, including pulleys, gears and levers – a core skill relevant in many technical jobs (SHL, 2000).

The test consists of thirty six (36) questions, each relating to a three-dimensional drawing. The time allowed to complete the test is eighteen (18) minutes.

SHL (2000) identified Mechanical Comprehension as an area where group differences often occur. SHL (2000) believed that different levels of interest, different levels of involvement by men and women and by people of different ethnic backgrounds have a definite influence on differences in the test scores of applicants.

- **Fault Diagnosis**

This test assesses the ability to identify faults in logical systems. Equipment, especially that associated with electronics and computing, brings with it a lot of challenges in terms of maintenance and servicing (SHL, 2000). The ability to isolate faults and problems in a system is therefore important.

The test requires a candidate to locate which element in an arrangement of coded symbols is not working as specified. No specialised knowledge of faultfinding or electronics is required; in this way applicants with little or no technical background can also be assessed (SHL, 2000). The ability to find and locate faults is applicable in machine maintenance and electronics, software, process control systems and communications.

The test consists of thirty six (36) items. The time allowed to complete the test is eighteen (18) minutes.

The Learning Potential Computerised Adaptive Test (LPCAT) (De Beer, 2000) is included in the selection process. The LPCAT is intended to serve as a screening instrument that can be used in the context of South Africa's past educational inequities (De Beer, 2000). The following features are specifically built into it (De Beer, 2000):

- It uses only figural non-verbal items. This method minimises the verbal-content or other school-related prior learning on the test performance.
- It provides a learning experience which allows applicants from disadvantaged backgrounds to indicate at what level they may be able to perform if relevant learning

opportunities are provided. A test-train-test approach is followed. The applicant's initial level of performance and his or her improvement when interpreting his or her overall level of cognitive development are taken into account. A standardised training method is used to improve the comparability of the test scores.

- Item Response Theory-based (IRT) scoring ensures the psychometric characteristics of the LPCAT. The difference score between the pre-test and post-test performance is improved by the IRT-based scoring.
- Computerized adaptive testing saves administration time spent on interpreting the results.
- This test is developed in South Africa. Multicultural groups were used in the item analysis, standardisation and validation of the test. This helps to support the LPCAT's validity in multicultural cognitive assessment.

Murphy and Maree (2006) claim that the socio-economic and educational disadvantages that many students have experienced in the past highlight the need to focus on future potential rather than on current ability. The non-verbal, diagrammatical format of the LPCAT limits the cultural bias that would be present if candidates were to be required to respond to items in a second or third language (Lopes et al., 2001; Taylor, 1999). This test is therefore included in the selection process in an attempt to negate the effect of language ability, past educational influences and competency on test scores (Van Eeden, et.al, 2001).

Based on the results of a study by Hugo and Claassen (1991) and general international consensus on culture-fair test content, De Beer (2000) decided to use non-verbal items of the figural type only. Verbal items and number series were excluded on purpose in an attempt to negate the effects of prior learning (De Beer, 2000).

The item types chosen for the LPCAT were figure series, figure analogies and pattern completion. The three item types used in the LPCAT can be described as follows (De Beer, 2000):

- Figure series  
A series consisting of four figures, each in a square, is presented. A fifth square is empty. In the Figure series the applicant must be able to make inferences from a rule from the part given of the series and to complete the series accordingly.

- **Figure analogies**

Two figures that correspond in a certain way are given in a block. In a second block, a third figure is given, and a fourth one must be selected so that it corresponds with the third figure in the same way as the second figure corresponds with the first. The first and third figures also correspond in a certain way (for example, size) – hence the second and fourth figures should correspond in the same way as the first and the third.

- **Pattern completion**

A block consisting of nine squares (in three rows and three columns) is presented. In each row and column, three figures form a pattern. The candidate must find the right figure in the last row and column. The applicant must be able to make inferences from a rule from the given part of the series and to complete the pattern.

The test consists of 270 items (there are 90 items in each of the three item types). The items vary in terms of their level of difficulty. Item difficulty and ability are measured on the same scale and the level of difficulty of each item is determined beforehand (De Beer, 2000). Items are selected to match the estimated ability level of each individual. A unique set of items can be administered to every person. By taking the ramifications of IRT into consideration, test scores are comparable, even when different sets of items are used to obtain the scores (Van Eeden et al., 2001).

To standardise the test administration, all the test administrators receive relevant training. The difference between the present level of performance and the potential future level of performance can be interpreted as the applicant's undeveloped learning potential (De Beer, 2000). The focus is thus on the potential to improve performance and not only on current abilities.

The composite score of the LPCAT is used to make recommendations on the most suitable students. The students must have the potential to function at college level.

## **PROCEDURE**

The Technical College uses a standardised selection procedure. The selection procedure includes a combination of measuring instruments. For the pre-selection purposes, an applicant must send his or her Curriculum Vitae to the Bursary Administrator. The Bursary

Administrator then selects the candidates who conform to the requirements. The prerequisites are the following:

- a matriculation certificate;
- a minimum of an N2 qualification; preferably an N3 qualification in the particular trade; and
- the completion of the four relevant subjects for the particular trade (in this study, the electrical trade) – maths, electronics, electro trade theory and engineering science.

All the applicants who conform to the requirements then do the TTB (SHL, 1996), and the LPCAT (De Beer, 2000). Candidates who measure up to the pre-determined norms attend a structured interview. Successful candidates must also pass a medical examination for final selection. Equal weights are assigned to all measuring instruments.

The objective of this study was to determine the empirical validity of the TTB and the LPCAT test results as predictors of the number of attempts a student makes to pass the various phase tests during the first year of the electrician learnership.

The selection process can be summarised as a successive-hurdle technique. This means that in order to be selected, candidates must successfully pass various screening steps. At each step or hurdle, some candidates are rejected (Van der Merwe, 2002). The selection process followed at the College is set out in Figure 1.

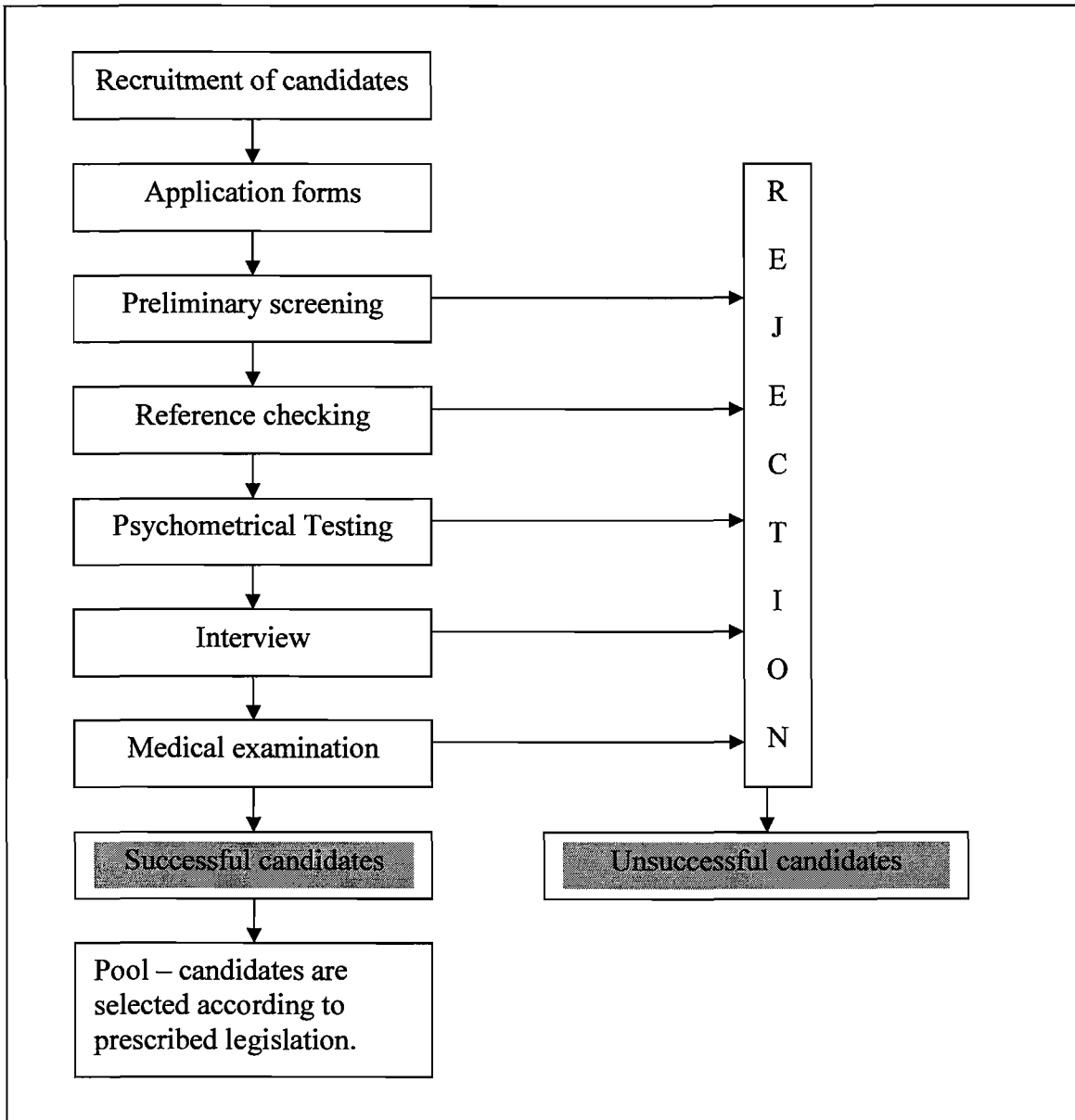


Figure 1. The selection process (adapted from Van der Merwe, 2002, p. 78).

## STATISTICAL ANALYSIS

The statistical analysis was done using Statistica version 7 (Statsoft Inc, 2007).

The data were analysed by means of descriptive statistics (means, standard deviations, minimums, maximums, skewness, kurtosis, percentages and frequencies). It was decided to investigate the criterion-related validity (in the form of empirical validity) of the measuring battery, the correlation between the test data and the criterion data by means of Pearson's Product Moment Correlation Coefficient. Since the criterion variable had a rather small

range of scores (1 to 3), and most respondents had one of two values (1 or 2) on this variable, analyses were also performed in which the criterion variable was treated as a categorical variable. These involved a discriminant analysis, t-tests and one way analyses of variance (ANOVAs).

The analyses also provided information about effect sizes and the statistical significance associated with predictor-criterion relationships. To detect the effect size, a number independent of the original measurement unit is needed (Cohen, 1988). This procedure is performed by standardising the raw effect size as expressed in the measurement unit of the criterion variable by dividing it by the standard deviation of the measures in the respective populations (the standard deviation is also in the original measurement unit). The effect size of the difference between the means is referred to as the d-statistic. The d-statistic is the difference between the means of the two variables, divided by the pooled standard deviation. To supply a common conventional frame of reference, the terms “small”, “medium” and “large” effect size were used for the magnitude of the differences between means. A small effect size is  $d \sim 0,20$ . This implies that the phenomena under investigation are either not under good experimental control or not under good measurement control, or both. A medium effect size is  $d \sim 0,50$ . A large effect size is  $d \sim 0,80$  or larger. Such a large effect size implies that two populations are so different that almost half of their areas do not overlap (Cohen, 1988). Effect sizes (based upon differences between means) can assist the researcher in making professional judgements about the strength of predictor-criterion relationships (Cohen, 1988). These judgements are based on the premise that a difference of at least a medium effect between two means is indicative of a difference that is of practical significance. Ideally, a researcher should perhaps only feel satisfied when the difference between two means indicates a large effect.

Relationships vary in strength and direction. Strength is concerned with how closely the two sets of data are related. The direction of relationships refers to whether one variable increases when the other increases (positive) or whether one variable decreases when the other increases (negative) (SHL, 2004). To provide a researcher with a frame of reference in which to appraise differences in the degree of correlation, specific values relating to a “small” effect size ( $\sim 0,10$ ), a “medium” effect size ( $\sim 0,30$ ) and a “large” effect size ( $0,50+$ ) serve as definitions and express relationships (Cohen, 1988).

To determine the degree of relationship between the scores obtained on the TTB and/or the LPCAT and the students' passing or failing the various modules and phase tests, the Pearson's Product Moment Correlation coefficient was used. The correlation coefficient obtained in this way is known as the validity coefficient. The higher such a (positive) correlation, the better the criterion-related validity of the instrument. (Huysamen, 1998). T-tests, ANOVAs and discriminant analysis were used to determine which independent variable(s) used independently (the TTB or the LPCAT), or in combination (the TTB and the LPCAT) best predict(s), the dependent variable (the success rate of first year apprentice electrician students). To measure the relationships referred to above, the mean test scores of the groups who needed one, two or three attempts respectively were compared to investigate whether there are any significant differences. This was done by means of one-way analyses of variance (ANOVAs) (where the sample sizes in the various criterion groups were sufficient) and t-tests (where the sample sizes were not sufficient). A discriminant analysis was performed to establish whether the predictor test results did in fact predict the academic success of first year apprentices in the electrical trade. The criterion variable was considered to be a categorical ordinal for the purpose of these analyses.

In this study, the criterion was restricted in range. Guilford (1965:342) stated that the coefficient of validity in a restricted group is "almost invariably smaller than what it would be in an unrestricted group". In such cases, the validity coefficients are therefore an underestimate of their true value (SHL, 2004). The formula used in this study to correct for the restricted range was the following (SHL, 2004, p.12):

$$r_{xy} = \frac{SD_y}{SD_x}$$

$$r_{cr} = \sqrt{1 - r_{XY}^2 + \left(\frac{SD_y^2}{SD_x^2}\right)}$$

Where

- $r_{cr}$  = Validity (correlation) coefficient corrected for restricted range
- $x$  = TTB and LPCAT scores of all applicants
- $X$  = TTB and LPCAT scores of selected applicants

Y = Criterion scores (academic success) of selected applicants

$r_{XY}$  = Validity coefficient

SD<sub>y</sub> = Standard deviation of the TTB and LPCAT scores of all applicants

SD<sub>x</sub> = Standard deviation of the TTB and LPCAT scores of selected applicants

## RESULTS

An independent sample t-test was performed in order to assess the differences between the mean test scores of the selected candidates and the non-selected candidates. The results of the t-test are reported in Table 2 below.

Table 2

*Differences between the mean scores of the selected candidates (N = 89) and the non-selected candidates (N = 283) on the subtests of the TTB and the LPCAT (t-test for independent samples)*

		Independent samples test					
		Levene's test for equality of variances		T-test for equality of means			
		F	Sig.	T	Df	Sig. (2-tailed)	Mean difference
		Lower	Upper	Lower	Upper	Lower	Upper
Spatial Recognition	Equal variances assumed	4,757	0,030	7,866	370	0,000	4,732
	Equal variances not assumed			8,725	177,773	0,000	4,732
Mechanical Comprehension	Equal variances assumed	0,003	0,959	5,687	370	0,000	3,382
	Equal variances not assumed			5,888	156,281	0,000	3,382
Fault Diagnosis	Equal variances assumed	0,117	0,732	9,733	369	0,000	7,637
	Equal variances not assumed			9,960	153,438	0,000	7,637
Numerical Reasoning	Equal variances assumed	2,052	0,153	9,923	370	0,000	3,409
	Equal variances not assumed			9,421	136,363	0,000	3,409
LPCAT	Equal variances assumed	17,876	0,000	4,797	370	0,000	3,403
	Equal variances not assumed			5,821	216,486	0,000	3,403

It is evident from the table above that all the differences were statistically significant ( $p < 0,001$ ). These differences were then investigated further by determining their effect sizes. The results are reported in Table 3 below.

