

**DYNAMIC ASSESSMENT OF COGNITIVE ABILITY:
INVESTIGATING THE CONSTRUCT VALIDITY OF THE
LEARNING POTENTIAL COMPUTERISED ADAPTIVE
TEST (LPCAT) WITHIN AN ACADEMIC CONTEXT**

GANIYAT OLUSHOLA ZURAKAT

**Submitted in partial fulfilment of the requirements for the degree of
Master of Social Science (Educational Psychology), in the school of Applied
Human Sciences, College of Humanities, University of KwaZulu-Natal,
South Africa**

Supervisor: Dr Carol Mitchell

Co-Supervisor: Dr Nicholas Munro

NOVEMBER 2018

DECLARATION

I declare that this dissertation: *Dynamic assessment of cognitive ability: Investigating the construct validity of the Learning Potential Computerised Adaptive Test (LPCAT) within an academic context*, is my original work. All the sources used have been indicated and acknowledged by means of complete referencing. This dissertation is being submitted as partial fulfilment of the requirements for the degree of Master of Social Science (Educational Psychology), in the school of Applied Human Sciences, College of Humanities at the University of KwaZulu-Natal, Pietermaritzburg, South Africa. No part of this work has been submitted previously for any degree or examination at any other University.

Student's signature

Date

Supervisors' approval of this thesis for submission

As the candidate's supervisors, we have approved this thesis for submission

Supervisor's signature

Date

Co-supervisor's signature

Date

ACKNOWLEDGEMENTS

He who is not grateful to man, will not be grateful to God

All praise and adorations belong to *Allah* (God), the lord of incomparable majesty. I am indeed grateful to Him for He it is who spared my life and bestowed on me the wisdom to complete this work. I owe the entirety of my life and success to Him! I would also like to acknowledge the following people who have contributed in a way or the other to other successful completion of this thesis:

I am wholeheartedly appreciative of my beloved, Ridwan Aiyetoro, for his unflinching support in all ramifications. Thank you dearie, for giving me the push and ensuring I did not rest on my oars, till the completion of the work. I am also indebted to my two little cuties - Zayd and Sumayyah, for accommodating my stealing out of the time we should have spent together while I completed this work. My love for you two knows no bounds!

I am also thankful to my parents and my siblings, for the bond we share, as well as the support and encouragement we often give one another. Thank you for always asking.... “how far with your thesis” this went a long way to boost my morale.

I must also acknowledge my parents and sisters in-law, for their support, encouragement and well-wish.

My supervisor, Dr. Carol Mitchell and my Co-supervisor, Dr. Nicholas Munro, you two really eased my task, with your constructive criticisms and reference sources. It would have been a lot more difficult completing this work, but for your guidance and timeous feedback. I am indeed grateful!

Cebisa Nkatu, for assisting with the test administration and seeing to necessary logistics during the data collection process.

Jopie Van Rooyen (JVR) test distributors for their generous donation of the SPM test booklets. Your generosity contributed a great deal to the success of this work.

Teslim Babalola, for his expertise on data analysis.

Dr. De Beer, for permitting me to sit in on the LPCAT training to enable me to administer it on my participants.

Jennifer Swift, for being my research buddy and keeping me abreast of deadlines.

I must also not forget to thank all the lecturers and school administrators, who offered me a platform to publicise and recruit participants for the study. I am particularly grateful to Nurudeen Ajayi, for his help in this regard.

Finally, I say thank you to the students who volunteered their time for this study, for they are part of the reason this thesis is completed!

DEDICATION

This is to my husband, the one person that encourages me to step out of my comfort zone, just to ensure I achieve my goal.... I admire and love the person that you are!

To my two sweet little bundles, Zayd and Sumayyah. I am eternally grateful for having you two.

This research work is also dedicated to all those who champion the cause of a fairer assessment of learners who are disadvantaged by their impoverished backgrounds.

ABSTRACT

No doubt, intelligence testing is vital for placement, intervention and other academic purposes, however, when factors that might disadvantage the individual, such as their impoverished socio-economic backgrounds are not considered in the process, such endeavours might further disadvantage the individual. In a bid to fill this assessment loophole, which is more peculiar to the static approach, the psychological assessment community witnessed the emergence of the dynamic approach to assessment, which gives hope for fairer assessments through its test-train-test approach. Although several measures using this approach have been developed, the dearth of empirical evidence on their psychometric properties still limits their popularity and acceptance across the globe. This study therefore investigated the construct validity of a locally developed dynamic measure of learning potential, the Learning Potential Computerised Adaptive Test (LPCAT), by comparing scores on it to those obtained on the Raven's Standard Progressive Matrices (SPM). The intent of this endeavour is for findings to lend credence to the use of the LPCAT within the South African context and invariably, the dynamic assessment approach. The study also investigated the influence of demographic factors (race, gender, socio-economic status (SES) and English language proficiency) on the LPCAT, with the intent of verifying its culture-fairness. The study adopted a quantitative approach and the sample included 92 undergraduate students, conveniently drawn across the Pietermaritzburg campus of the University of Kwa-Zulu Natal. Findings indicated a strong positive relationship between the LPCAT and the SPM, while none of the variables tested had a significant effect on the LPCAT scores, aside from English language proficiency. The study therefore affirmed the construct validity and culture-fairness of the LPCAT with a caution that the language proficiency of the examinee be considered to ensure a culture-fair assessment.

Keywords: Dynamic assessment, cognitive assessment, static approach, construct validity, culture-fairness, intelligence, zone of proximal development, learning potential

LIST OF ACRONYMS AND ABBREVIATIONS

ALD	Actual Level of Development
APA	American Psychology Association
APIL	Ability, Processing of Information and Learning Potential Battery
APM	Advanced Progressive Matrices
CHC	Cattell – Horn – Carroll theory
DA	Dynamic Assessment
ITC	International Test Commission
JSAIS	Junior South African Individual Scales
KAIT	Kaufman Adolescent and Adult Intelligence Tests
K-ABC	Kaufman Assessment Battery for Children
LPCAT	Learning Potential Computerised Adaptive Test
RPM	Raven’s Progressive Matrices
SES	Socio-economic status
SPM	Standard Progressive Matrices
SSAIS-R	Senior South African Individual Scales – Revised
TRAM	Transfer, Automatization, Memory and Understanding Learning Potential Battery
UKZN	University of KwaZulu-Natal
WAIS	Wechsler Adult Intelligence Scales
WISC	Wechsler Intelligence Scale for Children
XBA	Cross - Battery Approach
ZPD	Zone of Proximal Development

LIST OF TABLES

Table	Description	Page
Table 1	Age distribution of sample	48
Table 2	Description of sample by gender	49
Table 3	Racial distribution of sample	49
Table 4	Socio-economic distribution (indicated by school Quintile of sample)	49
Table 5	English language proficiency of sample	50
Table 6	Descriptive analysis of scores on the LPCAT and the SPM	50
Table 7	Correlation of LPCAT (pre – test) scores with SPM scores	51
Table 8	Model summary of regression (pre-test)	52
Table 9	Results of ANOVA ^a for regression (pre–test)	52
Table 10	Regression Coefficients ^a (pre–test)	53
Table 11	Model Summary of regression (Post–test)	53
Table 12	Results of ANOVA ^a for regression (post-test)	54
Table 13	Regression Coefficients ^a (post – test)	54
Table 14	Model Summary of regression (composite score)	55
Table 15	Results of ANOVA ^a for regression (composite score)	55
Table 16	Regression Coefficients ^a (composite score)	55

LIST OF FIGURES

Figure	Description	Page
Figure 1	Carroll's three-stratum model of human intelligence	15
Figure 2	The Cattell-Horn-Carroll model	16
Figure 3	The Zone of Proximal Development	33

TABLE OF CONTENTS

INTRODUCTION TO THE STUDY	1
1.0 Introduction	1
1.1 Statement of the Research Problem	1
1.2 Research Aims and Hypotheses	3
1.2.1 Primary research aim and hypothesis.	4
1.2.2 Secondary research aim and hypotheses.....	4
1.3 Research Methodology.....	5
1.4 Ethical Considerations.....	5
1.5 Theoretical Framework for the Study	5
1.6 Definition of Terms	5
1.7 Outline of the Thesis	6
LITERATURE REVIEW	8
2.0 Introduction	8
2.1 Overview of Psychological Assessment	8
2.1.1 Definition and forms of psychological assessment.....	8
2.1.2 Applications of psychological assessment.....	9
2.1.3 Standards and ethical issues in psychological assessment.....	10
2.1.4 Investigating the construct validity of assessment measures.....	10
2.2 Cognitive Ability in Focus	13
2.2.1 Theories of intelligence.	13
2.2.2 Assessment of cognitive ability.....	17
2.2.3 Controversies around intelligence testing	18

2.3	Paradigms of Psychological Assessment	20
2.3.1	Static assessment method.	21
2.3.2	Dynamic assessment method.....	22
2.3.2.1	Criticisms of dynamic assessment.....	25
2.4	The Learning Potential Computerised Adaptive Test (LPCAT).....	28
2.4.1	Description of the LPCAT.....	28
2.4.2	Review of literature on the LPCAT.....	28
2.5	The Raven’s Standard Progressive Matrices (SPM).....	29
2.5.1	Overview of the SPM.....	29
2.5.2	Review of literature on the SPM.....	30
2.6	Review of Literature on the LPCAT and the SPM	31
2.7	Theoretical Framework	32
2.7.1	Vygotsky’s theory of socio-cultural development.....	32
2.7.2	Some criticisms against Vygotsky’s zone of proximal development.....	34
2.8	Conclusion.....	35
	RESEARCH METHODOLOGY	36
3.0	Introduction	36
3.1	Research Aim and Hypotheses.....	36
3.1.1	Research questions and hypotheses.....	37
3.2	Research Design.....	38
3.2.1	Research paradigm.....	38
3.2.2	Quantitative research approach.....	388
3.2.3	Research location.....	38
3.2.4	Sample and sampling procedures.....	39
3.2.5	Data collection instruments.....	39

3.2.6	Procedure.....	40
3.2.6.1	Ethical approval/ recruitment of participants.....	40
3.2.6.2	Test administration.....	41
3.3	Data Analysis	42
3.3.1	Data cleaning.....	42
3.3.2	Quantitative data analysis.	43
3.3.3	Validity, reliability and rigor.....	44
3.4	Ethical Considerations.....	45
3.5	Limitations of the Study.....	47
3.6	Summary	47
	RESULT OF DATA ANALYSIS	48
4.0	Introduction	48
4.1	Descriptive Analysis	48
4.1.1	Statistical distribution of sample..	48
4.1.2	Basic descriptive analysis of main variables tested.....	50
4.2	Inferential Statistical Analysis	51
4.2.1	Results of correlation analysis.	51
4.2.2	Results of multivariate analysis (multiple regression analysis).	52
4.2.2.1	Hypothesis 1: Relationship between multiple variables (race, SES, English language proficiency and gender) and LPCAT pre-test scores.....	53
4.2.2.2	Hypothesis 2: Relationship between Multiple Variables (race, SES, English language proficiency and gender) and LPCAT Post - Test scores.....	53
4.2.2.3	Hypothesis 3: Relationship between Multiple Variables (race, SES, English language proficiency and gender) and LPCAT composite scores.....	56

4.3	Summary	57
	DISCUSSION OF FINDINGS	57
5.0	Introduction	57
5.1	Discussion of Results	58
5.1.1	Main research aim: construct validity of the LPCAT.	58
5.1.2	Secondary research aim: Effect of selected demographics on LPCAT pre-test, post-test and Composite scores.....	59
5.1.2.1	Hypothesis 1: The effect of race, SES, English language proficiency and gender on LPCAT pre-test scores.....	60
5.1.2.2	Hypothesis 2: The effect of race, SES, English language proficiency and gender on LPCAT post-test scores.....	61
5.1.2.3	Hypothesis 3: The effect of race, SES, English language proficiency and gender on LPCAT composite scores.....	62
5.2	Summary	63
	SUMMARY, CONCLUSION AND RECOMMENDATIONS.....	65
6.0	Introduction	65
6.1	Summary of Research Process and Findings	65
6.2	Research Conclusions and Implications.....	66
6.3	Limitations of the Study.....	67
6.4	Recommendations for Future Research	67
6.5	Recommendations for Practice.....	68
6.6	Conclusion.....	69
	REFERENCES.....	70
	Appendix A: UKZN HSSREC Ethical clearance (Initial Copy)	80
	Appendix B: UKZN HSSREC Ethical clearance (Subsequent copy)	81
	Appendix C: Gatekeeper’s permission	83

Appendix D: Recruitment poster83

Appendix E: Demographic details sheet.....84

Appendix F: Informed consent form.....85

Appendix G: Research Agreement (SPM).....88

Appendix H: Acknowledgement of LPCAT use91

Appendix I: LPCAT Training Certificate92

CHAPTER ONE

INTRODUCTION TO THE STUDY

1.0 Introduction

The validity of results obtained from intelligence testing, particularly when administered on individuals considered to be disadvantaged, either by their socio-economic status (SES) or previous educational opportunities has, within the last two decades, been an issue of concern among researchers and practitioners. The issue of contention is that when assessment is administered in a way that lends no credence to the heterogeneous composition of individuals tested, as in the static assessment method, test results yield unfair description of disadvantaged learners, thereby limiting their opportunities. In a bid to overcome this, Dynamic Assessment (DA), envisioned as an assessment method that is less susceptible to the individual differences, is beginning to gain favour, with more and more research conducted to investigate its prospects.