Table 3

*Differences between the scores of the selected candidates (N = 89) and the non-selected candidates (N = 283) on the predictors*

		Selected candidates (N = 89)				Non-selected candidates (N = 283)				
		Mean	Min	Max	SD	Mean	Min	Max	SD	d-stat
TTB	Spatial	19,45	9	30	4,22	14,72	2	30	5,16	-0,89***
	Recognition									
	Mechanical	15,38	5	27	4,65	12,00	3	32	4,97	-0,66**
	Comprehension									
	Fault Diagnosis	20,57	9	34	6,24	12,94	0	36	6,52	-1,06***
LPCAT	Numerical	12,67	3	21	3,04	9,27	4	20	2,76	-1,07***
	Reasoning									
		57,70	46	69	4,26	54,29	33	72	6,25	-0,57**

\* d ~ 0,20 (small effect)

\*\* d ~ 0,50 (medium effect)

\*\*\* d ~ 0,80+ (large effect)

The sample consisted of selected and non-selected applicants for an apprentice electrician learnership at the Technical College from 2003 to 2006. To comply with the Employment Equity Act (South Africa, 1998a), 75% of the total number of selected students have to be members of the designated groups (Africans, Coloureds, Asians, people with disabilities and women) and 25% may be members of the non-designated group. At least 10% have to be women. From the information supplied in Table 3, it is clear that the difference between the scores of the selected and non-selected candidates on the subtests of the TTB had a medium to large effect (one of a medium and three of a large effect). The difference between the mean scores of the selected and the non-selected groups on the LPCAT indicates a medium effect. That implies that the differences between the test scores of the selected and non-selected groups were so different that almost half of their areas did not overlap. From the results

reported in Table 2 and Table 3, it can therefore be inferred that there are practically significant differences between the selected and the non-selected candidates with regard to their test results. It must be noted that the observed validity coefficients may underestimate the predictor-criterion relationship due to the effect of range restriction.

Table 4

*Descriptive statistics in respect of the predictor variables (N = 89)*

	Mean	Min	Max	SD	Skewness	Kurtosis
Spatial Recognition	19,45	9	30	4,22	0,32	-0,05
Mechanical Comprehension	15,38	5	27	4,65	0,20	-0,29
Fault Diagnosis	20,57	9	34	6,24	0,11	-0,82
Numerical Reasoning	12,67	3	21	3,04	-0,05	0,66
LPCAT	57,70	46	69	4,26	-0,31	0,38

In Table 4, descriptive statistics in respect of the predictor variables are given. To evaluate the mean of the test results on the subtests of the TTB meaningfully, the researcher used the mean of the raw scores obtained by the candidates and compared the results to those obtained by an appropriate reference group. This group is known as the norm group and the test norms refer to the data representing their performance on the TTB (SHL, 2000). The mean scores of the selected candidates on Spatial Recognition, Mechanical Comprehension and Numerical Reasoning indicate that the selected candidates' score is a six (6) when the score is expressed as a sten value. On Fault Diagnosis the mean score of the selected candidates has a seven (7) sten value. On the LPCAT, scores ranged from 46 (academic level – Grade 9 or NQF Level 1) to 69 (academic level – university or NQF Level 6). The mean score of the selected candidates on the LPCAT was 57,7. This implies that the level of performance or training which can be considered appropriate for the group is at a Technical College or Technikon level. Skewness is the degree to which a distribution departs from symmetry about its mean value (Keller & Warrack, 2003). If it leans to the left, it is positively skewed; and if it leans to the right, it is negatively skewed. The term kurtosis refers to the peakness or flatness of a frequency distribution. Perfectly normal distributions have skewness and kurtosis values equal to zero (Ployard, Weekley, Holtz, & Kemp, 2003). Table 4 indicates that the scores on the ability tests' skewness and kurtosis are acceptable ( $-1,0 < \text{skewness} \text{ and } \text{kurtosis} < +1,0$ ).

Table 5

*Matrix of correlations between the scores on the predictor variables (N = 89)*

		TTB				LPCAT
		Spatial Recognition	Mechanical Comprehension	Fault Diagnosis	Numerical Reasoning	
TTB	Spatial Recognition	1,00				0,16
	Mechanical Comprehension	0,38**	1,00			0,20
	Fault Diagnosis	0,36**	0,26**	1,00		0,26*
	Numerical Reasoning	0,29**	0,56**	0,31**	1,00	0,11
LPCAT						1,00

\*  $p < 0,05$

\*\*  $p < 0,01$

In Table 5, the correlations between the scores on the predictors are shown. In this study, the correlation between the sub-tests of the TTB and the LPCAT was lower than expected, as the highest correlation was that between the LPCAT and Fault Diagnosis (0,26). The high correlation between the LPCAT and Fault Diagnosis was to be expected, as none of the tests requires previous relevant educational background. The other correlations ranged from 0,11 to 0,56. The low correlations confirm that the two tests measure different constructs of intellectual reasoning. The highest correlation as far as the sub-tests of the TTB are concerned was that between Numerical Reasoning and Mechanical Comprehension ( $r = 0,56$ ;  $p < 0,01$ ), and it was of a large effect. This means that there is less than one chance in a hundred of obtaining this correlation by chance. The remainder of the correlations ranged from 0,26 to 0,38 (all of which are significant at the 0,01 level). These significant correlations are in line with those found in other studies (SHL, 2000). They are, however, not of practical significance..

Table 6

*Mean number of attempts needed to pass the different phase tests (criterion variables) (N = 89)*

Phase Test	N	Mean	Min	Max	SD
Phase Test 1	89	1,54	1	3	0,69
Phase Test 2	89	1,53	1	3	0,59
Phase Test 3	89	1,80	1	3	0,73
Phase Test 4	89	1,66	1	3	0,62
Phase Test 5	89	1,37	1	3	0,55
Phase Test 6	88	1,52	1	3	0,61
Phase Test 7	89	2,12	1	4	0,91
Phase Test 8	88	1,18	1	3	0,42

In Table 6 the mean number of attempts needed to pass the phase tests is set out. The phase test with the highest mean was Phase Test 7, with a mean of 2,12. This implies that Phase Test 7 may be considered the most difficult test that the students have to pass. In a personal discussion with N. Faihst, the accredited assessor of the Electronics section (personal communication, 30 April 2007), he explained that the students are exposed to the Electronics discipline for the first time during the period preceding Phase Test 7. They encounter instruments that are new to them and that are quite complex. This may influence their performance in this test. The phase test that required the fewest attempts to pass it, and which can therefore be considered to be the easiest test, is Phase Test 8, with a mean of 1,18. The content of the modules that the students have to complete to pass Phase Test 8 is basically the same as the content required to pass Phase Test 7. All the information on electronics is clustered, and is recapitulated a few times in a more familiar and understandable way before the students do Phase Test 8.

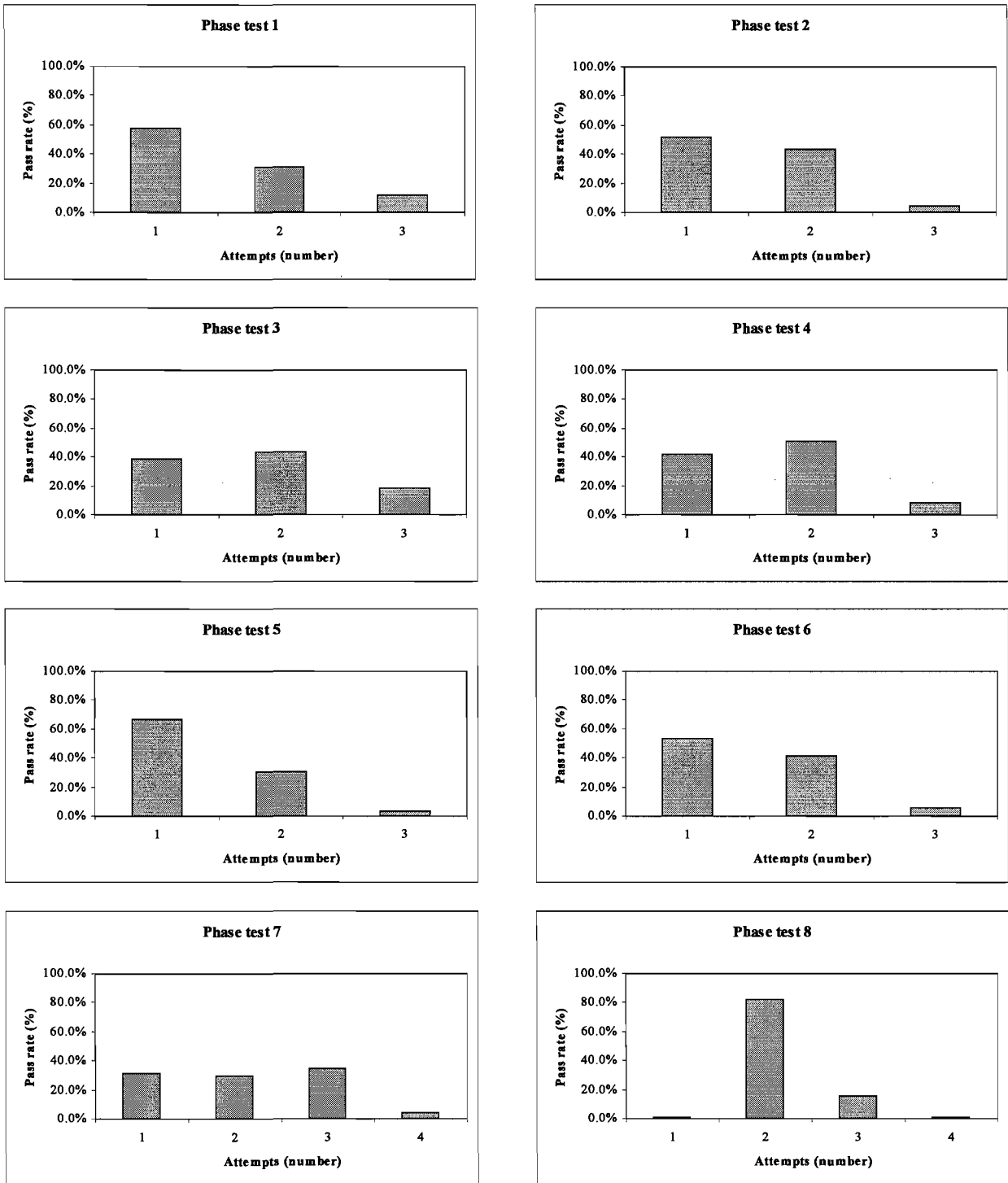


Figure 2. Number of attempts needed to successfully complete the phase tests

It is clear from the graphs in Figure 2 that most candidates needed to make one or two attempts to pass the tests, with a small portion needing three attempts. The exception is Phase Test 7, where about a third of the candidates only passed on the third attempt. Each candidate is granted three attempts to pass, but, as is indicated in the graphs, exceptions were made for Phase Tests 7 and 8. This was explored further. N. Faist (personal communication, 13 March 2007) explained that candidates who had performed well in all their previous phase tests were granted a fourth attempt to pass Phase Tests 7 and 8. At this stage, the candidates had almost completed their theoretical course and the Technical College had already invested significant amounts of funding, time and effort in the students; thus each case was evaluated according to its own merits.

Table 7

*Matrix of correlations between scores on the predictors (the TTb and the LPCAT) and the criterion variables (mean number of attempts needed to pass the different phase tests) (N = 89)*

	N	TTB				LPCAT
		Spatial Recognition	Mechanical Comprehension	Fault Diagnosis	Numerical Reasoning	
Phase Test 1	89	-0,22*	-0,16	-0,05	-0,14	-0,10
Phase Test 2	89	-0,14	-0,12	-0,13	-0,10	-0,29*
Phase Test 3	89	-0,07	-0,16	-0,29**	-0,18	0,04
Phase Test 4	89	-0,11	-0,20	-0,05	-0,14	-0,18
Phase Test 5	89	-0,13	-0,17	-0,17	0,01	-0,21*
Phase Test 6	88	-0,15	-0,31**	-0,26*	-0,22*	-0,20*
Phase Test 7	89	-0,12	-0,26*	-0,12	-0,16	-0,15
Phase Test 8	88	-0,27**	-0,23*	-0,25*	-0,00	-0,20*

\*  $p < 0,05$

\*\*  $p < 0,01$

Table 7 depicts the correlations between the predictor and criterion variables. Negative correlation coefficients between the predictor and criterion variables were to be expected because the higher the test performance of a candidate, the lower the number of attempts the candidate was expected to need to pass the criterion measure. The highest correlation between the TTb and a criterion variable was found in the case of Phase Test 6, with a correlation of -0,31 (medium effect size). There was a negative correlation between the LPCAT and all the criterion variables except in the case of Phase Test 3. The highest correlation between the LPCAT and the criterion,  $r = -0,29$ , which indicates a correlation of a

medium effect, was found with Phase Test 2. The results of the matrix of inter-correlations between the candidates' scores on the predictors (the TTB and the LPCAT) and the criterion variables (number of attempts needed to pass the different phase tests) must be interpreted with caution because the variance of the criterion variant was limited. For most of the phase tests it was one or two attempts; and for a small number of candidates it was three attempts. The criterion variant is categorically ordinal and therefore t-tests and one-way ANOVAs may be more suitable statistical methods to be used in the analysis of these data (see Table 9). The researcher used the different methods chosen (the correlation, t-tests and one-way ANOVAs) to view the data from different angles. It must also be noted that the observed validity coefficients may underestimate the predictor-criterion relationship due to the effect of range restriction. When an adjustment was made for this, the results relating to the correlations between the scores on the predictors (the TTB and the LPCAT) and the criterion (the mean number of attempts needed to pass the different phase tests) changed to those provided in Table 8.

Table 8

*Matrix of correlations between scores on the predictors (the TTB and the LPCAT) and the criterion variables (mean number of attempts needed to pass the different phase tests) (after adjustments for restriction of range) (N = 89)*

	TTB			LPCAT	
	Spatial Recognition	Mechanical Comprehension	Fault Diagnosis	Numerical Reasoning	
N	89	89	89	89	89
Phase Test 1	-0,27	-0,17	-0,05	-0,14	-0,14
Phase Test 2	-0,18	-0,14	-0,16	-0,10	-0,40
Phase Test 3	-0,08	-0,18	-0,34	-0,18	0,05
Phase Test 4	-0,13	-0,21	-0,06	-0,15	-0,24
Phase Test 5	-0,17	-0,18	-0,20	0,02	-0,28
Phase Test 6	-0,18	-0,34	-0,30	-0,23	-0,28
Phase Test 7	-0,16	-0,29	-0,13	-0,16	-0,22
Phase Test 8	-0,33	-0,25	-0,29	-0,01	-0,28

Table 8 indicates the correlations between the predictor and criterion variables, after adjustments had been made for a restriction of range. The table indicates that, as was expected, the correlations between the predictor and criterion variables strengthened. In the

case of Phase Tests 1, 3 and 7, only one of each of the sub-tests of the TTB indicated a statistically significant correlation with academic performance. In Phase Tests 2 and 5, only the LPCAT showed a statistically significant correlation with academic performance. In the case of Phase Tests 4, 6 and 8, both the predictors. In Phase Test 4 the LPCAT and Mechanical Comprehension showed a statistically significant correlation with academic performance after adjustments had been made for restriction of range. In the case of Phase Test 6, Spatial Recognition was the only predictor that did not show a statistically significant correlation with the criterion variable, while in the case of Phase Test 8 it was Numerical Reasoning. None of the aforementioned correlations, however, was of any practical significance. In **Appendix C** more details of the content of the different phase tests are presented. The results indicated in Table 8, combined with the results indicated in Table 9, are compared with the information supplied in Appendix C. The discussion of the relationships between the test scores on the predictors, the content of the different phase tests and the groups who needed one, two or three attempts to pass the phase tests can be seen below.

An appropriate approach to measuring the relationships explored above could be to compare the mean test scores on the predictors (the TTB and the LPCAT) of the groups who needed one, two or three attempts respectively to pass the phase tests, to investigate whether there are significant differences. Where a significant difference can be found, it would suggest that there may be a relationship between the score on the predictors and the number of attempts needed to pass the respective tests. These analyses could be used to confirm the correlations found above, but treated the criterion variables as ordinal, categorical variables. The sample size of the groups were, however, of such a nature that a one-way ANOVA could not always be used. Only when sample sizes were comparable was an ANOVA used (in the case of Phase Tests 1, 3 and 7). For the rest, the number of candidates who needed three attempts was so small that a t-test was used to compare the mean scores of the remaining two groups instead. In Table 9 the normal convention for significance is not used. In this case it was used in order to indicate findings which approach significance at the 0,05% level, but where the significance value is marginally larger than 0,05. These findings, as well of the findings of the matrix of correlations between the scores on the predictors and the criterion variables, are summarised in Table 9 below.

Table 9

*Comparison of the mean test scores on the predictors (the TTb and the LPCAT) of the candidates who needed one, two or three attempt(s) to pass the different phase tests*

	TTB			LPCAT	
	Spatial Recognition	Mechanical Comprehension	Fault Diagnosis	Numerical Reasoning	
<b>Phase Test 1</b>	ANOVA n.s.*	ANOVA n.s.	ANOVA n.s.	ANOVA n.s.	ANOVA n.s.
Mean 1 attempt	20,18	15,96	20,59	12,88	57,88
Mean 2 attempts	18,82	14,89	21,21	12,86	57,96
Mean 3 attempts	17,50	13,80	18,70	11,10	56,00
Correlation	-0,27++	-0,17+	-0,05	-0,14+	-0,14+
Effect size 1 vs 2	0,32*	0,32*	0,25*	-0,10	0,01
Effect size 1 vs 3	0,65**	0,45**	0,31*	0,58**	0,44**
Effect size 2 vs 3	0,33*	0,22*	0,38**	0,57**	0,45**
<b>Phase Test 2</b>	T-test sign. (p<0,05)	T-test sign. (p<0,05)	T test n.s.	T-test n.s.	T-test n.s.
Mean 1 attempt	20,33	16,22	21,33	13,00	58,48
Mean 2 attempts	18,21	14,18	19,95	12,31	57,54
Correlation	-0,18+	-0,14+	-0,16+	-0,10+	-0,40++
Effect size of difference	0,52**	0,45**	0,22*	0,23*	0,24*
<b>Phase Test 3</b>	ANOVA n.s.	ANOVA n.s.	ANOVA sign (p<0.05)	ANOVA n.s.	ANOVA n.s.
Mean 1 attempt	19,24	16,03	22,26	13,18	57,29
Mean 2 attempts	20,33	15,51	20,67	12,72	58,10
Mean 3 attempts	17,75	13,69	16,75	11,50	57,6
Correlation	-0,08	-0,18+	-0,34++	-0,18+	0,05
Effect size 1 vs 2	0,27*	0,11	0,25*	0,15	-0,19
Effect size 1 vs 3	0,35**	0,60**	0,91***	0,60**	-0,06
Effect size 2 vs 3	0,60**	0,36**	0,72***	0,38**	0,13
<b>Phase Test 4</b>	T-test n.s.	T-test sign. (p<0,05)	T-test n.s.	T-test sign (p<0.10)	T-test n.s.
Mean 1 attempt	19,89	16,81	20,49	13,38	58,65
Mean 2 attempts	19,29	14,18	21,16	12,07	57,07
Correlation	-0,13+	-0,21++	-0,06	-0,15+	-0,24++
Effect size of difference	0,14	0,57**	-0,11	0,43**	0,37**
<b>Phase Test 5</b>	T-test n.s.	T-test n.s.	T-test n.s.	T-test n.s.	T-test sign. (p<0,05)
Mean 1 attempt	19,80	15,93	21,29	12,66	58,44
Mean 2 attempts	18,96	14,41	19,44	12,63	56,07
Correlation	-0,17+	-0,18+	-0,20+	0,02	-0,28++
Effect size of difference	0,20*	0,33*	0,30*	0,01	0,57**

<b>Phase Test 6</b>	T-test sign (p<0.10)	T-test sign. (p<0,05)	T-test sign. (p<0,05)	T-test sign (p<0.10)	T-test sign (p<0.10)
Mean 1 attempt	20,09	16,81	21,96	13,30	58,49
Mean 2 attempts	18,50	13,86	19,00	12,14	56,89
Correlation	-0,18+	-0,34++	-0,30++	-0,23++	-0,28++
Effect size of difference	0,38**	0,67***	0,48**	0,40**	0,39**
<b>Phase Test 7</b>	ANOVA n.s.	ANOVA sign. (p<0.05)	ANOVA n.s.	ANOVA n.s.	ANOVA n.s.
Mean 1 attempt	20,07	17,25	22,25	13,43	58,86
Mean 2 attempts	20,00	15,08	19,58	12,50	56,62
Mean 3 attempts	18,23	14,06	19,68	12,19	58,13
Correlation	-0,16+	-0,29++	-0,13+	-0,16+	-0,22++
Effect size 1 vs 2	0,02	0,45**	0,40**	0,32*	0,51**
Effect size 1vs 3	0,41**	0,71***	0,41**	0,42**	0,30*
Effect size 2 vs 3	0,47**	0,24*	-0,02	0,10	-0,37**
<b>Phase Test 8</b>	T-test sign p<0.10	T-test sign. (p<0,05)	T-test sign. (p<0,05)	T-test n.s.	T-test sign (p<0.10)
Mean 1 attempt	19,93	15,96	21,27	12,78	58,03
Mean 2 attempts	18,00	13,43	17,29	12,86	55,86
Correlation	-0,33++	-0,25++	-0,29++	-0,01	-0,28++
Effect size of difference	0,48**	0,57**	0,65**	-0,03	0,52**

n.s. : not significant

+ r ~ 0,10 (small effect)

\* d ~ 0,20 (small effect)

++ r ~ 0,30 (medium effect)

\*\* d ~ 0,50 (medium effect)

+++ r ~ 0,50 (large effect)

\*\*\* d ~ 0,80 +(large effect)

It is important to note that the statistical methods used in the data analysis sometimes did not reveal statistical significance, but inspection of the mean scores on the predictor tests indicated a clear inclination downwards as the number of attempts increased. Thus, the more attempts the candidate needed to make to pass a phase test, the lower he or she scored on the predictor tests. This is evident in all the phase tests, with the exception of Phase Test 7, when the number of attempts needed to pass was compared with the candidates' performance on the LPCAT.

In Phase Test 1, students have to convert units and prefixes and explain, read and use different electrical concepts and instruments. They must also interpret technical drawings. The one-way ANOVA indicated that, as far as the performance on Phase Test 1 is concerned, not one of the predictor tests shows a statistically significant relationship with the criterion. The matrix of inter-correlations between the scores on the predictors and the criterion

variables indicates a statistically significant correlation of medium effect size between Spatial Recognition and the number of attempts the candidates needed to pass Phase Test 1. Spatial Recognition measures the candidates' ability to recognise and match shapes in various orientations. When the content of Spatial Recognition is compared to the content of Phase Test 1, the correlation can possibly be explained. Mechanical Comprehension, Numerical Reasoning and the LPCAT indicated a relationship of small effect size with the criterion. The difference in the mean scores between all the predictor tests and the criterion was of a medium effect size in the groups who needed one and three attempts to pass the phase test, although it seemed to be the strongest between Spatial Recognition and the criterion. The significant correlation is supported by the medium to large effect sizes of the differences found especially between the groups that needed one and three attempts to pass Phase Test 1.

Phase Test 2 is a practical test where the student has to test, fit, wire and connect different electrical instruments. The t-test showed that Spatial Recognition and Mechanical Comprehension seemed to indicate a relationship with the number of attempts the candidates needed to make to pass Phase Test 2. The content of the phase test can be linked with the candidates' ability to recognise and match shapes in various orientations. Mechanical Comprehension measures the candidates' ability to visualise movement and their understanding of basic mechanical principles and their applications to a number of devices – a core skill relevant in many technical jobs. This ability can also be linked with the practical application of the contents of Phase Test 2. The matrix of inter-correlations between the scores on the predictors and the criterion variables, as far as the candidates' performance in Phase Test 2 is concerned, indicated a statistically significant relationship of a medium effect size with the LPCAT. It indicated a relationship of small effect size with all the TTB subtests.

This finding contradicts the findings of the t-test, where the difference in the means scores between Spatial Recognition and Mechanical Comprehension and the criterion was of medium effect size. The effect size of the difference in the mean scores of Fault Diagnosis, Numerical Reasoning, the LPCAT and the criterion was small. The aforementioned results may indicate that Spatial Recognition and Mechanical Reasoning do have a statistically significant relationship with the number of attempts a candidate makes to pass this phase test.

In Phase Test 3, the student has to wire circuits and a semi-automatic start-delta starter. He or she has to explain star and delta connections, as well as induction motors contactors. The

contents of Phase Test 3 require students to be able to isolate faults and understand the way electrical equipment works. The ANOVA and the matrix of inter-correlation between the scores on the predictor tests and the criterion variable (performance on Phase Test 3) indicated a statistically significant relationship of medium effect size between Fault Diagnosis and the criterion. Mechanical Comprehension and Numerical Reasoning indicated a relationship of small effect size with the number of attempts needed to pass Phase Test 3. The difference in the mean scores between Fault Diagnosis and the criterion was of a large effect size in the groups who needed one and three, and two and three attempts respectively to pass the phase test. The difference in the mean scores between Spatial Recognition, Mechanical Comprehension, Numerical Reasoning and the criterion was of medium effect size in the groups who needed one and three and two and three attempts to pass the phase test. The difference in the mean scores between the LPCAT and the criterion did not indicate an effect on the number of attempts needed to pass Phase Test 3. This indicates that Fault Diagnosis does have the strongest relationship with the number of attempts a candidate needs to pass this phase test.

In Phase Test 4, students have to test, wire, explain and connect different electrical instruments. Students are expected to do simple calculations and word sums. The t-test showed a statistically significant relationship between Mechanical Comprehension and Numerical Reasoning and the criterion (number of attempts the candidates needs to pass Phase Test 4). The content of the phase tests may therefore be linked with the specific abilities. The matrix of inter-correlations indicated a relationship of small effect size between Spatial Recognition, Numerical Reasoning and the criterion. The matrix of inter-correlations indicated a relationship of medium effect size between Mechanical Comprehension, the LPCAT and the number of attempts needed to pass Phase Test 4. The correlation with the LPCAT contradicted with the findings of the t-tests. The effect size of the difference in the mean scores with regard to Mechanical Comprehension and Numerical Reasoning, the LPCAT and the criterion was medium. The effect size of the difference in the mean scores between Spatial Recognition and Fault Diagnosis indicated no relationship with the number of attempts needed to pass Phase Test 4.

Phase Test 5 is a practical test where the students have to wire, examine, connect, mount, remove and overhaul electrical equipment. Students are expected to explain transformers and single phase motors. The t-test and the matrix of inter-correlations indicated a statistically

significant relationship of medium effect size between the LPCAT and the criterion (performance on Phase Test 5). A relationship of a small effect size was indicated between Spatial Recognition, Mechanical Comprehension, Fault Diagnosis and the number of attempts needed to pass Phase Test 5. The difference between the candidates who needed one, two or three attempt(s), in the LPCAT mean scores created by the criterion was of medium effect size. A difference of a small effect size between the groups was found between the means scores of Spatial Recognition, Mechanical Comprehension and Fault diagnosis and the number of attempts the candidates needed to make to pass Phase Test 5. The effect size between the difference in the mean scores of Numerical Reasoning and the number of attempts needed to pass this phase test indicated no relationship of practical significance.

Phase Test 6 is a practical test where the students have to wire, connect, test and maintain electrical equipment. The students have to explain the theory of direct current machines. The t-test indicates all the predictor tests as having statistically significant relationships with the criterion (performance on Phase Test 6). The matrix of inter-correlations showed a statistically significant relationship of small effect size between Spatial Recognition and the criterion. The rest of the predictor tests indicated a relationship of medium effect size with the criterion. The difference in the mean scores of all the predictor tests and the number of attempts needed to pass Phase Test 6 indicated a relationship of a medium to large effect size.

Phase Tests 7 and 8 consist of the electronic training modules. The students must know, understand and apply electrical principles. The ANOVA indicated a statistically significant relationship between Mechanical Comprehension and the number of attempts needed to pass Phase Test 7. The matrix of inter-correlations demonstrated a relationship of medium effect size between Mechanical Comprehension, the LPCAT and the criterion. Mechanical Comprehension measures the understanding of basic mechanical principles and their applications to a number of devices – this is a core skill relevant in many technical jobs. A correlation of small effect size was found between the rest of the predictor tests and the criterion. The difference in the mean scores between Mechanical Comprehension and the candidates who needed one and three attempts to pass the Phase Test was of a large effect size. The effect size of the difference in the mean scores between candidates who needed one or three attempts to pass Phase Test 7 for the rest of the subtests of the TTB was medium.

Phase Test 8 requires an integration and application of all the electronic knowledge in all the training modules. The t-test and the matrix of inter-correlations indicated a statistically significant relationship between the criterion (performance on Phase Test 8) and all the predictor tests, except for Numerical Reasoning. The relationship between Spatial Recognition, Mechanical Comprehension, Fault Diagnosis, the LPCAT and the criterion was of a medium effect size. The difference in the mean scores of all the predictor tests, excluding Numerical Reasoning, and the number of attempts the candidates needed to pass Phase Test 8 was of medium effect size.

There was a difference between the results of the t-tests, the one-way ANOVA and the matrix of inter-correlations in Phase Test 1,2,3,4,5 and 7. The t-test and the ANOVA on the one hand and the matrix of inter-correlation on the other hand identified a statistically significant relationship between the same predictors and the number of attempts needed to pass a phase test for Phase Tests 6 and 8. Although the results of the different statistical methods are not always identical, it does indicate that the different predictor tests have statistically significant relationships with the criterion. The range of scores of the criterion was very small; therefore the correlation method was not as applicable as the t-test and ANOVAs. When a contradiction in the results is found one should rather consider the findings of the t-tests and ANOVAs. Sometimes only the TTB indicated a statistically significant relationship and sometimes only the LPCAT indicated a relationship, and sometimes both of the predictor tests displayed a statistically significant relationship. In the cases where significant differences were found, the group of candidates who passed the test at their first attempt always had a significantly higher score in the predictor tests than those who needed two or three attempts (except in the one case, where otherwise indicated). It can therefore be concluded that a relationship between the predictor tests and the criterion does exist, although it is not always statistically significant. The predictor tests should therefore be used in combination, because one or both of the tests indicated a relationship with the criterion in the different phase tests.

Table 10

*Differences between the means of the designated and non-designated groups regarding their scores on the predictors ((independent sample t-tests)*

		Independent samples test					
		Levene's test for equality of variances		T-test for equality of means			
		F	Sig.	t	df	Sig. (2-tailed)	Mean difference
		Lower	Upper	Lower	Upper	Lower	Upper
Spatial	Equal variances assumed	2,443	0,122	3,299	89	0,001	2,898
Recognition	Equal variances not assumed	3,176	59,63		0,002	2,898	0,912
Mechanical	Equal variances assumed	0,036	0,849	7,105	87	0	5,798
Comprehension	Equal variances not assumed			7,151	68,57	0	5,798
Fault Diagnosis	Equal variances assumed	2,600	0,11	1,895	87	0,061	2,557
	Equal variances not assumed			1,983	76,662	0,051	2,557
Numerical Reasoning	Equal variances assumed	1,379	0,243	6,042	87	0	3,407
	Equal variances not assumed	6,613	84,387	0	3,407		
LPCAT	Equal variances assumed	0,124	0,725	2,436	87	0,017	2,216
	Equal variances not assumed			2,467	69,888	0,016	2,216

An independent sample t-test was also performed in order to assess the differences between the means of the designated and those of the non-designated groups reported above. As is evident from Table 10 above, all differences were statistically significant, with the exception of Fault Diagnosis, which was not significant at the 5% level, but was significant at the 10% level. These differences were then investigated further by determining their effect sizes. The results are reported in Table 11 below.

Table 11

*Difference between the designated (N = 55) and non-designated (N = 34) groups regarding their scores on the predictors*

		Non-designated group (N = 34)				Designated group (N = 55)				
		Mean	Min	Max	SD	Mean	Min	Max	SD	d-stat
TTB	Spatial Recognition	21,15	13	30	4,36	18,40	9	27	3,80	0,65**
	Mechanical Comprehension	18,71	8	27	4,07	13,33	5	23	3,72	1,16***
	Fault Diagnosis	21,91	13	33	5,61	19,75	9	34	6,51	0,35*
	Numerical Reasoning	14,65	9	19	2,20	11,45	3	21	2,86	1,05***
LPCAT		58,97	50	69	4,02	56,91	46	66	4,25	0,48**

\* d ~ 0,20 (small effect)

\*\* d ~ 0,50 (medium effect)

\*\*\* d ~ 0,80+ (large effect)

In Table 11, the designated and non-designated groups' performance on the predictor tests is shown. From the analysis of the data the following differences became clear: in all cases, the mean of the non-designated group was higher than that of the designated group. The mean between the two groups differed by almost half a standard deviation. On the TTB, the differences between the means for the two groups on Mechanical Comprehension and Numerical Reasoning were of a large effect size. This implies that the differences are of practical significance. To put it another way, there is a 37,8% and 41,1% overlap between the means of the two groups on Mechanical Comprehension and Numerical reasoning respectively. On Spatial Recognition and the LPCAT, the difference is of a medium effect size (still of practical significance) and on Fault Diagnosis of a small effect size. There are also a 57% overlap for Spatial Recognition and the LPCAT and a 72,6% overlap for Fault Diagnosis.