This study was therefore positioned within the context of DA, as an inquiry into the validity of a dynamic assessment measure of learning potential, for use within the South African context. Specifically, the study investigated the validity of the Learning Potential Computerised Adaptive Test (LPCAT), by comparing it with the Raven's Standard Progressive Matrices (SPM), which adopts a static assessment method. This chapter sheds light on the research problem that motivated the study, as well as the objectives for the pursuit and the hypotheses that data collected will be tested against. A brief account of the research design, ethical considerations and the theoretical underpinning of the study is also presented in this chapter. Lastly, key terms, as operationalized within the study are highlighted.

1.1 Statement of the research problem

An aspect within the science of psychological assessment that continues to generate a lot of discourse is the distinction between static and dynamic assessment. In the static assessment approach, evaluation is based on accumulated skills and knowledge, without consideration of the individual's potential for learning (Cho, Compton, Gilbert, Steacy,

Collins & Lindström, 2015). Also, the individual's performance on an assessment is either correlated with those of other individuals who are considered representative of the intended group (norm-referenced) or correlated against a set criterion that the individual is expected to have met at a particular age (criterion – referenced; Brown & Knight, 2012). Most currently available and widely used tests of intelligence, locally and internationally, are designed as static measures and are norm- referenced. This includes the Wechsler Adult Intelligence Scales (WAIS), Raven's Progressive Matrices (RPM), Kaufman Adolescent and Adult Intelligence Tests (KAIT), Kaufman Assessment Battery for Children (K-ABC), Wechsler Intelligence Scale for Children (WISC), Junior South African Individual Scales (JSAIS), Senior South African Individual Scales – Revised (SSAIA-R), amongst others.

Although in the past decades, the static approach gained prominence globally, advancements in the field of assessment have unearthed some shortcomings. Some of these are its inability to give information about the individual's future capability (Hamers & Resing, 1993; Murphy & Maree, 2006; Sattler, 2008), culture-fairness issues (Van de Vijver, 2002), the lack of details on the cognitive deficits that mitigate performance on tests (Cho et al., 2015) amongst others. Therefore, for a country like South Africa, characterized by cultural, racial, socio-economic and linguistic diversity, the use of static assessment measure alone raises ethical questions, since the norm sample of these tests do not take into account the contextual diversity, whereas, culture-fairness and contextual appropriateness of tests are stipulated core values to the profession (APA, 2010; ITC, 2013).

Dynamic assessment approach on the other hand, enables the evaluation of the individual's potential for skill and knowledge acquisition, through its “test-train-test” approach to assessment. This approach implies that a training on the skills that are tested by the assessment is included within the assessment process. This therefore, gives the test taker an opportunity to ‘learn on the job’, reduces forms of unfairness and places test takers on an equal level (Cohen & Swerdlik, 2010). Reflecting on the use of this approach locally, De Beer (2003, p.718) stated that “a focus on future development by means of the measurement of potential can address both legislative and practical requirements in a country such as South Africa with its diversity of people”. She further stated that “the measurement of learning potential can also help to identify appropriate levels of training to be provided over a broad spectrum of ability, without necessarily relying on language proficiency or prior

formal training, thus providing useful information for the purposes of selection or training and development” (p. 178).

Although the Dynamic assessment approach has continued to enjoy attention, because of its capacity for educational, environmental and cultural fairness, the dearth of empirical evidence on its psychometric properties is still a limitation (De Beer, 2003; Lidz, 2009; Smit, 2010). Even though there has been a great deal of research that addresses this problem area (Caffrey, date? ; De Beer, 2010; Fuchs & Fuchs, 2008; Grigorenko & Sternberg, 1998; Guthke & Stein, 1996; Lidz, 2009; Murphy, 2002), there continues to be a need for more to be done to strengthen the evidence base of dynamic assessment. As Murphy and Maree (2006, p. 173) put it, “the ever-present issue with dynamic assessment is the fact that many dynamic batteries do not always address issues such as reliability and validity”, whereas, “although there is some disillusionment with psychometric tests, the scientific accuracy of many instruments belonging to this paradigm is a redeeming feature of this approach”.

The present study was therefore an attempt to bridge the gap within the practice of dynamic assessment, through investigating the construct validity of a dynamic measure of learning potential - The Learning Potential Computerised Adaptive Test (LPCAT). The endeavour not only addressed the issue of empirical evidence on dynamic assessment, it also evaluated the acclaimed culture-fairness of the dynamic assessment approach. This pursuit will hopefully add to the body of knowledge on the science and practice of dynamic assessment and ultimately, assist psychologists in making more informed decisions about LPCAT use, and in turn, DA assessment methods.

1.2 Research aims and hypotheses

The overarching aim of this study was to generate empirical evidence on the construct validity of the LPCAT. This was accomplished by comparing performance on it to that of the Raven’s Standard Progressive Matrices (SPM), which has a repertoire of empirical evidence supporting its validity. To achieve this, primary and secondary aims were designed, alongside null and alternate hypothesis, to guide the quantitative investigation of the constructs under study.

1.2.1 Primary research aim and hypothesis. The primary aim of this study was to determine if there is a significant relationship between undergraduate students' performance on the LPCAT and the SPM. Consequently, this was achieved by testing the following null and alternate hypothesis:

- **H₀:** There is no significant relationship between students' performance on the LPCAT and the SPM
- **H_a:** There is a significant relationship between students' performance on the LPCAT and the SPM.

1.2.2 Secondary research aim and hypotheses. The secondary aim of this study was to determine if selected demographic variables, specifically; race, SES, gender and language proficiency impact students' performance on the LPCAT, this is to provide elaborate and in-depth understanding of the research aim. To evaluate this, three secondary objectives and hypotheses were designed, as described below;

1. Do selected demographic variables have a statistically significant effect on students' LPCAT pre-test scores?

H₀: There is no statistically significant relationship between selected demographic variables and students' LPCAT pre-test scores.

H_a: There is a statistically significant relationship between selected demographic variables and students' LPCAT pre-test scores.

2. Do selected demographic variables have a statistically significant effect on students' LPCAT post- test scores?

H₀: There is no statistically significant relationship between selected demographic variables and students' LPCAT post-test scores.

H_a: There is a statistically significant relationship between selected demographic variables and students' LPCAT post-test scores.

3. Do selected demographic variables have a statistically significant effect on students' LPCAT Composite scores?

H₀: There is no statistically significant relationship between selected demographic variables and students' LPCAT Composite scores.

H_a: There is a statistically significant relationship between selected demographic variables and students' LPCAT Composite scores.

1.3 Research methodology

This research study adopted a quantitative approach to explore and measure the degree of relationship between the LPCAT and the SPM, as well as the degree of variance between the LPCAT and the selected variables. The sample for the study comprised undergraduate students from the Pietermaritzburg campus of the University of KwaZulu-Natal (UKZN). To achieve the study aim, a dynamic assessment instrument, the LPCAT and a static assessment instrument, the SPM were both administered to the participants (N=95), after which descriptive and inferential analysis were conducted, to examine the research hypothesis. A detailed account of the research method and design is contained in Chapter 4.

1.4 Ethical considerations

This research was conducted in accordance with the ethical policy of the institution within which it was conducted. As per requirement, ethical clearance was sought and granted (Appendix A & B) by the Human and Social Sciences Research Ethics Committee (HSSREC), for the conduct of the research. At the commencement of each assessment session, each participant signed an informed consent form, which detailed the research aim and their right to voluntary participation. Confidentiality of research participants was ensured throughout as no personal details was divulged in the presentation of findings or analysis of results.

1.5 Theoretical framework for the study

The study was underpinned by Vygotsky's theory of Socio-Cultural Development, particularly, and his notion of the Zone of Proximal Development (ZPD). This theory paved way for the understanding of dynamic assessment and its relevance to the practice of psychological assessment. A detailed description of this theory is contained in chapter 2.

1.6 Definition of terms

Pertinent concepts, as operationalised within the study context are stated below:

Psychological assessment

This refers to the scientific process of gathering information and gaining an understanding

of an individual's development and functional abilities across a myriad of areas of functioning.

Dynamic assessment

This approach to assessment aims at evaluating the potential of the individual, to acquire new knowledge and improve his/her current performance level, in addition to evaluating the current level of functioning. There is usually some degree of instruction and feedback, within the assessment process.

Static assessment

This describes an approach to the assessment of an individual's existing abilities that is conducted without assistance to the student. This kind of method therefore, gives information on crystallised abilities.

Cognitive ability

The mental process of knowing, including aspects such as awareness, perception, reasoning, and judgment.

Intelligence

Ability to understand what is learned, perceived or reasoned. It could also mean ability to apply knowledge to manipulate environment or things abstractly. Fluid intelligence of participants in this study was measured by their SPM and LPCAT pre-test scores.

Culture-fairness

This is one of the ethical considerations in the selection of tests for assessment, it implies the contextual appropriateness of the test, excluding it from bias of all nature.

Learning potential

Within this study, learning potential implies the capacity to gain knowledge and skills, if given training. Learning potential was therefore measured by scores on the LPCAT.

1.7 Outline of the thesis

Subsequent chapters of the research study are delimited as below;

Chapter 2 comprises a review of relevant literatures on dynamic assessment, particularly the LPCAT, highlighting their findings and the gaps therein. The chapter also discusses the theoretical underpinning of the study.

Chapter 3 is a presentation of the design and methodology adopted in the research investigation.

Chapter 4 entails a presentation of the research findings and analysis.

Chapter 5 is a discussion of the findings from the investigation, as well as presentation of the limitations, implications and suggestions for further research.

Chapter 6 is the concluding chapter

CHAPTER 2

LITERATURE REVIEW

2.0 Introduction

This chapter entails a discussion of topics relating to cognitive assessment. It begins with a general overview of the concept of psychological assessment, including its definition, forms, applications, and ethical issues, as well as the process of investigating the construct validity of measures. The chapter progresses with a focus on cognitive assessment, where the discussion covered the theories of intelligence, measures of intelligence and an expose on the current controversies around intelligence testing. Reference is also made in this chapter, to the contending issues within both the traditional and dynamic assessment paradigms. Towards the end of the chapter is an introduction of the two assessment instruments that are used in the study; the Learning Potential Computerised Adaptive test (LPCAT) and the Raven's Standard Progressive Matrices (SPM) and, a review of empirical studies that have been conducted on each and both of them. The chapter ends with a brief outline of the theoretical framework underpinning the study, with a highlight of some of its alleged shortcomings.

2.1 Overview of psychological assessment

2.1.1 Definition and forms of psychological assessment. Psychological assessment is one of the core aspects of psychological practice and it is described as a scientific process of gathering information and gaining an understanding of an individual's development and functional abilities (Lubbe, 2004). Aiken (2003) also mentions that psychological assessment implies that which is used to make decisions about people, perhaps in job settings, and to aid people in decision making, with regards to their future occupation, education or other status.