Table 12

*Significance of the differences between non-designated (N = 34) and designated (N = 55) groups regarding their scores on the criterion (number of attempts needed to pass the different phase tests) (t-tests for independent samples)*

		T-test for independent samples			T-test for equality of means		
		Levene's test for equality of variances					
		F	Sig.	t	Df	Sig. (2-tailed)	Mean difference
		Lower	Upper	Lower	Upper	Lower	Upper
Phase Test 1	Equal variances assumed	0,827	0,366	-1,208	87	0,230	-0,183
	Equal variances not assumed			-1,228	70,643	0,224	-0,183
Phase Test 2	Equal variances assumed	0,283	0,596	-0,533	87	0,596	-0,069
	Equal variances not assumed			-0,521	62,836	0,604	-0,069
Phase Test 3	Equal variances assumed	1,246	0,267	-1,006	87	0,317	-0,160
	Equal variances not assumed			-1,057	77,648	0,294	-0,160
Phase Test 4	Equal variances assumed	0,912	0,342	-2,119	87	0,037	-0,283
	Equal variances not assumed			-2,105	65,840	0,039	-0,283
Phase Test 5	Equal variances assumed	0,210	0,648	-0,093	87	0,926	-0,011
	Equal variances not assumed			-0,090	59,950	0,929	-0,011
Phase Test 6	Equal variances assumed	0,129	0,720	-1,557	86	0,123	-0,206
	Equal variances not assumed			-1,548	66,361	0,126	-0,206
Phase Test 7	Equal variances assumed	3,173	0,078	0,220	87	0,827	0,044
	Equal variances not assumed			0,212	59,625	0,833	0,044
Phase Test 8	Equal variances assumed	1,139	0,289	-0,283	87	0,777	-0,027
	Equal variances not assumed			-0,303	80,640	0,763	-0,027

The results of the t-tests showed significant differences between the non-designated and the designated groups with regard to Phase Test 4, while other group differences are not statistically significant. These differences were then investigated further by determining their effect sizes. The results are reported in Table 13 below.

Table 13

*Differences between the non-designated (N = 34) and the designated (N = 55) groups regarding their scores on the dependent criterion variant (number of attempts needed to pass the different phase tests)*

	Non-designated group (N = 34)					Designated group (N = 55)					d-stat
	N	Mean	Mi n	Ma x	SD	N	Mean	Mi n	Ma x	SD	
Phase Test 1	3 4	1,47	1	3	0,71	55	1,58	1	3	0,69	0,16
Phase Test 2	3 4	1,50	1	3	0,62	55	1,55	1	3	0,57	0,08
Phase Test 3	3 4	1,71	1	3	0,63	55	1,85	1	3	0,78	0,20*
Phase Test 4	3 4	1,47	1	3	0,61	55	1,78	1	3	0,60	0,50**
Phase Test 5	3 4	1,35	1	3	0,60	55	1,38	1	3	0,53	0,05
Phase Test 6	3 4	1,38	1	3	0,60	54	1,61	1	3	0,60	0,38*
Phase Test 7	3 4	2,15	1	4	0,99	55	2,11	1	4	0,88	0,04
Phase Test 8	3 4	1,15	1	2	0,36	54	1,20	1	3	0,45	0,14

\* d ~ 0,20 (small effect)

\*\* d ~ 0,50 (medium effect)

\*\*\* d ~ 0,80 + (large effect)

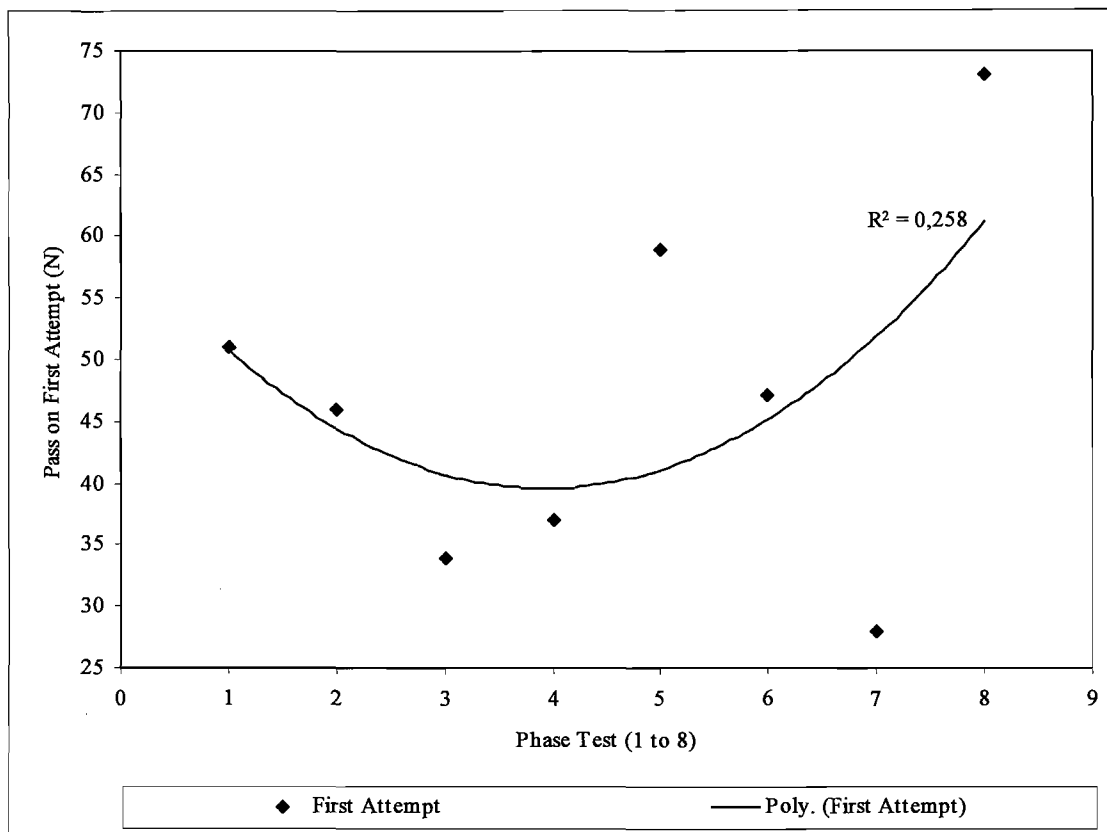
In Table 13, the non-designated and designated group's performance on the phase tests is depicted. For Phase Tests 3 and 6 the difference between the means of the non-designated group and the means of the designated group was of a small effect size, but as is indicated in

Table 12, significant difference was not confirmed by the t-test. For Phase Test 4, the difference was of a large effect size. The results of the t-tests depicted in Table 12 also revealed a statistically significant difference between the non-designated and designated groups with regard to Phase Test 4, while other group differences are not statistically significant. This strengthens the finding that the largest effect size occurred with regard to Phase Test 4. That implies that Phase Test 4 was the only phase test where the non-designated group needed a practically significantly smaller number of attempts to pass their phase test. It is also interesting to note that in the case of Phase Test 7 (the most “difficult” test) and Phase Test 8 the designated group’s mean number of attempts at passing was smaller than that of the non-designated group.

Table 14  
*Level of education of the selected candidates and their performance on the predictor variables*

		Sum of Squares	df	Mean square	F	Sig.
Spatial Recognition	Between groups	74,510	3	24,837	1,419	0,243
	Within groups	1435,024	82	17,500		
	Total	1509,535	85			
Mechanical Comprehension	Between groups	129,372	3	43,124	2,105	0,106
	Within groups	1679,930	82	20,487		
	Total	1809,302	85			
Fault Diagnosis	Between groups	127,740	3	42,580	1154	0,333
	Within groups	3026,632	82	36,910		
	Total	3154,372	85			
Numerical Reasoning	Between groups	35,453	3	11,818	1,286	0,285
	Within groups	753,477	82	9,189		
	Total	788,930	85			
LPCAT	Between groups	55,915	3	18,638	1,108	0,351
	Within groups	1379,015	82	16,817		
	Total	1434,930	85			

The educational level of the selected candidates was divided into the following groups: the group of candidates with a post-matric certificate, the group of candidates with an N4 to N6 qualification, the group of candidates with an N3 qualification and the group of candidates with an N2 qualification. The one-way ANOVAs performed with regard to both the predictor and criterion variables revealed that there were no statistically significant differences between the groups with different educational qualifications.



*Figure 3. Parabolic regression curve regarding the frequency of attempts at passing the different phase tests*

The difficulty of the contents of the different modules in the training of the first year electricians increases in complexity. Figure 3 indicates that, in the beginning, the students' pass rate on their first attempt was relatively high, where after their pass rate on their first attempts decreased. At the end of the year, although the content became more complex, the pass rate on the first attempt increased to higher levels than initially (with the exception of Phase Test 7).

To establish whether the test results predict success, discriminant analysis was used. The criterion variable was thus considered to be ordinal. Discriminant analysis was used to establish what measure of success one can have in classifying candidates as having to attempt their different phase tests once, twice or thrice on the basis of their test (predictor) results. The same principle was used as in the case of the ANOVA and t-test analyses, namely that for Phase Tests 1,3 and 7, where the number of attempts allowed for three groups that were comparable in terms of their size, three groups were used, resulting in two discriminant functions. For the remainder of the tests, only two groups were used, representing one and

two attempts respectively, resulting in one discriminant function. First, the percentage of the original cases that were classified as needing one, two or three attempts to pass the test is discussed. It is then interpreted against the discriminant analyses. The classification results for the respective phase tests are reported below.

Table 15

*Classification results of the discriminant analysis for Phase Test 1*

		Classification results <sup>(a)</sup>				
		Phase Test 1	Predicted group membership			Total
Original	Count		1	2	3	1
		1	18	16	17	51
		2	6	13	9	28
		3	2	1	7	10
	%	1	35,3	31,4	33,3	100,0
		2	21,4	46,4	32,1	100,0
		3	20,0	10,0	70,0	100,0

(a) 42,7% of original grouped cases were correctly classified

The classification results of the discriminant analysis for Phase Test 1 are reported in Table 15. Using the TTB subtests and the LPCAT as a predictor of group membership, 42,7% of the original cases were classified correctly as needing one, two or three attempts to pass the test. This is not a particularly high success rate. If one had simply guessed that all students would fall in the largest group (the group needing one attempt), one would have been correct 51 out of a possible 89 times – a success rate of 57,3%. The predictors therefore did not add much additional value as opposed to mere guessing. It would seem that greater success was achieved in predicting whether a candidate would need two or more attempts (a 46,4% success rate and a 70% success rate for two and three attempts respectively) than in predicting who would need only one attempt (a 35,3% success rate).

Table 16

*Classification results of the discriminant analysis for Phase Test 2*

		Classification results <sup>(a)</sup>			
Original	Count	Phase Test 2	Predicted group membership		Total
			1	2	
		1	27	19	46
		2	11	28	39
		Ungrouped cases	2	2	4
	%	1	58,7	41,3	100,0
		2	28,2	71,8	100,0
		Ungrouped cases	50,0	50,0	100,0

(a) 64,7% of original grouped cases were correctly classified.

Table 16 displays the classification results of the discriminant analysis for Phase Test 2. Using the TTB subtests and the LPCAT as a predictor of group membership, 64,7% of the original grouped cases were classified correctly. This is a relatively high success rate. If one had simply guessed that all students would fall into the largest group (the group needing one attempt), one would have been correct 46 out of a possible 89 times – a success rate of 51,6%. The TTB and the LPCAT therefore did add value as opposed to mere guessing more than half the time. In 71,8% of the cases, the predictors predicted which candidates would need two attempts. Greater success was thus achieved in predicting success for candidates who needed two attempts.

Table 17

*Classification results of the discriminant analysis for Phase Test 3*

		Classification results <sup>(a)</sup>				
Original	Count	Phase Test 3	Predicted group membership			Total
			1	2	3	
		1	20	5	9	34
		2	13	13	13	39
		3	1	5	10	16
	%	1	58,8	14,7	26,5	100,0
		2	33,3	33,3	33,3	100,0
		3	6,3	31,3	62,5	100,0

(a) 48,3% of the original grouped cases were correctly classified.

Table 17 depicts the classification results of the discriminant analysis for Phase Test 3. The predictor tests classified the original cases correctly 48,3% of the time. This is not a very high success rate. If one had simply guessed that all students would fall into the largest group (the group needing two attempts), one would have been correct 39 out of a possible 89 times – a success rate of 43,8%. The TTB and the LPCAT therefore did not add much additional value

as opposed to mere guessing. It would seem that the predictors were predicting more successfully whether a candidate would need one or three attempts (a 58,8% success rate and a 62,5% success rate for one and three attempts respectively) than predicting who would need two attempts (a 33,3% success rate).

Table 18

*Classification results of the discriminant analysis for Phase Test 4*

		Classification results <sup>(a)</sup>			
		Phase Test 4	Predicted group membership		Total
Original	Count		1	2	1
		1		24	13
	2		15	30	45
	Ungrouped cases		5	2	7
	%				
	1		64,9	35,1	100,0
	2		33,3	66,7	100,0
	Ungrouped cases		71,4	28,6	100,0

(a) 65,9% of the original grouped cases were correctly classified.

Table 18 demonstrates the classification results of the discriminant analysis for Phase Test 4. A total of 65,9% of the original grouped cases were correctly classified. This is a relatively high success rate. If one had simply guessed that all students would fall into the largest group (the group needing two attempts to pass successfully), one would have been correct 45 out of a possible 89 times – a success rate of 50,5%. The two predictor tests therefore added little additional value as opposed to mere guessing. Predicting whether a candidate would need one or two attempts to pass Phase Test 4 seemed more or less the same (a 64,9% success rate for one attempt and a 66,7% success rate for two attempts).

Table 19

*Classification results of the discriminant analysis for Phase Test 5*

		Classification results <sup>(a)</sup>			
		Phase Test 5	Predicted group membership		Total
Original	Count		1	2	1
		1	40	19	59
		2	11	16	27
		Ungrouped cases	1	2	3
	%	1	67,8	32,2	100,0
		2	40,7	59,3	100,0
		Ungrouped cases	33,3	66,7	100,0

(a) 65,1% of the original grouped cases were correctly classified.

Table 19 displays the classification results of the discriminant analysis for Phase Test 5. Using the TTB subtests and the LPCAT as predictors of group membership, 65,1% of the original cases were correctly classified as needing one or two attempts to pass Phase Test 5. This is a relatively high success rate. If one had simply guessed that all students would fall into the largest group (the group needing one attempt to pass the phase test), one would have been correct 59 out of a possible 89 times – a success rate of 66,2%. The two predictor tests therefore added additional value as opposed to mere guessing. In 67,8% of the cases, it predicted successfully that a candidate would need one attempt.

Table 20

*Classification results of the discriminant analysis for Phase Test 6*

		Classification results <sup>(a)</sup>			
		Phase Test 6	Predicted group membership		Total
Original	Count		1	2	1
		1	29	18	47
		2	12	24	36
		Ungrouped cases	2	4	6
	%	1	61,7	38,3	100,0
		2	33,3	66,7	100,0
		Ungrouped cases	33,3	66,7	100,0

(a) 63,9% of the original grouped cases were correctly classified.

Table 20 demonstrates the classification results of the discriminant analysis for Phase Test 6. Using the TTB subtests and the LPCAT as predictors of group membership, 63,9% of the original cases were correctly classified as needing one or two attempts to pass the test. This is a relatively high success rate. If one had simply guessed that all candidates would fall into the largest group (the group needing one attempt), one would have been correct 47 out of a possible 89 times – a success rate of 52,8%. The two predictor tests did not add much value

as opposed to mere guessing. It would seem that more or less the same percentage success was achieved in predicting whether a candidate would need one or two attempts (a 61,7% success rate and a 66,7% success rate for one and two attempts respectively).

Table 21

*Classification results of the discriminant analysis for Phase Test 7*

		Classification results <sup>(a)</sup>				
		Phase Test 7	Predicted group membership			Total
			1	2	3	1
<b>Original</b>	<b>Count</b>	<b>1</b>	14	7	7	28
		<b>2</b>	5	12	9	26
		<b>3</b>	6	7	18	31
		<b>Ungrouped cases</b>	0	3	1	4
	<b>%</b>	<b>1</b>	50,0	25,0	25,0	100,0
		<b>2</b>	19,2	46,2	34,6	100,0
		<b>3</b>	19,4	22,6	58,1	100,0
		<b>Ungrouped cases</b>	,0	75,0	25,0	100,0

(a) 51,8% of the original grouped cases were correctly classified.

The classification results of the discriminant analysis for Phase Test 7 are reported in Table 21. Using the TTB subtests and the LPCAT as predictors of group membership, 51,8% of the original grouped cases were correctly classified as needing one, two, three or more attempts to pass the phase test. This is not a high success rate. If one had simply guessed that all the students would fall into the largest group (the group needing three attempts), one would have been correct 31 out of a possible 89 times – a success rate of 34,8%. The two predictor tests therefore did not add any additional value as opposed to mere guessing. The largest success was achieved in predicting that a candidate would need three attempts (58,1%) to pass Phase Test 7.

Table 22

*Classification results of the discriminant analysis for Phase Test 8*

		Classification results <sup>(a)</sup>			
Original	Count	Phase Test 8	Predicted group membership		Total
			1	2	
		1	46	27	73
		2	5	9	14
		Ungrouped cases	1	1	2
	%	1	63,0	37,0	100,0
		2	35,7	64,3	100,0
		Ungrouped cases	50,0	50,0	100,0

(a) 63,2% of the original grouped cases were correctly classified.

Table 22 depicts the classification results of the discriminant analysis for Phase Test 8. Using the TTB subtests and the LPCAT as predictors of group membership, 63,2% of the original cases were correctly classified as needing one or two attempts to pass Phase Test 8. This is a relatively high success rate. If one had simply guessed that all students will fall into the largest group (the group needing one attempt), one would have been correct 73 out of a possible 89 times – a success rate of 82%. The two predictor tests therefore added additional value as opposed to mere guessing. More or less the same success rate was achieved in predicting whether a candidate would pass at the first or second attempt (a 63,0% success rate and a 64,3% success rate for attempts one and two respectively).

The discriminant analysis indicated that the predictor tests seemed to add value in predicting the number of attempts the candidates would need to make to pass the different phase tests. The added value may not always be significant. The results of the discriminant analysis confirmed the findings set out in Table 9, which suggested that there is a relationship between the predictor tests and the criterion, although it is not always statistically significant.

## DISCUSSION

The primary objective of this study was to determine the empirical validity of an ability test, the TTB, and a learning potential test, the LPCAT, as predictors of the academic success of first year apprentice electricians. The college under investigation provides learnerships for all the selected students in the various trades. It is therefore important to illustrate that the tests used to select these students are valid instruments for selecting students with the potential to succeed from a diverse group of candidates. It is vital for effective academic planning and it

addresses issues related to the implied cost of training the selected students. The TTB and the LPCAT were first included in the test battery at the college during 2003. N. de Jager, Head of the Technical College (personal communication, 26 April 2007), stated that the pass rate of the students had increased from an average of 60% to an average of more than 90% since the instruments had been adopted. This implicates a practically significant improvement of the pass rate of students. This study was designed to determine whether there was a correlation of statistical significance between the two predictor tests (TTB and LPCAT) and the academic success of first year apprentice electricians.

Because the administering the TTB and the LPCAT forms part of a series of steps in the selection of candidates who are allowed into the training programme (the selected candidates), it was deemed necessary to determine whether there are any significant differences between their scores on the predictor tests (the TTB and the LPCAT) and those of the non-selected candidates (those who are or were not allowed into the training programme). The t-test results of the differences between the mean scores of the selected and non-selected candidates indicated a statistically significant difference. The results of this comparison confirmed that the subtests of the TTB and the LPCAT discriminate clearly between the selected and the non-selected candidates.

The descriptive statistics showed that the selected candidates' level of performance is on a technical college or technikon level.

The researcher focused on the number of attempts the apprentice electrician students needed to pass a phase test to determine the predictive validity of the two tests. The number of attempts they needed to pass the phase tests suggested that the students found Phase Test 7 the most difficult to pass (the mean for the test was 2,12 attempts). Phase Test 8 seemed to be the easiest test to pass, with a mean of 1,18. Most of the candidates needed one or two attempts to pass the various phase tests. The exception was Phase Test 7, where about a third of the candidates needed three attempts to pass the test.

The correlation between the TTB and the LPCAT was low. This confirms that the two tests measure different aspects of intellectual functioning. De Beer (personal communication, 26 April 2007) stated that there is often a low correlation between the LPCAT and tests which largely reflect the individual's previous learning experience and not his or her potential for

future growth. The LPCAT focuses on non-verbal reasoning ability. She explained that the LPCAT measures fluid ability, while the TTB measures crystallized ability. The aim of the LPCAT is to identify students who will benefit most from training, rather than those who are already performing at a specific level (Van Eeden et al.,2001). The fact that the tests measure different constructs indicates that both tests may be used together in the selection of students. This finding is in line with the arguments of other researchers who have suggested that cognitive ability tests must be combined with other predictors in order to increase overall validity (Outz, 2002; Shochet, 1994; Taylor, 1994).

There was a negative correlation between the scores on the predictor tests and all but one of the criterion scores – the exception was the correlation between the LPCAT scores and scores on Phase Test 3. Negative correlation coefficients were expected because it was assumed that the higher the test performance of a candidate, the fewer attempts he or she would need to pass the criterion measure. After adjustments for restriction of range were made, only the TTB correlated with Phase Tests 1, 3 and 7. In Phase Tests 2 and 5 only the LPCAT showed a correlation. In Phase Tests 4, 5, 6 and 8, both the predictor tests displayed a correlation. This confirms that both tests correlate with the criterion tests, even though they measure different aspects of intellectual functioning.

It must be noted that the variance of the criterion is very limited. For most of the phase tests, the criterion was one or two attempts and for a small number of candidates, it was three attempts. The criterion variable is categorical ordinal and therefore t-tests and one-way ANOVAs may be more suitable statistical methods to use in the analysis of the above-mentioned data. The correlation matrix, t-tests and one-way ANOVAs were used to examine the data analysis from different angles. There was a difference between the results of the t-tests and the one-way ANOVAs on the one hand, and the matrix of inter-correlations on the other hand, in all the phase tests with the exception of Phase Tests 6 and 8. In the latter two cases a statistically significant relationship was identified between the same predictors and the number of attempts needed to pass a phase test. Although the results of the different statistical methods were not always the same, it does indicate that the different predictor tests have a statistically significant relationship with the criterion. Sometimes only the TTB displayed a statistically significant relationship and sometimes only the LPCAT displayed such a relationship; sometimes both the predictor tests displayed a statistically significant relationship. In the cases where significant differences were found, the group of candidates

who passed the test on their first attempt always had a significantly higher score in the predictor tests than those who needed to make two or three attempts (except in Phase Test 7 when compared with the candidates' performance on the LPCAT). It can therefore be concluded that there is indeed a relationship between the predictor tests and the criterion, although it is not always statistically significant. The predictor tests should be used in combination because one or both of the tests indicated a relationship with the criterion in the different phase tests.

From the analysis of the data it is clear that there is a difference between the means of the scores on the predictor tests for the designated and non-designated groups, with a practically significant difference of a medium effect size ( $d \sim 0,50$ ) in the case of Spatial Recognition and of a large effect size in the case of Mechanical Comprehension and Numerical Reasoning. All three the aforementioned constructs are part of the TTB. In all cases, the mean of the non-designated group was higher. Previous research has indicated that competence in the numerical and verbal domains is, to a very large extent, a function of quality and quantity of schooling (Taylor, 1994). Local and international research has found that people from disadvantaged backgrounds tend to score poorly when assessed by means of standard psychometric instruments (Ployhart, Ziegert, & McFarland, 2003; Ryan, 2001; Shochet, 1994; Van Eeden et al., 2005). Some research findings have indicated that students from disadvantaged backgrounds underperform on cognitive ability tests, compared to other students (Ryan, 2001). There is also a strong correlation between cognitive ability and positive performance (Carretta & Ree, 2000; Hunter, 1986; Schmidt & Hunter, 1998). The scores of the designated group may possibly reflect a lack of educational opportunities and not necessarily a lack of potential (De Beer, 2005; Visser & Hanslo, 2005).

As mentioned above, the non-designated group in this study performed better than the designated group on both the predictor tests used in the selection battery, and they needed fewer attempts to pass a phase test, with the exception of Phase Test 7, which is actually deemed to be the most "difficult" of all the phase tests. The fact that the non-designated group scored "higher" on both predictor variables thus correlates with the comparison between the means and standard deviations of the two different groups and the number of attempts they needed to pass the phase tests. All but one of the correlations between the scores on the TTB and the LPCAT and the criterion scores were negative, confirming that the

higher the scores on the predictors, the lower the number of attempts needed to complete the various phase tests (the criteria).

One way ANOVAs indicated that the level of previous education (the candidate's qualification) did not show a significant correlation with the predictor variables (the TTB and the LPCAT). In this study, the educational level of the students, which forms part of the pre-selection criteria, thus does not seem to predict the success in the training and completion of the electrical trade course. Although Visser and Hanslo (2005) claimed that the Matriculation Certificate Examination serves as the primary gatekeeper to students' access to training in higher institutions, perhaps one should combine it with a learning potential test to select students with potential.

In this study it is therefore concluded that the college under review seems to be using an accurate selection procedure which can predict the future performance of students in the electrical trade. A combination of both predictor tests (the TTB and the LPCAT) is necessary, as both tests measure different constructs of the intellect. The tests predict the number of attempts students will need to pass the phase tests and therefore they correlate with the success of the training. The outcome of this study may make a contribution to help the government, other training institutions and other organisations to predict the success of apprentice electricians. By reducing the risks attached to selecting candidates to study at a given institution by selecting candidates with the largest chance to succeed in their training South Africa may be able to reach its annual target of 12 500 trained artisans per year by increasing the throughput of the colleges where these candidates receive their training.

## LIST OF REFERENCES

### REFERENCES

- Anon. (2005/2006). Services Sector Education and Training Authority (SETA): Producing learnerships that are fit for purpose. *South Africa Business Guidebook*, 9, 134-135.
- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice-Hall.
- Andrews, S. G. (1996). *A small-scale investigation of the group administration of Feuerstein's learning potential assessment device*. Unpublished M.Ed. thesis, University of Natal (KwaZulu-Natal), Pietermaritzburg.
- Baron, H., & Butterworth, A. (1998, June). *Good practice for fair selection*. Paper presented at the BPS Test User Conference, Thames Ditton, UK.
- Boolsen, M., & Theron, C. C. 1996. Kennisbasiswa, gebruikspatrone en gebruikshindernisse ten opsigte van die riglyne vir die ontwikkeling en validering van personeel-keuringsprosedures. *Tydskrif vir Bedryfsielkunde*, 22(2), 7-14.
- Brown, J. D. (1988). *Understanding research in second language learning: A teacher's guide to statistics and research design*. New York: Cambridge University Press.
- Brown, J. D. (2000). Statistics Corner. Questions and answers about language testing statistics: How can we calculate item statistics for weighted items? *Shiken: JALT Testing and Evaluation SIG Newsletter*, 3(2), 19-21. Retrieved March 5, 2007, from the World Wide Web: [http://www.jalt.org/test/bro\\_6.htm](http://www.jalt.org/test/bro_6.htm).
- Brummer, H. L., Badenhorst, J. A., & Neuland, E. W. (2005, December). An evaluation of the most important competitive analysis methods applied by global mining firms to determine the future intent of a competitive force. *Southern African Business Review*, 10(3), 19-29.
- Carretta, T. R., & Ree, M. J. (2000, December). General and specific cognitive and psychomotor abilities in personnel selection: The prediction of training and job performance. *International Journal of Selection and Assessment*, 8, 227-235.
- Cascio, W. F. (1997). *Applied psychology in human resource management*. Englewood Cliffs, NJ: Prentice-Hall.
- Cattell, R. B. (1971). *Abilities: their structure, growth, and action*. Boston: Houghton Mifflin.

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum & Associates.
- Cummins, J. (1984). *Bilingualism and special education: Issues in assessment and pedagogy*. Clevedon, Somerset, UK: Multilingual Matters Ltd.
- Davies, T., & Farquharson, F. (2004). Implementing learnerships: Learner recruitment and selection – lessons learnt from the KwaZulu-Natal pilot projects. *South African Journal of Higher Education*, 18(1) (sic), 336-352. Pretoria: Unisa Press.
- De Beer, M. (2000). *Learning Potential Computerised Adaptive Test (LPCAT): Technical Manual*. Pretoria: Unisa Press.
- De Beer, M. (2002, June). *Utility of Learning Potential Computerised Adaptive Test (LPCAT) scores in predicting academic performance of bridging students: A comparison with other predictors*. Paper presented at the Fifth Annual Industrial Psychology Conference, Pretoria.
- De Beer, M. (2005). Development of the Learning Potential Computerised Adaptive Test (LPCAT). *South African Journal of Psychology*, 35(4)(sic), 717-747.
- De Jaeger, S., Hattingh, S., & Huster, E. (2002). *How to implement a learnership: A guide for employers and training providers*. Commissioned by the German Technical Co-operation in partnership with MERSETA. Pretoria: GTZ.
- De Jong, A., & Visser, D. (2000). Black and White employees' fairness perceptions of personnel selection techniques. *South African Journal of Psychology*, 30(4), 17-24.
- DEPARTMENT of Labour see South Africa. *Guide for the Development of Five Year Sector Skills plan and annual updates for 2005 to 2010*. Draft Document to the Mining Qualifications Authority 2005. Retrieved April 14, 2007, from the World Wide Web <http://www.mqa.org.za/content.asp?subID=37>
- De Villiers, A. B. (1999). *Disadvantaged student's academic performance: Analysing the zone of proximal development*. Unpublished D.Phil. thesis, University of Cape Town, Cape Town.
- DesJardins, S. (2002, June). *Studying the timing of student departure from college*. Paper presented at the Annual Forum of the Association for Institutional Research (AIR), Cape Town.
- Friedenberg, L. (1995). *Psychological testing: Design, analysis, and use*. Amherst, MA: Allyn and Bacon.

- Grey, S. S. (2000). *The psychometric properties of the Snijders-Oomen individual non-verbal Intelligence Scale Revised (SON-R) for secondary school learners from culturally diverse communities*. Unpublished masters dissertation, Rand Afrikaans University, Johannesburg.
- Guilford, J. P. (1965). *Fundamental statistics in psychology and education* (4th ed.). New York: McGraw-Hill.
- Hamers, J. H. M., & Resing, W. C. M. (1993). Learning potential assessment: Introduction. In J. H. M. Hamers, K. Sijtsma, & A. J. J. M Ruijsenaars, *Learning Potential assessment: Theoretical, methodological and practical issues*. Amsterdam: Swets & Zeitlinger.
- Henry, E., Hawkey, D., & Baron, H. (2000). *Technical Test Battery Manual and User's Guide*. Thames Ditton, UK: SHL Group.
- Hugo, H. L. E., & Claassen, N. C. W. (1991). *The functioning of the GSAT Senior for students of the Department of Education and Training*. Pretoria: Human Sciences Research Council.
- Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96, 72-98.
- Huysamen, G. K. (1996). *Psychological measurement: An introduction with South African examples*. Pretoria: Van Schaik.
- Huysamen, G. K. (1998). *Psychological measurement: An introduction with South African examples* (2nd ed.). Pretoria: Van Schaik.
- Huysamen, G. K. (2002). The relevance of the new APA standards for educational and psychological testing for employment testing in South Africa. *South African Journal of Psychology*, 32(2), 26-33.
- Kannevsky, L. (1995, February). Learning potentials of gifted students. *Roeper Review*, 17, 157-164.
- Kearney, S. (2005, Sept). Dropouts: How you can arrest the slide. *Process and Control Engineering*, 58(8), 30.
- Keller, G., & Warrack, B. (2003). *Statistics for management and economics* (6th ed.). California, Brooks/Cole: Thomson Learning.
- Kerlinger, F. N. (1986). *Foundations of behavioral research* (3rd ed.). Fort Worth, TX: Harcourt College.
- Kriek, H. (1998, Jul/Aug). Fairness and equity: New legislation will help to increase the best practice of occupational testing. *Newsline*, South Africa: SHL
- Leedy, P. D., & Ormrod, J. E. (2005). *Practical research* (8th ed.). New York: Merrill Prentice Hall.