Psychological assessment can be applied to evaluate personality traits, behaviour, cognitive functioning and other characteristics, in order to make predictions, or judgements about people (Aiken, 2003). However, the assessment of cognitive functioning features prominently in the science of psychological assessment, with the first measure dating back

to 1905, when Alfred Binet and Theodore Simon developed the first test of intelligence (Binet & Simon, 1916).

Assessment can be used to gather varying information about an individual, this ranges from cognitive functioning, visual and auditory perception, emotional skills, language abilities, motor skills to the learning needs of the individual (APA, 2013). Aside from these, behavioural and neuropsychological assessments can also be carried out on an individual (Groth-Marnat, 2009), depending on the referral question the assessment serves to provide answers for.

2.1.2 Applications of psychological assessment. Assessment can be conducted in several settings, including psychiatric, medical, legal, educational, and in a psychological clinic (Groth-Marnat, 2009; Merrell, Ervin, & Peacock, 2011; Rust & Golombok, 2014). Psychological assessment can also be applied in businesses, industries, military, forensic and civil services (Aiken, 2003). Within these settings, psychological assessments can be used as part of the process for selection (to determine if an individual has the minimum skills required to be regarded as being successful), placement (to determine if a student, for example, should take a particular subject, and at what stage) classification (to determine the place of an individual in a category) purposes, as well as for the evaluation of academic achievement (Huysamen, 1996; Groth-Marnat, 2009; Rust & Golombok, 2014). Huysamen (1996) further posits that assessments can be used to determine a child's readiness for school, for the provision of vocational guidance and remedial exercises and to function as a tool for diagnosing mental retardation by psychologists.

Lubbe (2004, p. 320) views the purpose of assessment to be “focused on obtaining an holistic view of the child in terms of competencies, assets, strengths and areas of difficulty”. She adds that assessment “determines the progress of significant developmental achievements, aids in placement and promotion decisions and in the diagnosis of learning, teaching and emotional or behavioural difficulties experienced by the child”. According to (Domino & Domino, 2006), psychological assessment can be beneficial for programme evaluation, scientific inquiry, self-understanding and classification purposes.

2.1.3 Standards and ethical issues in psychological assessment. During psychological assessments, psychologists are usually faced with a myriad of ethical challenges, ranging from issues around informed consent, computerised assessment, non-biased test instruments and use, to record keeping and who to include in the feedback session. To combat these challenges internationally, ethical guidelines emerged, as the practice of psychological assessment continued to evolve, to ensure appropriate professional conduct. For instance, the International Test Commission (ITC, 2013) laid down guidelines that must be adhered to with regards to the use of tests in the assessment process, some of which include;

- Ensuring sufficient validity and reliability indexes exist for tests used,
- Psychometric properties of tests must be available to the public,
- Test users must ensure they have the competency required for administering the test,
- Test materials should be well secured,
- Test results must be treated with confidentiality,
- Test use must be justifiable,
- Culture-fairness and contextual appropriateness of a test should be considered before a test is used,
- Adequate preparation must be made prior to the administration of tests, and
- The scoring and interpretation of tests should be adequately carried out.

In addition, the American Psychological Association (APA, 2010) stipulates that;

- Informed consent must be obtained prior to assessment,
- Test and assessment technique developers should ensure to use up-to-date procedures in their test construction, standardization and validation,
- No obsolete test or norms should be used, and that
- Adequate feedback must be given after assessment.

A breach in any of these could impact on the validity of the test or assessment process, hence, psychologists and test users must endeavour to adhere as much as they can.

2.1.4 Investigating the construct validity of assessment measures. In addition to understanding the operationalization, theoretical underpinning and standardization of test measures, reliability and validity evidences are also important factors to be considered in

test selection (Groth-Marnat, 2009). While reliability refers to the extent to which a test yields same result if used on different occasions, test validity implies the extent to which a test measures the variable it is intended to measure (Cook & Beckman, 2006), not only in terms of the abstract concept it measures, but also in terms of the context and group it is intended for (Groth-Marnat, 2009). The most important questions that should be asked with regards to the validity of a test measure therefore include; (i) what the criteria and procedures used to validate the test are and (ii) whether the test will suit the context and purpose for which it is intended. In relation to the aim of the current research work, validity of test measures will be expatiated, specifically construct validity.

In evaluating validity of test instruments, test validity is always viewed as a hypothesis, in the sense that the intended interpretative meaning of assessment data is at first hypothesized, thereafter data is collected and analysed to confirm or negate the hypothesis (Downing, 2003). Gliner, Morgan and Harmon (2001) elucidate three (3) main approaches to evaluating test validity; content validity, criterion and construct validity.

Content validity: This implies the appropriateness of the assessment measure to the concept that is intended to be measured. In the initial stage of test development, test developers need to first operationalize the variables that they intend to measure, thereafter generate items, based on their conceptualization of the variable (Groth-Marnat, 2009). Therefore, the detailed presentation of the operationalization of the construct, intended purpose of the instrument, item development and selection process, wording of items, as well as the qualification of test developers and reviewers are considered as evidence for determining the content validity of instruments (Cook & Beckman, 2006). Hence, the extent of agreement between experts on the subject matter, after a careful consideration of these details, is used to adjudicate the appropriateness of test items.

Criterion validity: Criterion validity entails the extent to which a measure relates to an outcome and it is usually divided into predictive and concurrent validity (Groth-Marnat, 2009). While predictive validity implies the extent to which a measure predicts a future outcome or performance, concurrent validity implies the correlation with another measure, administered at the same time. Criterion validity is usually measured by comparing score on an instrument with an outcome, to determine the degree of relationship.

Construct validity: Construct validity implies the extent to which a test score represents the construct or phenomenon that the test is intended to measure (Cook & Beckman, 2006). It usually entails measuring the strength of the relationship between scores on a test and those of another independent measure that has been adjudged as a standard against which the measure is validated. According to Foster and Cone (1995), it follows three (3) stages; (i) a theoretical understanding of the construct being measured (ii) an understanding of the relationship between the construct and other related variables (iii) hypothesis testing to confirm the existence of the relationship hypothesised. Therefore, the more research studies that confirm the hypothesized relationship, the higher the degree of confidence in the test.

A review of literature that has evaluated the construct validity of learning potential and intelligence indicates that a variety of methods have been used to determine construct validity. These include correlation studies to compare performance on a measure with another measure of similar construct, measure of the effect of interventions, as well as the evaluation of the extent to which a test positively correlates with similar variables and negatively correlates with variables that are not similar to it, referred to as convergent and discriminant validity respectively (De Beer, 2003; Maruff et al., 2009; Norris & Tate, 2000; Rushton, Skuy & Bons, 2004).

For instance, to investigate the construct validity of the LPCAT, De Beer (2003) compared results from the LPCAT post – test with those of other cognitive measures (Paper and pencil games test, General scholastic aptitude test, Cognitive processing profiler, Subtests of the Junior aptitude test and General mental ability). The construct validity of the LPCAT was confirmed by the statistically significant relationship found between the LPCAT and the other cognitive measures. A similar method was adopted by Rushton et al. (2004) in their investigation of the construct validity of the Raven’s Advanced Progressive Matrices (APM) for African and Non-African students. Construct validity was tested by comparing scores of the two groups of students on the test. Results supported an internal validity existed for the APM, since “items on the matrices “behaved” in the same way for the African students as they did for the Non-African students” (p. 227).

Another study conducted by Norris and Tate (2000) to investigate the construct validity of the Behavioural Assessment of the Dysexecutive Syndrome (BADS) was completed by comparing it with standard executive tests that measure same constructs. Correlation

indexes were also used to confirm test validity. Similarly, Maruff et al. (2009) conducted a study on the construct validity of a cognitive assessment battery (Cog state Brief Computerised Cognitive assessment Battery). This was achieved by analysis of the relationship between the battery and conventional neuropsychological measures that test similar construct. According to the authors, the neuropsychological measures that were used as comparisons were selected because of previous indications of their test validity. Pearson product moment correlation coefficients were therefore computed, and test validity was determined by the extent to which scores on each task of the battery related to those on the neuropsychological tests used as comparison. With results yielding acceptable coefficient indexes, construct validity of the instrument was documented.

Similar to previous research with the same aim, construct validity of the LPCAT, which is the primary aim of the current research work, was therefore investigated following the patterns that have been identified. In this context, the SPM is used as comparison. Detailed description of the research methodology is documented in chapter 3.

2.2 Cognitive ability in focus

2.2.1 Theories of intelligence. Cognition is a broad term for describing the ability to process information, solve problems, make decisions and expand knowledge (Human-Vogel, 2004). Cognitive ability therefore, is a multidimensional phenomenon that encompasses mental processes including knowledge, attention, memory, judgement and evaluation, reasoning and computation, problem solving and decision making, comprehension and, language production, which are believed to be positively correlated and represented by the full-scale score on tests of cognitive ability (Dickens, 2008). Cognitive assessment is usually undertaken to assess cognitive and intellectual levels of functioning, to make a diagnosis and to initiate an intervention plan (Van Eeden & De Beer, 2009).

Intelligence on the other hand, has long been a construct of controversy among psychologists, hence no single definition exists that best captures the term (Goldstein, Princiotta & Naglieri, 2015; Kline, 2013). Central to all definitions however, is that intelligence encompasses such abilities as abstract thinking, understanding, communication,

learning, reasoning, planning and problem solving (Gardner, 1993; Goldstein, Princiotta & Naglieri, 2015; Terman, 1922; Wechsler, 1944). Theories of intelligence and investigations in the field dates to the 19th century (Boake, 2002) and have witnessed a lot of advancements until the present time. An attempt is therefore made in this sub-section to briefly trace and outline the major developments in the conceptualizations of intelligence over time.

Spearman's two-factor theory (1904) has over time been a common point of departure for other intelligence theories and this usually reflects in the construction and development of assessment measures. According to Spearman (1904), people who are good at one type of mental activity tend to do well on others too, meaning that they have a positive correlation across other types of mental abilities. Thus, he hypothesized the concept of general mental ability and several other specific abilities. (Dickens, 2008). Spearman (1904), categorized intelligence into general ability (identified as g factor) and a group of specific abilities (s factor). The g factor includes aspects of intelligence like abstract thinking and problem-solving abilities while the s factor includes abilities such as motor speed, memory, attention etc. To date, there has been a lot of expansion and modification of the Spearman's g factor, but Cattell's (1987) categorisation into fluid (G_f) and crystalized (G_c) intelligence is relevant for this study.

Drawing from Spearman's g factor, Cattell (1987) proposed two kinds of g: The Fluid intelligence (G_f) and Crystallised intelligence (G_c). G_f is considered the ability to solve new problems through reasoning, which is impacted on by biological and neurological factors. G_c on the other hand, is knowledge-based and dependent on education and socialisation (Horn & Cattell, 1966). Cattell (1987) posits that G_c is evident in the areas of verbal, numerical reasoning, mechanical and experimental judgement while G_f is displayed when the task involved is presented in the form of series, classification, analogy, or topology. In other words, fluid ability is expressed where the need is to deduce complex relationships without recourse to a repertoire of knowledge, whereas, crystallised ability draws from previous learning or experience. Both the LPCAT and SPM, which are used in this study, are tests of fluid ability, meaning that an individual's performance on both tests is not intended to be influenced by previous learning or experience.