- Levy, J. S., Robert, M. D., Gillio, G., Dee, M. D., Fenner, E., & Cornman-Levy, D. (1998). Maximizing Learning Potential for Health Care Providers: Sound Educational Principles Coupled with Computer Technologies. *Journal of Gynecologic Surgery, 14*(3), 141-147.
- Lopes, A., Roodt, G., & Mauer, R. (2001). The predictive validity of the Apil-B in a financial institution. *Journal of Industrial Psychology, 27*(1), 61-69.
- López, J. C. L. (2004). *Multicriteria decision aid application to a student selection problem*. Unpublished M.Ed. thesis, University of Occidente, Sinaloa, Mexico.
- Mlambo-Ngcuka, P. (2007, March). *Progress Report on Joint Initiative On Priority Skills Acquisition (JIPSA)*. Parliamentary Monitoring Group, Pretoria: South Africa. Retrieved April, 15, 2007, from the World Wide Web: <http://www.pmg.org.za/briefings/briefings.php?id=325>
- Mouton, J., & Marais, H. C. (1992). Basic concepts: Methodology of the behavioural sciences. Pretoria: Human Sciences Research Council.
- Murphy, R., & Maree, D. J. F. (2006). A review of South African research in the field of dynamic assessment. *South African Journal of Psychology, 36*(1)(sic), 168-191.
- Nunns, C., & Ortlepp, K. (1994). Exploring predictors of academic success in Psychology 1 at Wits University as an important component of fair student selection. *South African Journal of Psychology, 24*(4)(sic), 201-207.
- Outtz, J. L. (2002). The role of cognitive ability tests in employment selection. *Human Performance, 15*, 161-171.
- Owen, K., & Taljaard, J. J. (1996). *Handbook for the use of psychological and scholastic tests of the HSRC*. Pretoria: Human Sciences Research Council.
- Pelser, M., Bergh, Z. C., & Visser, D. (2005). The concurrent validity of learning potential and of psychomotor ability measures for the selection of haul truck operators in an open-pit mine. *South African Journal of Industrial Psychology, 31*(2), 58-70.
- Ployard, R. E., Weekley, J. A., Holtz, B. C., & Kemp, C. (2003). Distributions and Mean Differences. *Personnel Psychology, 733-752*.
- Ployhart, R. E., Ziegert, J. C., & McFarland, L. A. (2003). Understanding racial differences on cognitive ability tests in selection contexts: An integration of stereotype threat and applicant reaction research. *Human Performance, 16*, 231-259.
- Robertson, I. T., Iles, P. A., Gratton, L., & Sharpley, D. (1991). The impact of personnel selection and assessment methods on candidates. *Human Relations, 4*, 963-982.
- Ryan, A. M. (2001). Explaining the Black-White test score gap: The role of test perceptions. *Human Performance, 14*, 45-75.

- Samuda, R. J., Kong, S. L., Cummins, J., Pascual-Leone, J., & Lewis, J. (1989). *Assessment and placement of minority students*. Toronto: C.J. Hogrefe.
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*, 262-274.
- Shochet, I. M. (1994). The moderator effect of cognitive modifiability on a traditional undergraduate admissions test for disadvantaged black students in South Africa. *South African Journal of Psychology*, *24*(4)(sic), 208-215.
- SHL. (1996). *Assessment materials*. Thames Ditton, UK: SHL group.
- SHL. (2000). *Technical Test Battery: Manual and User's Guide*. Thames Ditton, UK: SHL group.
- SHL. (2004). *Occupational Testing: Technical Manual*. Thames Ditton, UK: SHL group.
- SHL. (2006, May). Are tests fair? *Newsline*, Pretoria: SHL People Performance.
- Snelgar, R. J., & Potgieter, T. (2003). An exploratory study of group adverse impact in a recruitment and selection strategy. *Sajems NS*, *6*, 89-98.
- Statsoft Inc. (2007). *Statistica (data analysis software system), version 7*. Tulsa, OK: Statsoft Inc.
- Struwig, F. W., & Stead, G. B. (2001). *Planning, designing and reporting research*. Cape Town: Maskew Miller Longman.
- SOUTH AFRICA. (1996). New Constitution 108 of 1996. Pretoria: Government Press.
- SOUTH AFRICA. (1998a). Employment Equity Act 55 of 1998. Pretoria: Government Press.
- SOUTH AFRICA. (1998b). Skills Development Act 97 of 1998. Pretoria: Government Press.
- SOUTH AFRICA. (1999). Skills Development Levies Act. Pretoria: Government Press.
- Taylor, T. R. (1994, December). A review of three approaches to cognitive assessment, and a proposed integrated approach based on a unifying theoretical framework. *South African Journal of Psychology*, *24*(4)(sic), 184-193.
- Taylor, T. R. (1999). *Administrator's manual for Tram-1 Battery: Assessment of Learning Potential for Semi-literates*. Johannesburg: Arolab.
- Van Eeden, R., De Beer, M., & Coetzee, C. H. (2001). Cognitive ability, learning potential, and personality traits as predictors of academic achievement by engineering and other science and technology students. *South African Journal of Higher Education*, *15*, 171-179.
- Van der Merwe, D., & De Beer, M. (2006). Challenges of student selection: Predicting academic performance. *South African Journal of Higher Education*, *20*, 547-562.

- Van der Merwe, G. (2005). *The construct validity of the First View for selection purposes in South Africa*. Unpublished master's dissertation, University of Pretoria, Pretoria.
- Van der Merwe, G. (2005). *The construct validity of the first view for selection purposes in South Africa*. Unpublished master's dissertation, University of Pretoria, Pretoria.
- Van der Merwe, R. P. (2002). Psychometric testing and human resource management. *South African Journal of Industrial Psychology*, 28(2), 77-86.
- Visser, A. J., & Hanslo, M. (2005). Approaches to predictive studies: Possibilities and challenges. *South African Journal of Higher Education*, 19, 1160-1176.
- World Economic Forum. (2001, 2003). World competitiveness report. Available at: <https://members.weforum.org/sit/homeprivat.nsf/Content/Africa+Economic+Summit+2003%5C-Competitiveness+Profiles> Date of access: 25 Sept. 2006.

## **CHAPTER 3**

### **1. CONCLUSIONS, RECOMMENDATIONS AND LIMITATIONS**

The purpose of this chapter is to make conclusions based on the objectives of the study. The limitations of the research are also discussed, and recommendations are made for the technical college in question and for technical colleges in general. Finally, some recommendations for future research are made.

#### **1.1 CONCLUSIONS**

The aim of this study was to determine the empirical validity of the Technical Test Battery (TTB) and the Learning Potential Computerised Adaptive Test (LPCAT) as instruments that can be used to select apprentices for training as electricians at a technical college in the Northern Cape in the Republic of South Africa.

The first objective was to investigate the current local employment issues that affect selection for training in technical fields. The literature indicated that the selection challenges in South Africa are unique and complex. The country has 11 official languages, it must address a wide range of social and educational problems, and it must adapt to a variety of political and legislative changes (Zaaiman, 1998). Previous research indicates that organisations have to face multiple challenges when they have to make selection decisions. On the one hand, they need to operate within the ambit of the law and, on the other; they must maximize the probability of selecting students with the best potential to be successful in their training (Nunns & Ortlepp, 1994; Pelser, Bergh, & Visser, 2005).

At present South Africa is experiencing a serious employment crisis (World Economic Forum, 2001/2002). A significant proportion of the shortages that are experienced are skills needed by artisans (Mlambo-Ngcuka, 2007). Approximately 5000 artisans are currently trained every year, but research suggests that at least 12 500 artisans must be produced each year over the next four years to meet the local demand in the labour market (Mlambo-Ngcuka, 2007). This implies a projected increase of 30 000 artisans between the year 2007 and the year 2010. The electrician trade is identified as one of the priority trades where the training of artisans must be increased (Mlambo-Ngcuka, 2007).

In this study, the second objective was to examine the difference between the measurement of cognitive ability and the measurement of learning potential. According to Taylor (1994), almost all the published tests that are available for use in industry and many of those available for educational use are designed to measure constructs such as abilities. However, conventional ability measurement instruments have increasingly come under fire for being biased or unfair to disadvantaged groups (Taylor, 1994). Traditional measures often used in the assessment of cognitive functions of individuals are regarded as unsuitable for two reasons. Firstly, a particular language is used in the test items of and in the instructions for most cognitive tests, and therefore the overall lower test scores of culturally diverse individuals are often indicators of poor verbal knowledge. Hugo and Claassen (1991) have established that large differences exist and that they may influence test scores when tests are administered in English to black pupils whose first language is an African language. Secondly, the verbal-item content in such cases may fail to provide important information regarding the potential learning and reasoning abilities of a person from a different culture (Grey, 2000).

Conventional psychometric tests are therefore often not very suitable for the selection of students from disadvantaged backgrounds, because they are a static measure of current abilities which gives no indication of the students' potential to learn (Van der Merwe & De Beer, 2006). Poor scores often reflect a lack of educational opportunities rather than a lack of potential (De Beer, 2005; Visser & Hanslo, 2005). Tests of cognitive abilities generally measure the individual's current level of functioning and rely heavily on the assumption that all the examinees have had comparable opportunities to acquire the skills and abilities that are being measured (De Beer, 2005; Murphy & Maree, 2006; Shochet, 1994). It is important to note that cognitive ability tests produce racial differences that are three to five times larger than that of other predictors (Hunter & Hunter, 1984; Outtz, 2002). Notwithstanding that fact, an overview of the literature indicated that cognitive ability tests are among the most predictive and practically efficient predictors of job performance for most occupations (Hunter & Hunter, 1984; Schmidt & Hunter, 1998). Some researchers suggest that cognitive ability tests should be combined with other predictors, such as learning potential tests, in order to reduce any adverse impact, while increasing the overall validity (Outtz, 2002; Shochet, 1994; Taylor, 1994).

The third objective of this study was to understand the nature of the construct of cognitive ability and that of learning potential. Previous research by Murphy and Maree (2006) indicates that the terms represent different aspects of intellectual functioning and measurement. It is now known that intelligence is a complex and dynamic phenomenon (Levy, et al., 1998). A study of intelligence conducted at Harvard University identified “multiple intelligences” (Levy et al., 1998). Different psychologists produced different structures of the intellect (Taylor, 1994). In his article, Taylor (1994) mentioned that Cattell (1971) had developed a higher-order theory that distinguished two forms of intelligence, namely, fluid intelligence and crystallised intelligence. Crystallised intelligence is a basic inherited capacity as developed by an interaction with environmental characteristics which are found in any society. Ability tests would be classified as measures of crystallised intelligence; hence, they would be strongly affected by cultural influences and also by schooling (Taylor, 1994). Fluid intelligence refers to specialised skills and knowledge promoted by and required in a given culture. The TTB measures crystallised ability, while the LPCAT measures fluid ability (De Beer, personal communication, 26 April 2007).

In the empirical study reported in Chapter Two, the researcher set out to determine whether there is a relationship between the TTB and the LPCAT as two different predictors of the academic success of first year apprentice electricians. The correlation between the TTB and the LPCAT is in fact quite low. This confirms that the two tests measure different aspects of intellectual functioning. De Beer (personal communication, 26 April 2007) stated that there is often a low correlation between the LPCAT and tests which largely reflect the individual’s previous learning experience rather than the person’s future potential for growth. The LPCAT focuses on non-verbal reasoning ability. As mentioned before, De Beer (personal communication, 26 April 2007) explained that the LPCAT measures fluid ability, while the TTB measures crystallised ability. The aim of the LPCAT is to identify students who will benefit most from training, rather than those who are already performing at a specific level (Van Eeden, De Beer, & Coetzee, 2001). The fact that the tests measure different constructs indicates that both tests may be used in tandem in the selection of students.

The empirical study also set out to determine whether there is a difference in the empirical (predictive) validity of the psychometric selection method if only one of the tests is used, as opposed to if both are used in combination. The empirical study indicated a negative correlation between the scores on the predictor tests and all but one of the criterion scores.

The exception was the correlation between the candidates' LPCAT scores and their scores on Phase Test 3. Negative correlation coefficients were expected because it was assumed that the higher a candidate's test performance, the fewer attempts he or she was expected to need to make to pass the criterion measure. After adjustments had been made to restrict the range, only the TTB correlated with Phase Tests 1, 3 and 7. Only the LPCAT correlated with Phase Tests 2 and 5. Both the predictor tests correlated with Phase Tests 4, 5, 6 and 8. This confirms that both tests correlate with the criterion tests, although they measure different aspects of intellectual functioning.

The empirical study also intended to determine whether there were any differences regarding the scores on the TTB and the LPCAT of students from the designated group, as opposed to the scores of students from non-designated groups. From the analysis of the data (see Chapter Two) it is clear that there was a difference between the means of the scores on the predictor tests for the designated and non-designated groups. In all cases, the mean of the non-designated group was higher. Other studies have indicated that competence in the numerical and verbal domains is very much a function of candidates' quality and quantity of schooling (Taylor, 1994). In the current study, the differences in the effect sizes of the mean score differences were confirmed by the significant correlation between the designated and non-designated groups and the scores of the number of attempts needed to pass the various phase tests.

As has already been mentioned above, the non-designated group performed better on both the predictor tests used in the selection battery and they needed fewer attempts to pass each phase test, with the exception of Phase Test 7, which is actually deemed to be the most "difficult" of all the phase tests. The fact that the non-designated group scored "higher" on both predictor variables thus correlates with the comparison between the means and the standard deviations of the two groups and the number of attempts they needed to pass the phase tests. All but one of the correlations between the scores on the TTB and the LPCAT and the criterion scores were negative, which confirms that the higher the scores on the predictors, the lower the number of attempts needed to complete the various phase tests (the criterion).

To determine whether the TTB and the LPCAT are valid predictors that can be used as selection instruments to predict the success of apprentice electricians, a variety of statistical methods were used. To determine whether there are any significant differences between the

scores of the selected and non-selected candidates on the predictor tests (the TTB and the LPCAT), t-tests were done. The t-test results of the differences between the mean scores of the selected and non-selected candidates indicated a statistically significant difference. The results of the aforementioned comparison confirm that the subtests of the TTB and the LPCAT discriminate clearly between the selected and the non-selected candidates.

The researcher focused on the number of attempts the apprentice electrician students needed to pass a phase test to determine the predictive validity of the two tests. The results of the different statistical methods were not always the same. This suggests that the different predictor tests have a statistically significant relationship with the criterion. Sometimes only the TTB displayed a statistically significant relationship, and sometimes only the LPCAT displayed such a relationship; sometimes both the predictor tests displayed a statistically significant relationship with the criterion. In the cases where significant differences were found, the group of candidates who passed the test on their first attempt always had a significantly higher score in the predictor tests than those who needed two or three attempts (except in the case of Phase Test 7 when compared with the candidates' performance on the LPCAT)

The conclusion of this study is that the technical college investigated seems to have chosen an accurate selection procedure which can predict the future performance of students in the electrical trade. A combination of both predictor tests (the TTB and the LPCAT) is necessary, as each of these tests measures a different construct of the intellect. The tests do predict the number of attempts students will need to pass the phase tests and therefore the test results do correlate with the success of the training.

The investigated college wants the selection methods it uses to be fair and effective in allowing opportunities for access and success in the learnerships the college provides. The results of this study show that the candidates' test performance on the TTB and the LPCAT correlated significantly with the academic success of the first year apprentice electricians. The results indicated that high scores on the predictors (the TTB and the LPCAT) were related to a reduction in the number of attempts needed to pass the phase tests (the criterion).

Skilled, qualified and experienced electricians are greatly in demand. To meet new and existing challenges in organisations and to comply with legislation, organisations and training

institutions must face huge challenges. It is evident that when individuals with high potential are identified accurately, time, effort and resources can be conserved. Organisations will have a much higher return on their investment if they can predict which candidates will be successful in their training. The evidence from this study suggests that the predictor tests (the TTB and the LPCAT) used in the selection battery of the investigated college can make a positive contribution in predicting the success of first year apprentice electrician students. The TTB and the LPCAT as selection instruments proved to be worthy instruments to be used to solve the problem of selecting candidates who will be successful in their training.

## **1.2 LIMITATIONS**

Several limitations can be reported regarding this study.

First, the sample size was relatively small. However, this study has made a start in the quest for accurate predictors of the academic success of first year apprentice electrician students.

The variance of the criterion was limited, and this complicated the statistical analysis. The criterion variable was categorically ordinal. For most of the phase tests, it consisted of one or two attempts; and for a small number of candidates it consisted of three attempts. Different statistical methods therefore had to be used to view the data analysis from different angles.

The study was limited to apprentice electrician students and the results may not apply to the other four trades (fitter and turner, boilermaker, diesel mechanic and millwright). Although the findings could, to a certain extent, be generalised to the other trades, which are also of a technical nature, more evidence of validity is required for those trades.