Also worthy of mention in this study is the Three-stratum theory of cognitive ability, proposed by American Psychologist, John Carroll (1997). The theory derives primarily from

an amalgam and expansion of Spearman's g and Cattell's Gf and Gc. The model proposes that intelligence is best understood as a three-layered hierarchy, that there are a fairly large number of distinct individual differences in cognitive ability, and that the relationships among them can be derived by classifying them into three different strata as pictured in figure 1 below;

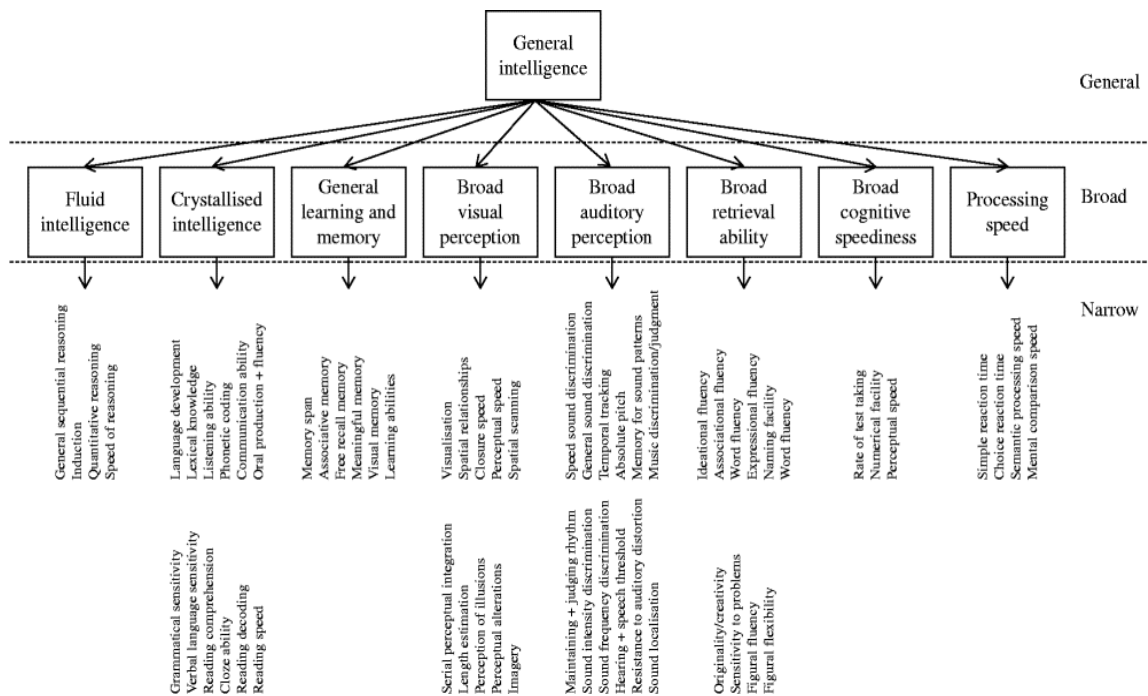


Figure 1: Carroll's three-stratum model of human intelligence (Pase & Stough, 2014).

According to Carroll (1997), stratum I represents 'narrow' abilities; stratum II- 'broad' abilities; and stratum III- a single 'general' ability. Stratum II is composed of eight broad abilities; Fluid intelligence (Gf), Crystallized intelligence (Gc), General memory and learning (Gy), Broad visual perception (Gv), Broad auditory perception (Gu), Broad retrieval ability (Gr), Broad cognitive speediness (Gs) and Processing speed (Gt), stratum I consists of 70 specific factors under each of the factors in stratum II, while stratum III reflects general intelligence, which accounts for the relationship among the broad abilities in stratum II (Carroll, 1997). A major contribution of the three-strata theory appears to be its pivotal role in the application of factor analysis in present day research in cognitive ability (Benisz, Dumont & Willis, 2015).

Going forward, Carroll's three-Stratum theory and Cattell-Horn's theory of fluid and crystalized intelligence have been merged, resulting in the Cattell-Horn-Carroll Theory

(CHC Theory; (Benisz et al., 2015; Flanagan, McGrew, & Ortiz, 2000). Unlike the two parent theories that started their conceptualization of intelligence from Spearman’s Gf, the CHC theory proposes a Two-Stratum model of intelligence; a broad stratum (II) and a narrow stratum (I). While stratum II consists of 10 broad factors of intelligence (Fluid intelligence (Gf), Quantitative knowledge (Gq), Crystallized intelligence (Gc); Reading and writing (Grw), Short-term memory (Gsm), Visual processing (Gv), Auditory processing (Ga), Long-term storage and retrieval (Glr), Processing speed (Gs), and Decision speed/reaction time (Gt), stratum I consists of specific abilities, as in Carroll’s Three—strata theory (Flanagan, Ortiz & Alfonso, 2007). Thus, a third stratum representing Spearman’s Gf is not emphasised in the CHC theory.

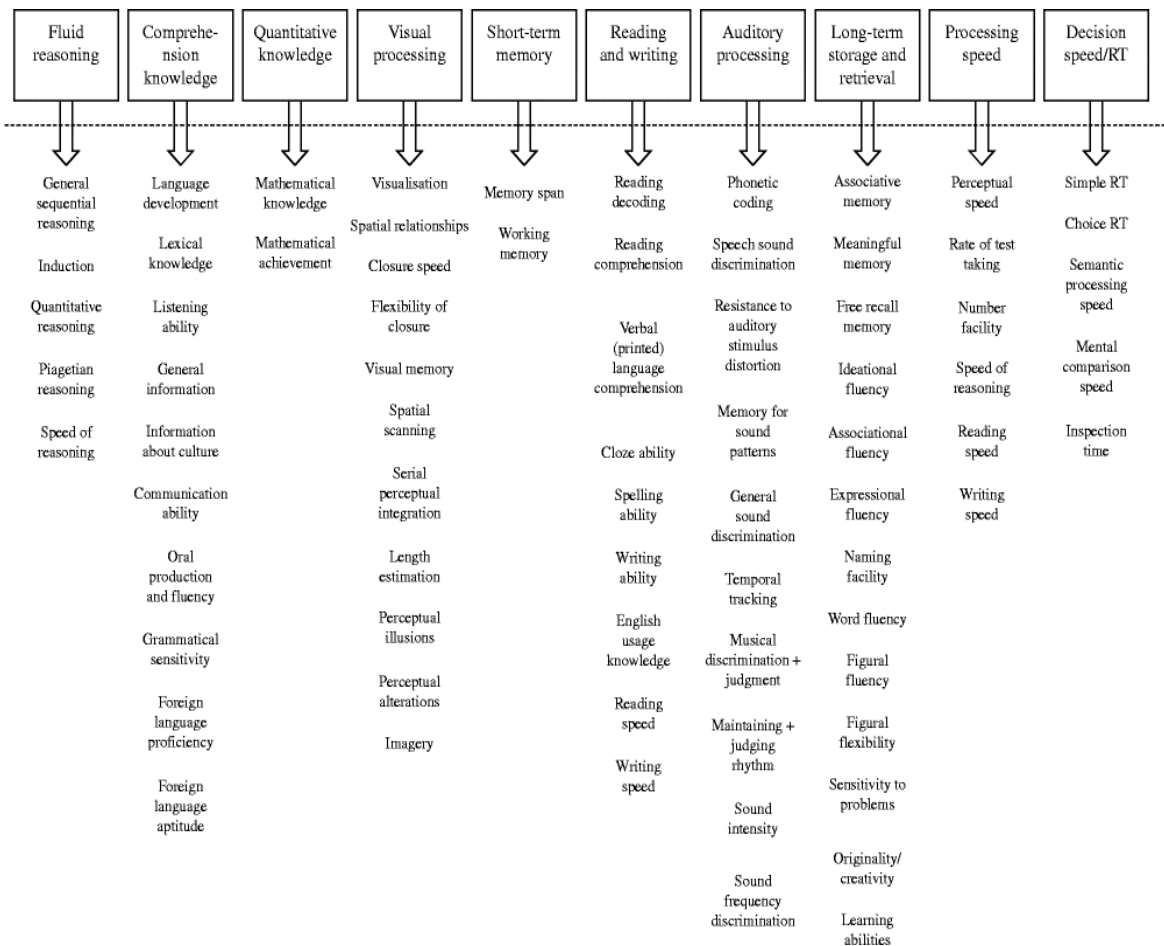


Figure 2: The Cattell–Horn–Carroll model (CHC; Pase & Stough, 2014)

Therefore, in line with the CHC theory, most contemporary IQ tests yield about four to seven cognitive ability scores, rather than previously emphasised single IQ score, notable of which are: the Stanford-Binet, fifth edition (Roid, 2003), the Kaufman Assessment

Battery for Children, second edition- KABC-II (Kaufman & Kaufman, 2004) and the third edition of the Woodcock-Johnson- WJ-III (Woodcock, McGrew, & Mather, 2001). Little wonder the CHC theory is considered the most advanced conceptualization of intelligence and most influential in the construction of contemporary IQ tests (Kaufman, Kaufman & Plucker, 2013).

2.2.2 Assessment of cognitive ability. The development of cognitive assessment measures is usually informed by the assumptions of intelligence theories. Although the suggestion of a theory and measurement of human ability is credited to Francis Galton (Thorndike, 1997), the first test of intelligence is credited to Alfred Binet and Theodore Simon (1904), who developed a test that measured intelligence, for the purpose of identifying learners who needed intervention. The test measured mental processes like comprehension, problem solving, analogical and logical reasoning and yielded a mental age score. The Binet-Simon test was not based on Spearman's *G_f*, rather they aimed at evaluating a number of intellectual abilities (Greenwood, 2015).

In 1914, William Stern came up with the notion of mental quotient- derived by dividing chronological age by performance on the Binet-Scale test (Stern 1914). Intelligence Quotient -IQ Score- as it is used today was however coined by Terman (1922) and, it is the product of a multiplication of mental age (Stern, 1914) by 100, such that the average IQ for any mental age is 100. Following this development, IQ scores and ranges began to be used as measures of intelligence and, although slight variations in terminology might be present across different tests, IQ ranges are generally the same (Greenwood, 2015). Following the Binet-Simon test, other measures of intelligence have been developed, however, the more contemporary Wechsler scales will be briefly discussed, given the vast research and popularity among psychologists.

The Wechsler Scales are prominent measures of cognitive ability, new editions of the test (Wechsler 2008; 2012; 2014) are derived from the original Wechsler-Bellevue Intelligence Scale (WBIS; Wechsler, 1931). The Wechsler scales continue to gain prominence because of its pivotal role in the paradigm shift in the way psychologists view intelligence test use, from a rather psychometric to a clinical view (Kaufman, 2009). The WBIS was developed in order to address the shortcoming Wechsler identified in previous tests, which is their

heavy loading on verbal items that resulted in discrepancy between scores on tests and real-life functioning of individuals tested (Benisz et al., 2015). To resolve this discrepancy therefore, the WBIS used standard scores (where the mean score is arbitrarily set at 100 with a standard deviation of 15), rather than the previously used ratio scores (Kaufman, 2009). The advantage of this being that the statistical significance of a score on the Wechsler test does not differ from year to year, since scores are quite constant, regardless of age. Besides, the use of deviation IQ was better able to evaluate the rate of cognitive development, which is not usually uniform with age (Benisz et al., 2015). The WBIS was constructed, guided by a combination of theories, current editions (Wechsler 2008; 2012; 2014) have however been updated, in line with the provisions of the CHC theory (Flanagan et al, 2013).

The most recent development in the process of intelligence testing however, is the shift from the use of a single test of intelligence, such as the Wechsler scale, to the adoption of the Cross-Battery Approach to assessment (referred to as XBA), first introduced in the 1990's by Dawn Flanagan, Samuel Ortiz and Kevin McGrew (Flanagan & McGrew, 2012; Flanagan, Ortiz, & Alfonso, 2013). The XBA is the process of including multiple IQ tests in an assessment, to garner a more comprehensive cognitive profile of the individual (Flanagan, Ortiz & Alfonso, 2007). The basis for the approach is for intelligence assessment to evaluate broader ability areas, since no single IQ test developed prior to 2000 was designed to assess all of the abilities postulated by the CHC theory (Woodcock, 1990), whereas, they are relevant for better understanding and prediction of cognitive functioning, The XBA approach to assessment is therefore considered a more theoretical and psychometrically valid method of evaluating cognitive ability (Flanagan, Ortiz & Alfonso, 2007).

2.2.3 Controversies around intelligence testing. Although intelligence testing has come a long way, with several refinements and advancements along the line, controversies continue to trail its use. This perhaps might be because of the nefarious ways in which it has been used to support prejudiced views. For instance, among the earliest use of intelligence testing was to identify “mental defectives”, which would inform legislative actions (Benisz et al., 2015). There have also been, and continue to be, instances where intelligence tests have been used to support racially prejudiced views (Brigham, 1923; Rushton & Jensen, 2005), whereas methodological reviews (Kamin, 2006; Ortiz, Ochoa &

Dynda, 2012) of such pursuits reveal malpractices. The resulting effect of these malpractices however, is the scepticism around the practice of cognitive testing, which borders on issues such as the validity of measures, group differences, cultural bias, amongst others.

With regards to the validity of measures, concerns have been raised that even though general cognitive ability plays a role in the different degrees of academic achievement (ranging from high to low), important life outcomes including income, educational attainment, job performance, social behaviours and others, other cognitive and non-cognitive factors also have a role to play (Heckman, 1995). Hence, IQ scores alone should not be considered in predicting future achievement levels.

Whether or not there is an actual distinction between tests of achievement and tests of ability (cognitive tests) is another issue of concern, as performance on tests of abilities are to some extent, impacted by accumulated knowledge (which achievement tests measures). This is because ability tests also measure the subject's achievements in the aspects of verbal, reading comprehension, arithmetic, amongst others, and involve some tasks that require knowledge of arithmetic, geometry etc. (Dickens, 2008). Hence, it appears that both achievement and ability tests have confounding elements.

Like the above, it is also argued that when tasks that are influenced by accumulated knowledge are included in measures of cognitive ability, it gives room for cultural bias, as some knowledges are readily available to people of a cultural background than others (Seifert, 2011). A typical instance of such cultural bias, which is usually a point of reference, is the about one standard deviation difference in the mean scores of American Whites and Blacks, in favour of Whites, on tests of cognitive ability (Brown, Reynolds & Whitaker, 1999). A counter argument to this criticism however, is that members of different racial groups who have similar scores on cognitive tests also have similar patterns of right and wrong responses. Whereas, if certain questions are culturally-biased as claimed, then the response pattern shouldn't be similar, in the sense that the biased group should find such questions more difficult than would the favoured group, resulting in dissimilar response patterns (Dickens, 2008).

In relation to the above counter submission, there is also the argument that differences in cognitive ability is more a function of genetics, rather than test composition (Neisser et al., 1996; Plomin et al. 2001; Rushton & Jensen, 2005). To support this claim, Rushton and Jensen (2005) reported on a number of studies conducted using different versions of the Raven's matrices [Advanced progressive matrices (APM), Standard Progressive Matrices (SPM) and Coloured Progressive Matrices (CPM)] and other cognitive tests (such as the WAIS), which according to them found an average IQ of 75 for Africans.

These claims have however been refuted by Kamin (2006), based on the identification of a number of methodological and ethical flaws in the manner in which the assessments were conducted and findings reported. Among the flaws noted was that test scores of less or uneducated Africans were compared to those of American norms, without consideration for the cultural differences (Fahrmeier, 1975; Glewwe & Jacoby, 1992) between the two groups nor for the poorer quality of education of the African groups (Skuy, Gewer, Osrin, Khunou, Fridjhon & Rushton, 2002). Therefore, the argument that cognitive test instruments (especially when conducted in the traditional manner most of them have been designed to be administered) lend credence to cultural-bias appears to still hold water.

These controversies notwithstanding, the benefits of IQ testing cannot be underestimated (Kaufman 2009), but will be appreciated when necessary caveats to its use are observed. MacCallum (2003, p. 113–115) captures this succinctly when he stated that "our models are implausible if taken as exact or literal representations of real world phenomena. They cannot capture the complexity of the real world which they purport to represent. At best, they can provide an approximation of the real world that has some substantive meaning and some utility". Therefore, with advancement in the research and practice of psychological assessment, as well as in recognition of the limitations in the prevailing approach to assessment, the dynamic method of assessment has been proposed, in the view that it offers the possibility of fairer assessment practices. This approach to psychological assessment, as well as the static approach are discussed in subsequent sections.

2.3 Paradigms of psychological assessment

Contemporary intelligence assessment measured however vary in their theoretical approach. In recent times, the hot debate is on static and dynamic approaches (Benisz et al.,

2015). These approaches are discussed subsequently. Within the practice of psychological assessment, static and dynamic methods have been adopted. Both methods are briefly outlined below.

2.3.1 Static assessment method. Static assessment methods, which aim to evaluate a child's ability, in relation to a norm group, have until recent times, dominated the practice of psychological assessment, particularly in the aspects of intelligence testing. Measures such as the Raven's matrices, Stanford-Binet Scales of Intelligence, Wechsler tests, are some of the mostly used instruments in this regard. Static assessment methods usually measure existing abilities of the individual - what has been learnt- without assistance to the student and therefore, give information only on crystallised abilities (Cho et al., 2015).

The static assessment method gained prominence and continues to be applied in the assessment of cognitive ability because of its alleged reliability and validity in assessing whether an individual currently possesses the proficiency required for certain tasks or not. Nonetheless, the use of a static method of assessment endures criticisms such as its inability to reveal the individual's future capability (Hamers & Resing, 1993; Murphy & Maree, 2006), yield information on whether cognitive deficits are intrinsic to the individual or are environmentally based (Cho et al., 2015), as well as problematic outcomes when assessment is done in a cross-cultural context, since the language proficiency and the educational and SES of test takers may negatively implicate the test results of learners from disadvantaged backgrounds (Van de Vijver, 2002).

One of such criticisms for instance is that static assessment methods, as exemplified by IQ tests, assume intelligence to be a stable or fixed characteristic within an individual, which is expected to develop in a predictable manner (Sattler, 2008). This implies that the approach makes uncontroversial, the expectations of children at different ages and stages, with the implication that children are identified and labelled as lacking in ability, with low expectations for their future ability (Lunt, 2001). Whereas there is evidence that children develop at a rather different pace and are therefore not expected to achieve milestones at the same age, but within a range of years (Vereijken, 2010).

Aside from these, the use of static assessment methods neither provide specific information about the individual learning process, about the cognitive functions that mitigate effective learning, nor do they inform strategies that are effective for remedial learning. This therefore results in findings that cannot be translated into practice nor used to improve teaching (Tzuriel, 2001). Thereby, creating a gap between psychologists and teachers.

Another argument in this regard is that static assessment methods are best used for classification and discriminatory labelling of individuals, based on their assigned level of intelligence, without a consideration of their untapped intelligence (Tzuriel, 2001). An instance of such discriminatory use of assessment can be seen in the report of Rushton and Jensen (2005), which claims that the IQ of Africans is 75, which they interpreted as “retarded”. This assessment process was heralded by inaccurate and unjustifiable methodology, as pointed out by Kamin (2006) and alluded to in 2.2.2.

Meanwhile, in recognition of the difference in linguistic, SES and ethnic backgrounds of children from minority groups, some states in the USA and Local Education Authorities (LEA) within the UK have put a hold on the use of IQ tests for minority groups, in the believe that the product and process of assessment for these group of children will most likely be unfair, perhaps discriminatory (Lunt, 2001). In the same vein, Murphy and Maree (2006, p. 169) opine that “if educators continue to assess learners based on static assessments, such as conventional intelligence quotient (IQ) tests, and from this infer who will receive special treatment (and in so doing stigmatize learners), some learners might be put at a disadvantage in terms of not being able to develop their full potential”.

These drawbacks of the static assessment method form part of the basis for the paradigm shift from static to a more dynamic assessment method.

2.3.2 Dynamic assessment method. The feasibility of making assessment procedures more adaptive to the performance level of the individual being assessed, via the incorporation of learning opportunities, has been an interesting area of investigation in recent years. This has brought about the introduction and thriving of the concept of Dynamic Assessment (DA), which has evoked a lot of debate across the globe. DA is a psychological approach to assessment which adopts a test-train-test pattern of assessment

(Cohen & Swerdlik, 2010). This pattern is usually followed to give the test taker an opportunity to ‘learn on the job’, reduce all forms of unfairness and place all test takers on an equal level. DA measures learning potential, rather than current level of intelligence, through the inclusion of a training within the assessment process (Murphy, 2002).

In other words, it aims at evaluating the potential of the individual to acquire new knowledge and improve his/her current performance level, in addition to evaluating the current level of functioning (De Beer, 2007). In summary therefore, the goal of DA is to provide information on the current abilities of an individual, define cognitive functions, identify cognitive deficiencies, describe change process and provide information about mediational strategies that promote cognitive modifiability (Tzuriel, 2001). Accordingly, the basic ideas that underpin DA are described by Haywood and Tzuriel (2002, p. 41-42) below;

- “Accumulated knowledge is not the best indication of one's ability to acquire new knowledge, although the two are highly correlated.
- Everybody functions at considerably less than 100% of full capacity; therefore, everybody can do better.
- The best test of any performance is a sample of that performance itself, therefore, assessment of learning abilities can be accomplished effectively with the use of learning tasks, especially those involving teaching-a condition that characterizes school learning.
- There are identifiable obstacles to one's access to and effective application of one's intelligence. Such obstacles include ignorance; impulsivity; impoverished vocabulary; cultural differences in learning habits, styles, and attitudes; poor self-concept as learners; and a host of motivational variables; plus, of course, inadequate development of important cognitive and metacognitive structures and strategies. By removing some of those obstacles, one can reveal the ability to function more adequately”.

The difference between static and dynamic assessment methods is twofold. First is in the difference in test situation, in the sense that while static assessment discourages interaction between the assessee and the assessor, to allow for near perfect creation of a similar test situation and standardization procedure, DA views assessment as a collaborative activity. Hence the inclusion of a training aspect within the assessment process. The second difference is in the nature and exploration of the construct investigated, goals, instruments,

test situation, process and interpretation of scores (Lunt, 2001). These differences nonetheless, proponents of DA warn against replacing traditional assessment with DA, but to rather use it as an add on, since both yield different kinds of information (Caffrey et al., 2008; Elliot, 2000a, 2000b; Haywood & Lidz, 2005). To put this into perspective, Elliot (2000a, 2000b) warns that psychologists should be wary of viewing DA as basically a superior tool of similar use as IQ tests (for the purpose of selection, classification, placement, prediction of future outcome), arguing that the implication will be that “a paradigmatic shift will not have occurred and the true potential of dynamic assessment will not be revealed” (Elliot 2000b p. 734). Rather, the goal should be a paradigm shift wherein Dynamic methods will be used by teachers and psychologists, in a collaborative manner, for the design of classroom – based interventions (Eliot, 2003).

Stressing the benefits of DA, proponents claim that it takes into cognisance the fact that test takers are not homogenous in the level of prior opportunities they might have had, which might impact on their current performance level (De Beer, 2007; Murphy, 2002). Therefore, by creating a training process within the test process, the potential of the learner is assessed, in addition to their current levels of performance, thereby bridging the gap observed with the use of traditional measures of cognitive ability. In addition, as it is impossible to determine whether poor performance on tests is based on environmental disadvantage or inherent disability, when assessment is completed in the static manner, focus on the process as in DA and not the product, provides a clearer picture of learning potential (Elleman, Compton, Fuchs, Fuchs, & Bouton, 2011). Aside these, DA gives room for identifying the difference between low performances based on cultural difference, it uncovers limitations to performance and prescribes unique remedy for deficiencies and ways to enhance learning process (Amod, Heafield & Seabi, 2018; Tzuriel, 2001).

It is also argued that when tests of abilities are administered in the static manner, they only indicate crystallised knowledge and current abilities, which is of no value to the teacher’s instructional process (Grigorenko, 2009). Whereas the real ability of such children (i.e. learning potential) is better identified when tests are administered using DA approach, since the main goal of DA is to capture the future potential of the child, if learning is tailored to their needs, (Sternberg et al., 2002) and at the same time, provide data that is usable in intervention planning (Amod, Heafield & Seabi, 2018).

2.3.2.1 Criticisms of dynamic assessment. Despite the values of the DA method, a number of concerns have been raised regarding its use, which include: the longer time it takes to administer a test, compared to when assessment is conducted using the static method; standardisation issues which limit generalization; the high cost of training needed for the examiner; as well as limited published empirical psychometric evidence (De Beer, 2005; Grigorenko & Sternberg, 1998; Haywood & Tzuriel, 2002; Karpov & Tzuriel, 2009).