The focus of the study was only the TTB and the LPCAT as predictors of academic success. In future research, the whole selection battery, which includes the pre-selection methods, as well as the interview results, should be investigated.

## **1.3 RECOMMENDATIONS**

### ***1.3.1 Recommendations to the technical college in question, as well as to technical colleges in general***

- The academic success of students may be influenced by multiple factors (for example, students' attitudes, personalities, cultural background and interests). The current selection battery of the college should improve if these factors could be measured.
- An industrial psychologist or human resources practitioner should be present during the interviews. Such specialists in the behavioural sciences can assist the bursary administrator and senior training officer in selecting students.
- Considering the current skills shortage, more emphasis should be placed on the training of tradesmen rather than on all the racial and legal issues but is in legally possible?
- The difficulty of the language in a theoretical test may affect candidates' success. The technical college should enhance students' language proficiency by providing suitable bridging courses to students whose first language is not English.
- Considering the cost of training candidates and the current skills shortage, more colleges should incorporate the TTB and the LPCAT as part of their selection battery for apprentice electrician students.

### ***1.3.2 Recommendations for future research***

- Future research needs to examine how the academic success of students may be influenced by factors such as their attitudes, personalities, cultural background and interests. The focus of such studies should be on how to measure these factors in a culture fair way.
- Further research with an expanded sample of candidates could be conducted to determine the predictive validity of the selection battery for the four other trades (fitter and turner, boilermaker, diesel mechanic and millwright).
- Research could be undertaken to examine whether interventions such as bridging classes really improve the academic performance of students.
- Research could be undertaken to evaluate how appropriate and effective the use of psychological assessment instruments really is when they are used cross-culturally.

#### 1.4 LIST OF REFERENCES

- Cattell, R.B. (1971). *Abilities: their structure, growth, and action*. Boston, MA: Houghton Mifflin.
- De Beer, M. (2005). Development of the Learning Potential Computerised Adaptive Test (LPCAT). *South African Journal of Psychology*, 35, 717-747.
- Grey, S. S. (2000). *The psychometric properties of the Snijders-Oomen individual non-verbal Intelligence Scale Revised (SON-R) for secondary school learners from culturally diverse communities*. Unpublished MA Dissertation, Rand Afrikaans University, Johannesburg.
- Levy, J. S., Robert, M. D., Gillio, G., Dee, M. D., Fenner, E., & Cornman-Levy, D. (1998). Maximizing learning potential for health care providers: Sound educational principles coupled with computer technologies. *Journal of Gynaecologic Surgery*, 14, 141-147.
- Hugo, H. L. E., & Claassen, N. C. W. (1991). *The functioning of the GSAT Senior for students of the Department of Education and Training*. Pretoria: Human Sciences Research Council.
- Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96, 72-98.
- Mlambo-Ngcuka, P. (2007, March) Progress report on Joint Initiative on Priority Skills Acquisition (JIPSA). *Government Communications (GCIS)* <http://www.pmg.org.za/briefings/briefings.php?id=325> Date of access: 18 April 2007.
- Murphy, R., & Maree, D. J. F. (2006). A review of South African research in the field of dynamic assessment. *South African Journal of Psychology*, 36, 168-191.
- Nunns, C., & Ortlepp, K. (1994). Exploring predictors of academic success in Psychology 1 at Wits University as an important component of fair student selection. *South African Journal of Psychology*, 24, 201-207.
- Outtz, J. L. (2002). The role of cognitive ability tests in employment selection. *Human Performance*, 15, 161-171.
- Pelser, M., Bergh, Z. C., & Visser, D. (2005). The concurrent validity of learning potential and of psychomotor ability measures for the selection of haul truck operators in an open-pit mine. *South African Journal of Industrial Psychology*, 31, 58-70.
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274.

- Shochet, I. M. (1994). The moderator effect of cognitive modifiability on a traditional undergraduate admissions test for disadvantaged black students in South Africa. *South African Journal of Psychology*, 24, 208-215.
- Taylor, T. R. (1994, December). A review of three approaches to cognitive assessment, and a proposed integrated approach based on a unifying theoretical framework. *South African Journal of Psychology*, 24, 184-193.
- Van der Merwe, D., & de Beer, M. (2006). Challenges of student selection: Predicting academic performance. *South African Journal of Higher Education*, 20(4), 547-562.
- Van Eeden, R., De Beer, M., & Coetzee, C. H. (2001). Cognitive ability, learning potential, and personality traits as predictors of academic achievement by engineering and other science and technology students. *South African Journal of Higher Education*, 15, 171-179.
- Visser, A. J., & Hanslo, M. (2005). Approaches to predictive studies: Possibilities and challenges. *South African Journal of Higher Education*, 19, 1160-1176.
- World Economic Forum. (2001/2002). *World competitiveness report*. Available at: <https://members.weforum.org/site/homeprivate.nsf/Content/Africa+Economic+Summit+2003%5C-Coompetitiveness+Profiles> Date of access: 25 Sept. 2006.
- Zaaiman, H. (1998). *Selecting students for mathematics and science. The challenge facing high education in South Africa*. (Abstract) Pretoria: HSRC Press. Retrieved September 21, 2006, from the World Wide Web: [http://www.hsrcpublishers.co.za/full\\_title\\_info.asp?id=2037](http://www.hsrcpublishers.co.za/full_title_info.asp?id=2037)

**APPENDIX A**  
(SHL, 2000)

**1. Reliability of the TTB**

Reliability can be defined as the degree to which scores for a group of assesseees are consistent over one or more potential sources of error (for example time, raters, items and conditions of measurement) in the application of a measurement procedure (Society for Industrial & Organisational Psychology of South Africa, 2005). To maximize reliability, all TTB tests have gone through the formal process of test construction, namely: item generation, initial trailing, item analysis and final trailing, to ensure that all the items are of an appropriate level and all correlate with the rest of the items in a test (SHL, 2000).

Reliability is normally expressed in terms of a correlation coefficient, which is called a reliability coefficient. Internal consistency, test retest reliability and alternate form reliability are three of the main methods of estimating reliability (SHL, 2000). Internal consistency reliability estimates for all the TTB test which are being used by this specific training institution are reported in Table 1. Coefficients have been calculated using Cronbach's Coefficient Alpha.

Table 1

***Internal consistency reliability estimates for the TTB test***

Test	N	X	sd	r	SEm (Rs)	SEm (Ts)
Numerical reasoning (NT6.1)	374	11,58	4,09	0,80	1,83	4,47
Spatial Recognition (ST8.1 and ST8.2)	454	21,55	5,38	0,81	2,35	4,36
Mechanical Comprehension (MT4.1 and MT4.2)	455	18,58	6,18	0,83	2,55	4,12
Fault Diagnosis (FT7.1)	242	21,24	8,05	0,92	2,28	2,83

Standard Errors of Measurement (SEm), expressed in raw score units and T-scores, are also given in Table 1. For the majority of tests, the standard error of measurement is between 1,8 and 2,5 for raw scores and between 3,0 and 5,0 for T-scores. Two standard errors of measurement should be allowed when considering the abilities of two candidates on the same test before inferring that there is a meaningful difference between them (SHL, 2000).

## **2. *Validity of the TTB***

The validity of a test is concerned with the extent to which it measures what it has been designed to measure (SHL, 2000). Validity can be defined as the degree to which accumulated evidence and theory support specific interpretations of scores from a selection procedure entailed by the proposed uses of that selection procedure (Society for Industrial & Organisational Psychology of South Africa, 2005).

The validity of a test can take a number of different forms, depending on the specific question being asked (SHL, 2000). All types of validity are relevant to an occupational test, but SHL (2000) sees content, concurrent and predictive approaches as the most important.

### **2.1 Content validity of the TTB tests**

Content validity is concerned with whether the test items are representative of the domain to be measured (SHL, 2000). The content validity of the test is satisfactory to the extent that experts agree

- i) whether the chosen items adequately represent the tasks in the universe as defined by the test constructor; and
- ii) whether the written items indeed require the execution of these tasks (Huysamen, 1998).

The TTB tests have been designed to measure core abilities required in a wide range of technical occupations. Job analysis of a variety of such posts has shown that verbal, numerical and visual skills are often, although not always, key attributes associated with successful job performance. The content for all of the TTB tests has been taken from materials commonly found in the technical workplace (SHL, 2000).

## **2.2 Criterion related validity of the TTB tests**

Huysamen (1998) defined criterion-related validity of a test as *the extent to which its scores accurately predict scores on the relevant criterion*. It can be noted that a number of validation studies have been carried out using the TTB tests. The results of these studies are reported in 2.4.1.

## **2.3 Construct validity of the TTB tests**

Construct validity is concerned with the extent to which a test indeed measures the hypothetical construct it purports to measure (Huysamen, 1998). SHL conducted a number of studies, which make it possible to examine the extent to which the original and new TTB tests are related to each other and to other tests (SHL, 2000). A sample of this information is reported in **Appendix B**.

## **2.4 Criterion related validity of the TTB tests (Henry et al, 2000).**

### *2.4.1 Study 1*

*Craft apprentices for a government agency.*

This long-term study was designed to assess the effectiveness of a number of (specially designed) TTB tests in the selection of apprentices by examining whether significant relationships existed between test scores and training progress. The apprentice intake of 1981 to a range of government establishments made up the main core of the sample and this group was followed up to the completion of their apprenticeships in 1985. Data was also collected in the 1982 and 1983 apprentice intakes and added to the overall results.

The tests used were equivalent versions of BT1 (Verbal Comprehension), NT2 (Numerical Computation) and MT4 (Mechanical Comprehension).

A predictive approach was used, with the initial test scores, collected at the time of selection, being correlated with training performance obtained at regular intervals (for example, 12 months, 24 months, 36 months etc.) throughout the apprenticeship. There was some restriction of range in the sample as the test results, obtained at selection, had been used to pre-screen applicants.

Data was collected for a total of 590 apprentices drawn from the 1981-1983 intake. There was a wide range of trades represented by this sample and they were spread across 38 different establishments. Most of the sample were male and were between 16 and 18 years old. Nearly all of the sample had some educational qualifications, these being a mixture of GCE Ó"levels and CSEs. The average number of GCE Ó level exams passed was 3.4 and for CSE's it was 4.3.

Two sets of criteria were used, the first being rating of Workshop Performance and the second academic performance in (First Year) Technical College exams. The Technical College exams were grouped under four main headings: Maths, Main Technical Subject, Engineering Science and Engineering Drawing. Workshop Performance was assessed against 16 skill-based criteria plus an overall assessment of craft apprentice performance. Each of the 16 skill-based criteria was assessed on a continuous scale and these were then converted into ten point ratings for the purpose of analysis. These ratings showed some degree of central tendency as well as a strong degree of intercorrelation. Apprentice performance on GCEÓ level exams was also collected. Because these results were not available at the time of selection, they did not influence final selection decisions. The following significant correlations were obtained between the criteria and the test scores.

## Study 1 Results

Table A1

*First Year Workshop Performance Ratings (N=397-444)*

	Eq VT1	Eq NT2	Eq MT4
Ability to use hand and powered tools	-	-	0,16*
Knowledge of appropriate equipment	-	0,15*	0,09
Ability to produce two dimensional drawings	-	0,20*	-
Ability in work requiring numeracy	0,13*	0,26*	-
Ability to work without supervision	-	-	0,11
Absorption of new techniques	-	0,1	0,15*
Ability to diagnose problems	-	-	0,11
Ability to produce accurate work	-	-	0,17*
Ability to plan and organise	-	-	0,11
Understanding of technical work	0,09	0,20*	-
Suitability as a craft apprentice	-	0,17*	0,09
Overall performance	-	-	0,20*

Table A2

*First year technical college results (N=52-255)*

	Eq VT1	Eq NT2	Eq MT4
Maths (n=251)	-	0,26*	-
Main technical subject (n=255)	-	-	0,21*
Engineering science (n=172)	-	-	0,25*
Engineering drawing (n=144)	-	-	0,37*
City and Guilds - General (n=54)	-	-	0,37*

City and Guilds - Mechanical (n=52)	0,28	-	0,42*
-------------------------------------	------	---	-------

Table 3

*GCE 'O' level educational results (N=500)*

	Eq VT1	Eq NT2	Eq MT4
GCE 'O' level Maths grade	0,31*	0,48*	0,14*
GCE 'O' level English grade	0,33*	0,26*	-
GCE 'O' level Physics grade	0,14*	0,13*	0,22*
No of GCE 'O' levels	0,36*	0,15*	0,38*

All quoted correlations are significant at 5% level, one-tailed \* significant at 1% level, one-tailed.

Workshop performance was best predicted by the (equivalent) NT2 and MT4 tests. For the Technical College Results, the (equivalent) MT4 was the best predictor. For GCE 'O' Level performance, the (equivalent) VT1 and NT2 tests were the best predictors, reflecting the academic nature of the criteria.

#### 2.4.2 Study 2

##### *Applicants for Mechanical Engineering Apprenticeships*

This study was carried out in February 1992 and was designed to evaluate the likely effectiveness of using four tests from the (revised) Technical Test Battery in the future selection of applicants for mechanical engineering apprenticeships.

A concurrent validation approach was used involving the testing of incumbent apprentices and the collection of current job performance data.

The sample size was 97, of which 95% were males. The average age was 16,73 years (sd=0,69) and 87% of the sample were educated to at least 'O' level/GCSE standard. When compared to a

norm group of engineering apprentices, the sample scored well above average on Verbal Comprehension (VT1.1) and above average on the remaining three tests. The means and standard deviations for each test are shown in Table 4.

## Results of Study 2

Table A4

*Means and Standard Deviations for TTB tests*

Test	X	sd
Verbal Comprehension (VT1.1)	28,32	4,05
Numerical Computation (NT2.1)	22,86	4,39
Visual Estimation (ET3.1)	23,34	4,85
Mechanical Comprehension (MT4.1)	21,05	5,00

One overall performance rating was used for each apprentice.

Table A5

*Correlations between the TTB tests and overall performance rating (N=97)*

	Verbal Comprehension VT1.1	Numerical Computation NT2.1	Visual Estimation ET3.1	Mechanical Comprehension MT4.1
Overall Performance Rating	-	-	0,31*	0,42*

\* significant at 1% level, one-tailed

Two TTB tests, namely Visual Estimation (ET3.1) and Mechanical Comprehension (MT4.1), correlated significantly with the overall performance rating. Verbal Comprehension (VT1.1) and

Numerical Computation (NT2.1) correlated positively with the overall performance rating but not at a statistically significant level.

### 2.4.3 Study 3

#### *Mechanical Engineering Apprentice Applicants*

This concurrent study was completed in June 1992 and involved all eight tests in the revised Technical Test Battery. The aim was to identify those tests which were likely to be effective in the future selection of mechanical engineering apprentices.

A total of 89 apprentices, all of whom has been pre-selected using other psychometric tests, were involved. All were male and white European, 37 were 1st or 2nd year apprentices and 52 were 3rd or 4th year. Of the total sample, 87% were educated to at least 'O' level/GCSE standard. The overall performance of the group was above average on all tests when compared to an engineering, apprentice norm group. The sample was well above average on Mechanical Comprehension (MT4.1), Technical Understanding (VT5.1) and Numerical Reasoning (NT6.1). The means and standard deviations for each test are shown below.

### Results of Study 3

Table A6  
*Means and Standard Deviations for TTB Tests*

Test	X	sd
Verbal Comprehension (VT1.1)	30,11	3,92
Numerical Computation (NT2.1)	23,81	5,16
Visual Estimation (ET3.1)	24,79	4,38
Mechanical Comprehension (MT4.1)	23,72	5,34
Technical Understanding (VT5.1)	30,79	4,32
Numerical Reasoning (NT6.1)	17,28	3,00
Fault Diagnosis (FT7.1)	24,44	6,90
Spatial Recognition (ST8.1)	23,26	4,55

Each apprentice was rated by two independent assessors (instructors) on eleven performance criteria deemed to be relevant to training success. The two sets of ratings were then added together to give a more balanced view and compensate, as far as possible, for any inconsistencies between assessors. Seven different assessors were involved in rating 1st and 2nd year apprentices, with the correlations between raters on the same criterion ranging from 0,65 to 0,82. Far more assessors were responsible for rating of the 3rd and 4th year apprentices (approximately 40). This resulted in much less consistency with the inter-rater correlations varying between 0,14 and 0,53.

The significant correlations obtained between tests and performance ratings are presented below.