Reliability and validity issues

A major criticism usually raised against DA is the lack of empirical evidence on its reliability and validity (Caffrey et al., 2008; Elliot, 2000a; Haywood & Tzuriel 2002; Lidz, 2009). While reliability implies consistency with which a test measures the same construct over time, validity refers to the extent to which a test measures what is intended for (Merrell, Ervin, & Peacock, 2011). A lack of evidence in these regard is considered problematic given the heavy reliance on empiricism, within the natural science, including concepts of reliability and validity (Merrell, Ervin & Peacock, 2011). However, some DA proponents argue that validity and reliability concepts shouldn't be used in DA, given its theoretical underpinning, which differs from that of static approach (Caffrey et al., 2008; Murphy & Maree, 2006; Poehner, 2010). The argument is that since DA considers cognitive ability as modifiable, prone to change and not static, evaluating such things as reliability and consistency therefore, have no place in in the theory of DA. Besides, evaluating test validity requires comparison with another measure of same construct, whereas, given the difference in constructs between DA and static measures, establishing reliability and validity with same rigour as static assessment may be of no added value (Poehner, 2010). It is on this premise that Poehner (2010) suggested that rather than modifying DA such that it is amenable to psychometric models of evaluation, new criteria for evaluating reliability and validity could be developed, such as in-depth case study analysis, detailed description of the measures.

Notwithstanding this viewpoint, given that item scores of individuals with same ability but of different groups need to be amenable to comparability (Differential Item Functioning), so as to eliminate bias and increase test validity (Kanjee, 2001), item response theory has been introduced within the DA framework, to investigate DIF, thereby addressing the issue of test validity and reliability (Murphy & Maree, 2006).

Time constraints, labour intensiveness and cost

Another issue that has been identified with the use of DA instruments is that it involves more administration time, skill, experience and more investment of effort, all of which are not cost effective, compared to static assessment. DA requires greater skill and effort in the sense that the practitioner needs to evaluate not only cognitive ability, but also identify deficient cognitive functions, determine beneficial mediation processes to implement, understand the cognitive process to recommend relevant remedial strategies, all of which require time and know-how. Whereas, most graduate psychology programmes do not include this training in their curriculum, hence, practitioners need to undergo further intensive training after qualifying (Haywood Tzuriel, 2002). Aside this, in an international survey of trainers in DA, Haywood and Lidz (2005) discovered that in addition to the challenge of training in DA, the need to help assessors “unlearn” the traditional approaches they have been used to, was another challenge. They therefore concluded that “it would be important to help new learners develop DA technique early, so that there is no need to “unlearn” (Haywood & Lidz, 2005, pp. 194-195).

In addition, DA proponents argue that the solution to the issue of cost lies in identifying the use to which results are intended. The claim is that when the intent is to evaluate current functioning, static assessment might be relevant, but when the goal is to uncover learning potential, evaluate the learning process to inform prescription of learning strategies, then DA is more reliable. Tzuriel (2002) puts this more succinctly in the following words:

A short, “instant” assessment might be cheaper in the short run but superficial, wasteful, and less effective in the long run. DA, on the other hand, is lengthier and more expensive but provides in-depth and qualitatively better results, which ensure accurate future intervention procedures. Psychologists, educators, and policy makers should be convinced first that the information derived from DA is worth the investment required to get it, and that the information achieved will then be used in a way that will have an impact on specific learning strategies and academic achievements” (p. 425).

It therefore appears that to combat the challenge of cost, the foremost task should be an identification of the goal of assessment.

Generalization of benefits

Another issue with regards to the process used within DA is the extent to which benefits from the mediation phase can be generalized to other contexts (Karpov & Tzuriel, 2009).

This raises not only theoretical questions but has implication on test design and mediation process. However, based on evidence from a number of studies conducted in Bar Ilan University laboratory, Haywood and Tzuriel (2002) claim that metacognitive skills develop with age and “had a consolidating effect by “bridging” cognitive modifiability across different domains” (p. 426). Therefore, aiding the transfer of what is learnt in one domain to other domains.

Construct fuzziness

Construct fuzziness is another often mentioned criticism of DA (Caffrey et al., 2008; Karpov & Tzuriel, 09). Construct fuzziness implies that the theoretical underpinnings, methods, goals and procedure of DA are characterized by a lot of difference and overlap (Karpov & Tzuriel, 2009), which brings about lack of clarity in the methods and objectives (Karpov & Tzuriel, 2009; Kozulin, 2011). This may be due to a lack of concise definition of DA, its goals and procedures (Karpov & Tzuriel, 2009), which results in a broad interpretation of DA and an overlap of terms. For instance, while Kozulin mentions a difference between DA and learning potential assessment, Lidz (p.238, 2009) opines that “all learning potential assessment is dynamic assessment, whereas not all dynamic assessment is learning potential”. Also, while Poehner (2010) opines that DA and dynamic testing should not be viewed as separate terms, Sternberg and Grigorenko (2002) allude to the difference between dynamic testing and DA, claiming that dynamic testing is a subset within DA. Accordingly, Poehner (2010) posits that this kind of fuzziness in construct can lead to confusion. As it stands therefore, it appears that research that focuses on evaluating differences in DA concepts might be needed to address the issue of construct fuzziness, just as Kozulin (2011) researched learning potential and cognitive modifiability and found both constructs to be distinct concepts.

Despite these criticisms, DA continues to gain prominence and scores of DA measures, assessing learning potential, have been developed locally, these include; the Learning Potential Computerised Adaptive Test (LPCAT) (De Beer, 2000a), the Ability, Processing of Information and Learning Potential Battery (APIL) (Taylor, 2006), the Transfer, Automatization, Memory and Understanding Learning Potential Battery (TRAM) (Taylor, 1999) among others. Only the LPCAT will be described here, as it is the focus of this study.

2.4 The learning potential computerised adaptive test (LPCAT)

2.4.1 Description of the LPCAT. The LPCAT is a dynamic measure of learning potential, developed against the framework of Vygotsky's theory of the Zone of Proximal Development (ZPD) (1978), which is also credited as the major theory around which the assessment of learning potential and dynamic assessment are woven. The LPCAT attempts to assess learning potential in the fluid reasoning domain of cognitive functioning. It utilizes non-verbal figural test items, in a test-train-retest format, so that the influence of language proficiency, SES and prior learning experiences on test results are significantly reduced (De Beer, 2006).

According to De Beer (2005), one of the reasons for assessing learning potential, rather than intelligence, is that there is evidence which indicates that IQ scores are not static but tend to change when there is an increase in the socio-economic level, educational opportunity and language proficiency of the individual. De Beer (2005) therefore remarks that it would be better not to confine individuals to an IQ score which might have been affected by their being disadvantaged, whereas they could have a higher potential to learn. As such, the LPCAT is designed to evaluate the general level of reasoning of the individual as well as the level at which it could be developed, given access to training.

The use of nonverbal-figural items in the construction of the LPCAT makes it amenable to comparison with other static measures of intelligence that have been similarly constructed, such as the Raven's Standard Progressive Matrices (SPM). For this reason, this study will attempt to establish the construct validity of the LPCAT, through comparison with the Raven's Standard Progressive Matrices (SPM), subsequently described.

2.4.2 Review of literature on the LPCAT. Following the development of the LPCAT, there has been several empirical studies to validate its use within educational and work settings in South Africa, most of which report positive findings. In a longitudinal study to evaluate its predictive validity, De Beer (2010) compared the LPCAT to two other static measures (English Proficiency Test and certain subscales of a standardised Aptitude Test) and reported that the LPCAT better predicted academic performance. This is corroborated by findings from another study (Schaap & Luwes, 2013), which also reports a positive

correlation between the LPCAT and academic performance, although the study also noted a difference in mean scores based on race, whereas another measure used in the study (Academic Aptitude Test-Math) generated equal scores for all the races. In line with this, the LPCAT was also reported to be a valid predictor of work performance, but also with significant differences among gender and race groupings (Mphokane, 2014).

It is important to also note that the study by Schaap and Luwes (2013) reports that the LPCAT made no significant contribution in predicting academic performance, when a regression analysis was done with the inclusion of two other verbal measures of proficiency and skill acquisition (AAT-Math and English Language Proficiency Test (ELSA)). In addition, Lohman (2005) contends that the incremental validity of figural reasoning or non-verbal tests is low when they are used in combination with verbal and quantitative reasoning tests. Perhaps the study conducted by Van de Merwe (2006) on the predictive validity of the LPCAT among university students may provide a better indication of comparability of measures. The study found that in addition to positively correlating with academic performance, the LPCAT also had a higher predictive value of academic performance, compared to the Potential Index Battery which is also a test of potential, but this time, potential to succeed in the workplace.

Aside these, De Beer (2003) investigated the criterion-related validity of the LPCAT, for different groups of students, across educational levels. Results indicate that the LPCAT scores correlate with that of the criterion measure, although, there was a significantly lower correlation of scores at the university level, compared to that at the high school level.

Thus, although most studies lend credence to the validity of the LPCAT, they also trigger some questions regarding the culture-fairness of the LPCAT, when the mean scores of different racial groups and gender are compared. Findings from the current study will hopefully add to the body of evidence in this regard.

2.5 The raven's standard progressive matrices (SPM)

2.5.1 Overview of the SPM. The Raven's Progressive matrices (RPM) (Raven, 1976) is a static measure of fluid ability and was constructed based on Spearman's (1904) g factor of intelligence. The RPM is available in three forms; the Standard Progressive

Matrices (SPM), the Coloured Progressive Matrices (CPM) and the Advanced Progressive Matrices (APM), and they can be administered to people of all ages, young children and clinical populations, and adults of average or more than average intellectual abilities respectively. For the purpose of this study, the SPM was adopted as it is more appropriate for use with people of all ages.

The SPM was published in 1938 and was developed based on Spearman's (1904) use of geometric figures in the assessment of modes of thinking, as well as his investigations on the nature of intelligence. The SPM consists of 60 items, grouped as five sets, each of which has 12 items, it starts with a relatively easy question but becomes progressively more difficult.

The SPM was designed to evaluate an individual's general range of ability, particularly, the current ability to perceive and think clearly, past experiences or current verbal abilities, notwithstanding. This makes it a measure of fluid general intelligence, according to Cattell's (1987) categorisation.

According to Raven (1976), the SPM, on its own, does not produce a general IQ score but it is to be used to complement other vocabulary measures. The SPM is regarded as a culture-fair measure because it provides an index of intellectual capacity that is not influenced by cultural background (although this is contested, as subsequent section indicates), which perhaps accounts for its wide usage across the globe.

2.5.2 Review of literature on the SPM. Over and above the LPCAT, the Raven's SPM has been widely researched, locally and internationally. Among several others, findings indicate a general construct validity for the SPM, meaning that it is a good measure of *g* (Owen, 1992, Pind, Gunnarsdóttir & Jóhannesson, 2003), as well as internal consistency and factorial validity (Abdel-Khalek, 2005). However, a large mean difference has been observed in the scores of Black and White South African pupils (Owen, 1992) and university students (Rushton et al., 2004). In addition, a UK study also observed a five-point IQ difference between the scores of males and females, aged 15 and above, in favour of males, although the same study also reported no observed difference between the ages of 6-14 (Lynn & Irwing, 2004). Aside from these, another study (Gunnarsdóttir & Jóhannesson,

2003) reported a positive criterion validity of the SPM, when it was compared to scholastic achievement tests.

As indicated above, there is an ongoing debate around the extent of culture-fairness of the SPM and Raven's Matrixes in general, since some studies have documented a difference in mean scores based on race, age, socio-economic and cultural background. For instance, in a study conducted on the performance of African and White South African university students on the SPM, Rushton et al. (2004) reported a difference in the g factor, with an average score at the 14th and 61st percentile respectively. Owen (1992) also found a difference in mean scores among the racial groups in South Africa, precisely, a -.52 difference was found between White and Indian, -1.35 between White and Coloured and a -2.78 between White and African.

In addition, Lynn (2002) argues that there is a difference in test scores based on sex, with claims of a male advantage. This position is also supported by findings from other meta-analyses (Irwing & Lynn, 2005; Lynn & Irwing, 2004), as well as by other studies that found a difference of 5.7 points between men and women (Sellami, Infanzón, Lanzón, Díaz & Lynn, 2010) and 4.4 points among 17-year-old males and females (Spanoudis, Natsopoulos, & Lynn, 2016). Iftikhar (2012) also found a gender difference in scores of 8 – 11-year-old primary school kids, but this time, girls performed higher than boys. In addition, Páchová (2013) queries the use of the SPM as a culture-fair test as he observed that the socio-cultural background of test takers played an important role in their performance. Khaleefa, Amer and Lynn (2014) also reported difference in IQ scores of university students, based on departments. While engineering and medical students scored higher, students from primary education department scored lower.

2.6 Review of literature on the LPCAT and the SPM

There are no published studies on the relationship between the LPCAT and the SPM. As both tests claim to measure the same construct (fluid intelligence) it will be interesting to investigate the relationship between them. In addition, the call for more empirical studies on the validity of dynamic assessments (Caffrey et al., 2008; De Beer, 2010; Guthke & Stein, 1996; Lidz, 2009; Murphy, 2002; Murphy & Maree, 2006) makes it equally important to conduct a study in this regard.

The study therefore evaluated the construct validity of the LPCAT by comparing test scores on it to that of the SPM. This is significant as it will not only add to the body of empirical evidence available on the subject matter, it will also assist psychologists in practice in making more informed decisions as to adopting dynamic assessment methods.

2.7 Theoretical Framework

2.7.1 Vygotsky's theory of socio-cultural development. According to Lev Vygotsky (1978), learning and thinking are functions of formal instruction, culture and social interactions (Rowe & Wertsch, 2002), meaning that an individual's thought patterns will always reflect his/her cultural and social practices. This opinion formed the basis of the development of Vygotsky's theory of socio-cultural development.

Vygotsky further postulates that within cultures, there are cognitive devices and procedures (such as rules, writings, gestures, numbers etc.) through which people relate with the world around them and which shapes their mental functioning (Grader & Shields, 2004). According to Vygotsky, these cognitive devices and procedures, which he terms "psychological tools", are introduced to children through social interactions with their parents and formal interactions at school and later on become internalized and automatically form the basis of the child's cognitive processes (Snowman & McCown, 2013).

Furthermore, unlike most cognitive development theorists, such as Piaget, who assume that social development is aided by cognitive development, Vygotsky believed the opposite is the case, meaning that cognitive development is primarily aided and influenced by social interaction (Snowman & McCown, 2013). Hence, he believed that the rate of cognitive development differs among cultures, forming the basis for the study's exploration of the effect of demographic variables on the LPCAT scores.

Vygotsky is also credited with the concept of the Zone of Proximal Development (ZPD), which is described as the difference between the level of performance that an individual can attain without help and the level that can be attained with help from a more knowledgeable other (Vygotsky, 1978). The ZPD according to Vygotsky, houses the thinking patterns and abilities that are in the process of maturing, therefore, when assistance is offered herein,

cognitive development is enhanced (Snowman & McCown, 2013). Vygotsky (1978) therefore opined that rather than basing intelligence on an individual's current knowledge, it is better to evaluate the ability to solve problems individually and with help.

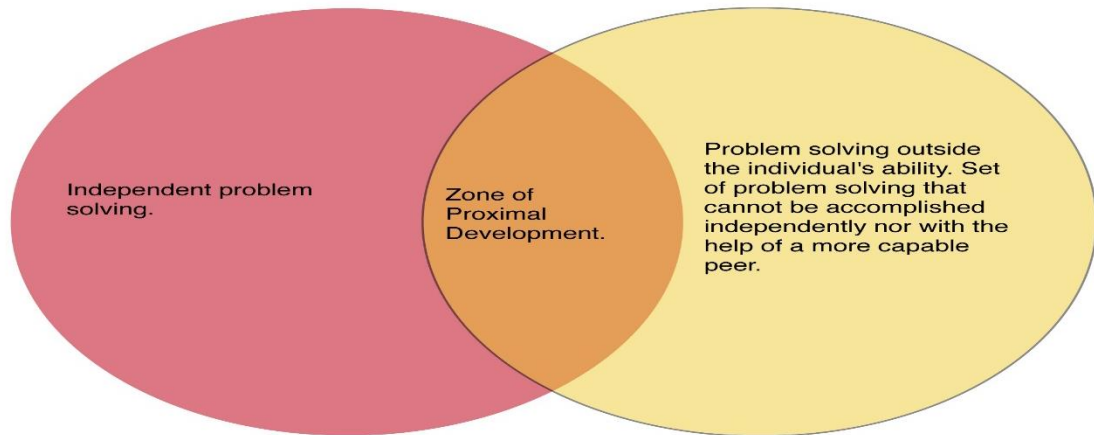


Figure 3. The Zone of Proximal Development (Campbell, 2008).

Vygotsky (1978) further posited that at least two developmental levels must be determined to ascertain an individual's level of mental development; the Actual Level of Development (ALD) and the Zonal Level of Proximal Development. At the first level (ALD) are the mental functions that are intact, following completion of developmental cycles, and it is that which is derived from testing (mental age), while the second level (ZPD) is the distance between the ADL and the potential development of a child. The ZPD accounts for abilities that are still in the process of developing. Vygotsky (1978) therefore posits that determining mental ages (ALD) only, rather than the ZPD would only give a summary of development, whereas the ZPD can play an important role in enhancing the effective diagnosis of educational problems, since information will be gathered, not only on how much learning has taking place up to the point in time, but also about how much can be learnt when relevant learning opportunities are provided.

According to Vygotsky's theory therefore, the ZPD should be considered in the measurement of cognitive ability because;

we can take account not only of the cycles and maturation processes that have already been completed but also those processes that are currently in a state of formation, that are just beginning to mature and develop . . . allowing not only for what already has been achieved developmentally but also for what is in the course of maturing (Vygotsky 1978, p.87).

This concept of measuring the individual's ZPD, in addition to the ALD, therefore forms the theoretical underpinning of the DA method and the current research, since it claims to attempt to guide the individual through their ZPD and evaluate their potential, by including a training within the assessment process; the test-train-retest approach (Murphy, 2002). The test-train-retest approach of DA also seems to be in line with Vygotsky's theory of socio-cultural development, as dynamic assessments not only measure the individual's ALD, which could have been influenced by their disadvantaged educational and socio-cultural backgrounds, but also their potential for learning, which is less influenced by previous learning experiences.

Relating the foregoing to the LPCAT, the pre-test aspect therefore signifies the ALD of the individual, as it provides information on the current level of development, while the post-test gives information on the potential for learning, after which training has been given. Therefore, accounting for Vygotsky's ZPD.

2.7.2 Some criticisms against Vygotsky's zone of proximal development. Some issues have been identified as limitations of Vygotsky's concept of the ZPD. The first major criticism is that the identification or evaluation of an individual's ZPD alone doesn't give information about the individual's learning ability, learning style and process nor current developmental level, compared to other individuals of the same age and motivational levels. Thus, only a limited developmental picture of the individual is derived (Chaiklin, 2003).

There is also the argument that there is a lack of a common metric scale to measure an individual's ZPD. Although Vygotsky measured using age at times, critics argue that this is too general a metric as differences between different ages cannot be considered statistically equal nor generalised (Chaiklin, 2003). Another issue of concern is the generalizability and stability or otherwise, of an individual's ZPD. This is in the sense that it is difficult to determine if an individual has the same ZPD across all domains, if the size of the ZPD changes or whether the improvement when help is given is short or long-termed (Chaiklin, 2003).

If these criticisms are related to dynamic assessment, then a major concern would be whether the improvement in the individual's test score, following training or help, can be

generalised to other ability areas that the current test doesn't measure or whether the improvement is limited to only the current ability tested.

2.8 Conclusion

This chapter contains an overview of what psychological assessment is, as well as its applications and ethical issues. The chapter also focused on cognitive assessment, with a brief highlight on some of the current issues around cognitive testing, methods of investigating construct validity, theories of intelligence, controversies around intelligence testing and the prevailing paradigms of assessment, including criticisms. A brief description of the two measures used in the study; the LPCAT and SPM is also contained in the chapter, with a review of some of the studies that have been carried out on both and questions that such findings raise, which the current study might also help to answer. The chapter concluded with an overview of the theoretical underpinning of dynamic assessment method, including some of the criticisms that have been raised against the theory.

CHAPTER THREE

RESEARCH METHODOLOGY

3.0 Introduction

This chapter details the design of the study, the process adopted for data collection and a description of how data collected was analysed to arrive at a conclusion on the set hypothesis. The chapter also covers a discussion on how validity and reliability were ensured in the study and how ethical issues regarding data collection were addressed. As with all research work, the current study also had some limitations, these are also discussed in this chapter.

3.1 Research aim and hypotheses

The main purpose for which this research was conducted is to evaluate the construct validity of the LPCAT, by comparing scores on it to those on the SPM. University undergraduates from the University of KwaZulu – Natal made up the sample for the study. In addition to the primary aim, the following secondary aims were also intended for this study:

1. To determine if selected demographic variables have a significant effect on students' LPCAT pre-test scores.
2. To determine if selected demographic variables have a significant effect on students' LPCAT post-test scores.
3. To determine if selected demographic variables have a significant effect on students' LPCAT composite scores.

It is anticipated that findings from this study will not only provide evidence on the construct validity of the LPCAT but will also add to the body of empirical evidence on the LPCAT and dynamic approach to assessment as a whole, to give practitioners aligned with the approach more confidence in its use.

3.1.1 Research questions and hypotheses. The following questions are intended to be answered by this research, by testing the corresponding hypotheses:

Main question and hypothesis

Is there a statistically significant relationship between undergraduate students' performance on the LPCAT and the SPM? Consequently, this will be achieved by testing the following null and alternate hypothesis:

- **H₀:** There is no statistically significant relationship between students' performance on the LPCAT and the SPM
- **H_a:** There is a statistically significant relationship between students' performance on the LPCAT and the SPM.

Secondary research questions and hypotheses:

1. Do selected demographic variables have a statistically significant effect on students' LPCAT pre-test scores?

H₀: There is no statistically significant relationship between selected demographic variables and students' LPCAT pre-test scores.

H_a: There is a statistically significant relationship between selected demographic variables and students' LPCAT pre-test scores.

2. Do selected demographic variables have a statistically significant effect on students' LPCAT post- test scores?

H₀: There is no statistically significant relationship between selected demographic variables and students' LPCAT post-test scores.

H_a: There is a statistically significant relationship between selected demographic variables and students' LPCAT post-test scores.

3. Do selected demographic variables have a statistically significant effect on students' LPCAT Composite scores?

H₀: There is no statistically significant relationship between selected demographic variables and students' LPCAT composite scores.

H_a: There is a statistically significant relationship between selected demographic variables and students' LPCAT composite scores.

3.2 Research design

The design of a research work entails a detailed description of the general outline of a study, including the epistemological grounding, the approach to the study, sample and sampling methods, data collection, amongst others, aimed at gathering valid evidence to answer the research questions (McMillan & Schumacher, 2001). Given this definition, details of the research design are provided below;

3.2.1 Research paradigm. Filstead (1979, p. 34) defines a paradigm as “a set of interrelated assumptions about the social world which provides a philosophical and conceptual framework for the organized study of that world”. In other words, the research paradigm is the researcher’s worldview, which guides the investigation (Guba & Lincoln, 1994). This study was viewed within the lens of the positivism paradigm, which holds the assumption that there is only one true reality which is “apprehendable, identifiable and measurable” (Ponterotto, 2005, p. 130), that this truth can be uncovered objectively through a rigorous and standard procedure, by an unbiased and uninvolved researcher (Ponterotto, 2005). Since the paradigmatic assumption of a study determines the nature of the phenomenon under study, as well as influences and reflects in the general design and methodology of the research (Denzin & Lincoln, 2000), the general design and methodology of this research was therefore influenced by the assumptions of the positivist paradigm.

3.2.2 Quantitative research approach. According to Denzin and Lincoln (2000) and Creswell (2012), a quantitative approach is used when the aim of the research is to describe a phenomenon, and patterns of relationship among variables in the study. A quantitative approach is also appropriate for the exploration of the possible relationships among variables and set of scores, as well as in identifying and measuring the degree of such relationship (Bickman & Rog, 2008). Since the purpose of the current study is to evaluate the degree of relationship between performance on the LPCAT and the SPM, as well as the degree of relationship between the LPCAT and selected variables, a quantitative approach was adopted. Also, in line with the quantitative approach, the research was designed to be a correlational study, to determine the pattern of variance among the variables and set of scores (Creswell, 2012).