Table A7

*Correlations between the TTB Tests and Performance Ratings (N=89)*

Performance Rating	ET3.1	MT4.1	VT5.1
Understanding oral instructions	-	-	0,29
Absorbing new technology	0,3	0,027	0,26
Overall level of job performance	-	0,25	-

All quoted correlations are significant at 5% level, one-tailed.

Despite inconsistencies in the performance data, as illustrated by the low inter-rater correlations, all eight tests correlated positively with several of the criteria. Rating inconsistencies and the restriction in range due to pre-selection both tended to depress the correlations between tests and ratings. It is reasonable to expect that, but for the effect of these two factors, many more significant correlations would have been found.

### **3. Relationship of the TTB tests with other measures.**

A number of studies have been conducted which make it possible to examine the extent to which the original and new TTB tests are related to each other and to other tests. A sample of this information is reported in Table B1 below. All correlations quoted are significant to at least the 5% level, unless otherwise stated.

Table A8

*Correlations between Original TTB Tests and the Applied Technology Series Tests (based on a combined and pre-selected group of technician apprentices, N=31-567)*

ATS Tests		VT1	NT2	ET3	MT4	VT5	NT6	ST7	DT8
Following instructions	VTS1	0,43	-	0,34*	0,34	0,43	-	-	-
Numerical Estimation	NTS2	0,39	0,62	-	-	0,29	0,42	-	-
Mechanical Comprehension	MTS3	0,46*	0,24*	-	0,48*	-	0,25*	0,40*	-
Fault Finding	FTS4	0,34*	0,36*	0,22	0,33*	-	0,41*	0,34*	0,46
Spacial Checking	STS5	-	0,33	-	-	-	-	0,48	0,43
Diagrammatic Thinking	DTS6	0,26	-	-	0,25	0,55	-	0,29	0,42

\* weighted average correlations from a number of different sample groups.

Table A9

*Correlations between Original TTB Tests and the Work Skills (Production) Series Tests (based on a sample of manufacturing staff and technicians, N=256-271)*

WSS-Production Tests	MT4	DT8
Understanding Instructions VWP1	0,32	0,46
Working with Numbers NWP2	0,33	-

Table A10

*Correlations between Original TTB Tests and the Critical Reasoning Test Battery (based on a sample of trainee supervisors within the transport sector, N=160)*

CRTB Test		MT4	VT5	NT6	ST7
Verbal Evaluation	VC1	0,53	0,64	0,52	0,49

Table A11

*Correlations between TTB Tests, AH4 and Bennett's Mechanical Comprehension Test (BMCT)(based on a group within the manufacturing industry, N=28-110)*

	NT2	MT4	VT5	NT6	ST7
AH5	-			0,66	
AH4-1	0,36		0,20	0,39	
AH-41G	0,38	0,34		0,24	0,37
AH4-2	0,21	0,34		0,24	0,37
AH4-2G					
AH4-3	0,52	0,25	0,43	0,43	0,27
AH4-3G	-				
BMCT(A)	-	0,47			0,39
BMCT(B)	-	0,47			0,38

Table A12

*Correlations between Mechanical Comprehension Test (MT4), VTS1 and MTS3 from the Applied Technology Series and Bennett's Mechanical Comprehension Test (BMCT) (based on a sample of incumbent machine operators in a textile firm, N-104-135)*

		MT4
Following Instructions	VTS1	0,49
Mechanical Comprehension	MTS3	0,53
Bennett's Mechanical Comprehension Test	BMCT	0,73

Table A13

*Correlations between all eight new TTB Tests and Birkbeck Tests 1-5 (based on a group of engineering apprentices, N-89)*

	VT1.1	NT2.1	ET3.1	MT4.1	VT5.1	NT6.1	FT7.1	ST8.1
B1 (General)	0,48	0,43	-	0,35	0,47	0,50	0,41	-
B2 (Mechanical)	0,34		-	0,68	0,33	-	0,35	-
B3 (Arithmetic)		0,54	-	-	-	0,43	-	-
B4 (Maths)	0,39	0,74	0,28	0,27	0,33	0,69	0,44	-
B5 (Spacial)	0,33	0,36	0,34	0,55	-	0,42	0,47	0,40

Table A14

*Correlations between seven new TTB Tests and their nearest equivalents from the original battery (based on part of the standardisation group, N=26-59)*

	VT1.1	NT2.1	ET3.1	MT4.1	VT5.1	NT6.1	FT7.1
VT1	0,80						
NT2	(N=59) 0,94						
ET3		(N=52) 0,76					
MT4			(N=59) 0,73				
VT5				(N=58) 0,67			
NT6					(N=26) 0,73		
DT8						(N=48) 0,61	
							(N=29)

The pattern of correlations found with other established tests ties in well with the nature of the constructs the TTB aids to measure. For example, taking the constructs of verbal ability, numerical ability, mechanical ability, visual / spatial ability and diagrammatic / fault finding ability, it is interesting to see how established tests of these tie in with their TTB counterparts.

In the verbal domain, VT1.1 and VT5.1 correlate quite highly with each other and with their verbal equivalents from other batteries such as Birkbeck B1 (0,48 and 0,47).

The numerical domain, (NT2.1 and NT6.1) shows an even higher degree of overlap including correlations of 0,54 and 0,43 with Birkbeck B3 and 0,74 and 0,69 with Birkbeck B4. In the mechanical domain (MT4.1) has a strong correlation with Birkbeck B2 (0,68). The visual / spatial area (ET3.1 and ST 8.1) is supported by correlations with Birkbeck B5 (0,55 and 0,40) and the diagrammatic / fault finding domain (FT7.1) by a correlation of 0,69 with FTS4.

Overall, the pattern of intercorrelations between original and revised versions of the TTB tests and with their closest equivalents from other batteries supports the construct validity of the TTB.

## APPENDIX B (De Beer, 2005)

### 1. *Reliability of the LPCAT*

Test-retest reliability, parallel forms reliability and split-half reliability do not apply to computerised adaptive testing (De Beer, 2005). This is because of the interactive selection of items from an item bank which results in different sets of items being administered to each examinee. According to Anastasi and Urbina (1997) one classical test theory method of evaluating reliability that can be applied is the internal consistency or coefficient alpha index, which also reflects the homogeneity of content. This aspect relates to the one-dimensionality requirement for using item response theory.

The reliability indices available for the LPCAT are the coefficient alpha values (See Appendix B, Table B1). The item response theory equivalent to test score reliability is the test information function, which allows for the calculation of the standard error at specific ability levels because it is dependent upon the level of test information at that particular level. De Beer (2005) states that the standard error of measurement at any level of ability is the reciprocal of the square root of the amount of test information at that level. The accuracy of the ability estimation is used as one of the termination criteria in adaptive testing, therefore equivalent accuracy of measurement at different ability levels is more attainable than with standard tests. Based on the test information available in the LPCAT pre-test and post-test item banks respectively, in the pre-test, roughly 68% of the *t*-score estimates will fall between -2,4 and +2,4 *t*-scores from the estimated ability level, and for the post-test, roughly 68% of the *t*-score estimates will fall between -1,7 and +1,7 *t*-scores from the estimates pos-test ability level (De Beer, 2005). The information levels at the extremes of ability are lower than in the centre region, this fact means that more items will have to be administered to examinees who perform at either of the extremes to reach the required levels of accuracy for their ability estimation (De Beer, 2005).

Table B1

*Coefficient Alpha values for the two paper-and-pencil test forms for two different groups.*

Group	Form A (168 items)		Form B (168 items)	
	N	Alpha	N	Alpha
Total group	1277	0,981	1173	0,978
African	639	0,975	554	0,971
Coloured	303	0,969	296	0,970
White	335	0,925	323	0,926
African language group	639	0,975	554	0,971
English/Afrikaans language group	638	0,973	619	0,971
Male	636	0,981	589	0,979
Female	640	0,980	584	0,978
Grade 7	622	0,980	600	0,977
Grade 9	653	0,981	572	0,979

Based on all the items that were administered for item analysis, the coefficient alpha values ranged between 0,925 and 0,979 for the various sub-groups. The alpha value for the total group was 0,981 for form A and 0,978 for form B, indicating high internal consistency. Coefficient alpha is regarded as an index of reliability in standard test and, according to Gregory (2000), can be seen as an index of the degree to which a test measures a single factor. Table C1 provides the coefficient alpha values for the two paper-and pencil test forms for other the total group and various sub-groups. The high values obtained for the coefficient alpha provide support for the one-dimensionality of the LPCAT items – a requirement for the use of the three-parameter item response theory model.

## **2. Content validity of the LPCAT**

The content of the LPCAT were evaluated by a panel of experts from the Human Sciences Research Council and approved for assessing general, nonverbal reasoning ability by means of figural content (De Beer, 2000). That means that the content of the test indeed determine the

universe of behaviour it has been designed to sample. The LPCAT items provide acceptable content validity for measuring of, similar to other culture-fair tests like Raven's Progressive Matrices (Raven, Court & Raven, 1977) and Cattell's Culture-fair Intelligence Test (Cattell, 1963).

### ***3. Criterion-related validity: Using criterion-prediction procedures to evaluate predictive utility***

According to De Beer (2000) criterion-related validity is evaluated by comparing the test scores with one or more external variables or criteria which provide a direct measure of the characteristic or behaviour in question. This is done by using test scores to identify present performance on a criterion (concurrent validity) or to predict future performance on a criterion (predictive validity) (De Beer, 2000). The correlation of a new test against previously available tests is also often included as evidence of criterion-related validity. The main aim of criterion-related validity is to assess the practical validity and utility of a test for a specified purpose (De Beer, 2000). The LPCAT results were compared with ABET training results for an adult group and with academic results for secondary and tertiary groups. For the low literate adult group, correlations between LPCAT results and training results ranged between 0,398 and 0,610, while for a secondary school level sample, correlations of academic results and LPCAT performance ranged between 0,439 and 0,543 (De Beer, 2000). In a separate study with a group of bridging students, the correlation of LPCAT and academic results ranged between 0,313 and 0,525 (De Beer, 2002) These results provide support for the validity of the LPCAT in the multicultural South African context.

### ***4. Construct validity of the LPCAT***

Factor analysis was done to determine the construct validity of the LPCAT. By comparing LPCAT results with results of existing cognitive instruments, correlations were statistically highly significant and ranged between 0,400 and 0,645 for comparison with the Paper-and-Pencil Games (Claassen, 1996), and between 0,567 and 0,691 for comparison with the General Scholastic Aptitude Test (Claassen, De Beer, Hugo & Meyer, 1991). The LPCAT factor analysis results indicate a one-dimensional factor structure for the total group as well as for specific

subgroups. It therefore indicates construct validity in that the same theoretical construct is indicated for the total group and for the subgroups (De Beer, 2000).

## APPENDIX C

### PHASE TEST 1

#### BASIC TRAINING MODULES

<u>MODULE</u>	<u>CONTENT</u>
CCD	Course Control Document
ST	Perform safe work practices (Acts & Regulations)
UP - T	Convert units and prefixes
MA - T	Explain electromagnetism
EU - T	Analyse electrical units
PS - T - 1	Explain single phase power supplies
PS - T - 2	Explain three phase power supplies
EM - T - 1	Explain the use of electrical measuring instruments
EM - T - 2	Read energy meters
TI - 1	Use electrical testing instruments
PW - 1	Measure resistance
CC - T	Select conductors and cables
DRG	Read and interpret technical drawings
MMI - E	Measure and mark off material
MG	Use and maintain a pedestal grinder
OPE-1	Use and maintain a portable drilling machine
OPE-2	Operate a portable angle grinder
HM-2	Use and maintain a pedestal drill
HM-1	Use hand tools

### PHASE TEST 2

#### BASIC TRAINING MODULES

<u>MODULE</u>	<u>CONTENT</u>
CT - 1	Test a low tension cable
CA - 1	Fit a 16 ampere plug top
CA - 2	Fit a compression gland

CA - 3	Fit a gland
CA - 4	Join a low tension cable
EL - T	Explain single phase earth leakage protection circuits
EL - T - 1	Explain three-phase earth leakage protection circuits
CW - T	Wire a panel
TC - 1	Connect transformers
FF - 1 - E	Fault find basic electrical circuits
LS - 1	Light switching
LC - 1	Identify lights and lighting circuits
SW - 1	Wire a three-heat stove switch
SW - 2	Wire a five-hat stove switch
SW - 3	Wire a thermostatic control switch
CW - 1	Construct and wire installations
IT - 1	Test a single phase installation
CO-T	Identify contactors and overloads for various applications

**PHASE TEST 3**

**BASIC TRAINING MODULES**

**MODULE**

**CONTENT**

EC-1	Wire circuits
SD - T	Explain star & delta connections
CM-T	Explain induction motors contactors

**ADVANCED TRAINING MODULES**

**MODULE**

**CONTENT**

SD - 2	Wire a semi-automatic start-delta starter
--------	---

**PHASE TEST 4**

## **BASIC TRAINING MODULES**

### **MODULE**

### **CONTENT**

MT-2

Test three-phase motors

MP-1

Phase out an induction motor

## **ADVANCED TRAINING MODULES**

### **MODULE**

### **CONTENT**

SD – 3

Wire a fully automatic start-delta starter

SL – 1

Wire a resistance starter to a slip ring motor

SL – 2

Wire a liquid starter to a slip ring motor

AT – 1

Wire a fully automatic auto-transformer starter

PM – T

Explain pole changing motor theory

PM – 1

Connect a three-phase pole changing motor to a starter and selector switch

PM – 2

Connect a three-phase pole changing motor to contactors

PM – 3

Connect a three-phase changing motor to a fully automatic starter

## **PHASE TEST 5**

## **BASIC TRAINING MODULES**

### **MODULE**

### **CONTENT**

EM - 1

Connect a single phase energy meter

TT - 1

Explain single phase transformers

TT - 2

Explain three-phase transformers

SP-T

Explain single phase motors

DO-L

Wire a direct-on-line starter

PW-3

Wire a three-phase reversing starter

PW-2	Wire sequence starters
BIF	Examine bearings
BRR-1	Mount a bearing manually
BRR-2	Remove a bearing mechanically
BRR-3	Remove and mount bearings
MO-1-MW	Overhaul a three-phase motor

## **PHASE TEST 6**

### **BASIC TRAINING MODULES**

#### **MODULE**

#### **CONTENT**

MT-1	Test a single phase motor
SP-1	Connect a flexible cord to a single phase motor
SP-2	Connect a single phase motor to a selector switch
SP-3	Connect a single phase motor to reversing

### **ADVANCED TRAINING MODULES**

#### **MODULE**

#### **CONTENT**

PW - 4	Wire a three-phase balanced load
PW - 5	Connect three single phase energy meters
IT - 2	Test a three-phase installation
FF - 2 - E	Fault find live electrical circuits
DC - T	Explain the theory of direct current machines
DC - 1	Connect a load to a DC generator
DC - 2	Connect a DC motor and starter
DC - 3	Connect a Ward-Leonard speed control
EM - 2	Connect a three-phase energy meter
EM - 3	Connect a three-phase energy meter with range extension
RA - E - 2	Maintain auto-electrical systems

SD - 1	Wire a manual start-delta starter
SL - 3	Wire a grid and controller to a slip ring motor
SL - 4	Wire a vapour matic starter

### **PHASE TEST 7 & 8**

#### **BASIC ELECTRONICS TRAINING MODULES**

<b><u>MODULE</u></b>	<b><u>CONTENT</u></b>
HM - 3	Solder components
BE - 1	Identify and use resistors
BE - 2	Identify and use inductors
BE - 4	Identify diodes
TI - 2	Use an oscilloscope
BE - 5	Construct rectification circuits
BE - 6	Identify and use Zener diodes
BE - 7	Construct a voltage doubler
BE - 8	Identify transistor action
BE - 9	Identify transistor configurations
BE - 10	Test regulated power supplies
BE - 11	Test thyristors
BE - 12	Test thyristor phase control
BE - 13	Explain and apply PLC operating principles
FF - 1	Fault find circuits 1

#### **ADVANCED ELECTRONICS TRAINING MODULES**

<b><u>MODULE</u></b>	<b><u>CONTENT</u></b>
IE - 1	Test an amplifier
IE - 2	Use a transistor as a switch
IE - 3	Test multi vibrators
FF - 2	Fault find circuits 2

IE - 4	Identify operational amplifiers
PCC	Construct a continuity tester
IE - 5	Test a field effect transistor
IE - 6	Test logic gates
FF - 3	Fault find circuits 3
IE - 7	Test electronic systems