3.2.3 Research location. The study was conducted within the Pietermaritzburg campus of the University of KwaZulu – Natal. This campus is home to four Colleges; Science and Agriculture, Education, Law, Human and Management Sciences. The research

sample was intended to be drawn across the four faculties, however, given that the convenience sampling method was adopted, the study respondents were mainly from the faculties of Humanities, Law and Science and Agriculture.

3.2.4 Sample and sampling procedures. The sample of a research study comprises a selected group of people deemed representative of the larger population a study aims to cover (Mitchell & Jolley, 2007). The sample for this study was initially intended to be first year students only, however, following difficulties with data collection due to protest action on the study location, the sample was expanded and therefore included undergraduate students currently studying in any of the Colleges on the Pietermaritzburg campus of the University of KwaZulu-Natal.

Based on the central limit theorem which predicts an increase in the likelihood of a normal distribution of sample as the sample size increases (Howell, 2012), the sample size for the study was initially set within the range of 100 to 200. However, due to the nationwide university student's protest action that lasted for a better part of the data collection period, thus affecting student's attendance at school and recruitment of participants, the actual sample size was 95. This size was still considered appropriate given that although 'as sample size increases beyond $n = 1$, the sample becomes a more accurate representative of the population, and the standard error decreases" (Gravetter & Wallnau, 2012, p.208), researchers can significantly reduce error by increasing sample size to about $n = 30$, since a sample size greater than $n=30$ does not add much to the generalizability of the sample (Gravetter & Wallnau, 2012).

To select the sample, a non-probability convenience sampling method was adopted, which according to Creswell (2012), implies using a group of people who are willing, accessible and possess characteristics needed for the study. In line with this, a recruitment poster (Appendix D) was designed and pasted on notice boards across the four campuses earlier mentioned.

3.2.5 Data collection instruments. Data for this study was collected using the LPCAT and the SPM. A demographic details sheet (Appendix E) was designed by the researcher, to gather demographic details of the participants. A brief description of the design of both measures is given below:

Learning Potential Computerised Adaptive Test (LPCAT)

As stated in previous chapters, the LPCAT is a non-verbal figural dynamic measure of learning potential that was developed specifically for South Africa and normed with the multicultural characteristics of the context in mind (De Beer, 2005). The test assesses fluid intelligence and does not require language proficiency nor scholastic background to be able to complete it, which accounts for the non-verbal format adopted (De Beer, 2007). The LPCAT is a computerised assessment test and takes about an hour to complete. Test items are structured in the form of matrices and the test taker needs to select a missing pattern, from four options, to complete a series (De Beer, 2005; 2007). The test is structured to consist of two linked adaptive tests (pre-test and post-test), with a training session between them. Hence, the LPCAT produces a pre-test score (present level of performance), a post-test score (potential for learning), a difference score (undeveloped potential) and a composite score (a combination of the pre-test, post-test and difference scores), which makes it a useful measure of an individual's fluid reasoning level as well as the level at which they can be developed, given necessary training. In addition to being interpreted quantitatively, results on the LPCAT can also be interpreted qualitatively as the visual presentation of results provide qualitative information on the individual's performance levels throughout the test administration process (De Beer, 2007).

Raven's Standard Progressive Matrices (SPM):

Developed by John C. Raven (1936) in the United Kingdom (UK), the SPM is a well-researched and commonly used non-verbal measure of fluid intelligence which can be administered to individuals or groups between the ages of five and beyond. It was designed to be culture-fair as it requires little verbal instruction (Raven, 2000), accounting for its wide use in South Africa. The SPM comprises 60 multiple choice items arranged in order of difficulty and takes about 30 to 40 minutes to complete (Raven, 2000).

The SPM is administered as a paper and pencil test where the subject is asked to select the missing item to complete a group of patterns which are presented in 6x6, 4x4, 3x3 or 3x2 matrices. All items in the test are presented in black ink on a white background.

3.2.6 Procedure.

3.2.6.1 Ethical approval/ recruitment of participants. An application for ethical

clearance was submitted to the Human and Social Sciences Research Ethics Committee (HSSREC) of UKZN. After this was granted (Protocol Reference Number: HSS/0577/016M; both initial and subsequent approval copies are contained in Appendix A & B) and following the registrars' permission (Appendix C) announcements were made at various undergraduate lecture rooms (with the consent of the lecturers in charge), where the research aim and procedure was explained to the students. In addition, recruitment posters (Appendix D) were placed on noticeboards across the campuses to describe the study and invite students to participate. Volunteer participants were requested to report for assessment at set dates and times.

3.2.6.2 Test administration. Prior to the commencement of assessments, the researcher attended a training on the LPCAT, as per the test requirement. No training was attended on the SPM as this was not required, however, the test was administered according to the instructions contained in the manual. Assessments were conducted at the PsycLab (a psychology laboratory within UKZN Pietermaritzburg campus Psychology building), which was chosen because it is equipped with computers which can be used to administer the LPCAT and because it offered increased confidentiality for test takers.

On the assessment date(s), the aim of the study was re-explained to participants. Their voluntary participation was ensured by providing them with a detailed information sheet, while their informed consent was documented by them signing an informed consent form (Appendix F). Test administration was conducted by the researcher (who was a registered student psychologist), with the help of a trained research assistant, and under the supervision of the research supervisors (registered psychologists). Assessment was commenced as the participants arrived and no more than 20 participants were assessed at any given time. The LPCAT and SPM were administered to participants in no particular order, as participants could choose which they preferred to complete first. The LPCAT was completed individually, on computers which had been pre-loaded with the application, while each participant was handed test booklets and response sheets, for the completion of the SPM. Each participant took between 1.5 hours to 2 hours to complete both tests and was given a R30 meal voucher at the end of assessment, to compensate for time.

At the completion of each assessment process, each participants' LPCAT scores were generated and saved on a memory stick, while the SPM responses were marked and scored by the researcher. Scores and biographical details of each participant were later transferred to an excel sheet, in preparation for data analysis. It is however worthy of mention that ethical issues regarding test results were adhered to throughout the data collection and analysis process, as only the researcher and her supervisors had access to the data.

3.3 Data analysis

Data analysis encompasses the “process of organizing data into categories and identifying patterns and relationships among the patterns” (McMillan & Schumacher, 2010 p. 368). This was accomplished by first going through a process of data cleaning, and then analysis of the data, in line with the research questions. These processes are further described below:

3.3.1 Data cleaning. Data cleaning entails the process involved in checking data for possible errors, after it has been entered into the computer, in order to ensure that invalid data are corrected immediately (Chambliss & Schutt, 2012). The data cleaning process in this study involved five steps:

Firstly, the data cleaning process brought to light that three of the participants had no biographical details recorded for them in the spreadsheet. These participants were contacted, and they provided the needed information, hence, they were included in the final data. Secondly, it was discovered that two of the participants did not complete either the LPCAT or SPM. Since this would make it difficult to compare their results on both tests, these participants' data were removed from the database. Thirdly, the ages of participants were recalculated to ensure that the ones stated in their biographical details forms were accurate. Upon crosschecking, it was noted that some differed slightly, when considering the date of assessment. The defaulting data were therefore amended to avoid inaccurate descriptive statistics. Aside these, while inputting the response of each item in the SPM, it was discovered that some final scores were miscalculated, as they did not match that of the correct response counts. This was corrected as appropriate to ensure all scores were accurate. Finally, the last stage of data cleaning revealed that one of the research participants

was a master's student. Since the sample for the study included only undergraduate students, this participant's data was excluded from the final sample used for the study.

Following the above data cleaning processes, the final quantitative sample used for the descriptive and correlation analyses therefore included LPCAT and SPM biographical details and scores of 92 undergraduate students from the Pietermaritzburg campus of UKZN.

3.3.2 Quantitative data analysis. Using the Statistical Package for the Social Sciences (SPSS) software, three types of analysis were used for this study:

Firstly, a descriptive statistical analysis was conducted to provide information on the profile of students that made up the sample. This included their age, gender, race, SES and language proficiency. This analysis also provided information on the general tendencies (mean, median, mode), spread of scores (variance, standard deviation) and the relative standing of individual scores compared to other scores (Z score, percentile ranks), within the data collected (Creswell, 2012).

After this was completed, a Pearson Coefficient Correlation analysis was computed to determine the relationship between the LPCAT and SPM scores. In line with the aim of the present study, the pre-test scores on the LPCAT, which represents current achievement levels, were considered as the current intelligence levels and were thus correlated with the total scores on the SPM. Correlation coefficients were determined, and hypotheses were tested at a significance level of 0.05 which has been adjudged as standard (Creswell, 2012).

The last stage of analysis in the study involved a multiple regression analysis of variance, which was conducted to determine the degree of relationship between multiple independent variables and a single dependent variable (Martin & Bridgmon, 2012). Four independent variables were considered in the study namely: race, SES, English language proficiency and gender. For the purpose of this study, SES was measured by the quintile of the high school participants attended, as quintile rankings (Q1-Q5) give an indication of the SES of a school, with Q5 schools ranked as least poor and Q1 as very poor. The rankings of school in South Africa is usually determined by the poverty level and literacy level of the community in which the school is located, as well as the geographical positioning of the school (Khumalo,

2014). Therefore, this study considered participants from Q1 schools to be from very low SES homes and those from Q5 schools to be from high SES homes.

English language proficiency in this study was measured by the language percentage score obtained on the LPCAT, which is measured during the assessment and assigned based on the participant's level of understanding of the language and concepts used in the explanations and feedbacks. A score lower than 75% is considered as limited understanding of the language of assessment, in this case, English language (De Beer, 2016).

For the multiple regression analysis, the combined effect of the demographic variables on each of the LPCAT pre-test, post-test and composite scores was investigated. The effect of the individual variables on each of the LPCAT scores was also evaluated.

3.3.3 Validity, reliability and rigor. Reliability and validity in this research is established by the reliability and validity of the measures used. Durrheim and Painter (2006) opine that the validity of a measure can be established by gathering evidence available on it, in addition to noting its common contextual use. Consequently, the validity and reliability indices of the LPCAT and SPM are highlighted below to justify their use for this study;

LPCAT

Research on the psychometric properties of the LPCAT reveal a high internal consistency of items used, with a coefficient alpha score between 0.925 and 0.981 (De Beer, 2000a). It should be noted however that only internal reliability of the test was assessed since other indices (test-retest, parallel form and split-half) cannot be applied to computerised adaptive test formats (De Beer, 2005). In addition, content, construct and criterion-related validity of the test have also been evaluated (De Beer, 2002; 2003) and results indicate 0.4 and 0.7 correlation scores for the construct validity (De Beer, 2000b; 2006) and statistically significant values of 0.1 and 0.5 for the concurrent and predictive validity, with higher correlation values for the post-test and composite values compared to the pre-test scores (De Beer, 2000b; Van de Merwe & De Beer, 2006).

SPM

The reliability and validity of the SPM has been tested using a wide range of populations,

including non-Whites (sic) and Africans (Raven, 1998). Pind, Gunnarsdóttir and Jóhannesson (2003), document a correlation score ranging from 0.38 – 0.75 in a criterion-related validity study which compared the scores on the SPM to other scholastic achievement scores. Owen (1992) also report a retest reliability score of 0.88 and 0.93 with approximately one-year interval between test administrations, while Abdel-Khalek (2005) also indicate a high internal consistency for the SPM, with a retest reliability score ranging from .69 to .85. In addition, a factorial validity score ranging from .73 to .89 was also reported by Abdel-Khalek (2005).

Rigour is indicated by the congruency in the general design for the study viz problem statement, research question and method, (Durrheim & Painter, 2006). In a bid to achieve this and to indicate objectivity and reduce researcher bias, the entire thesis design, particularly the data collection and analysis process was adequately supervised by the research supervisors, who are trained in the use of both tests. For instance, data collected was scrutinised for possible errors after they were entered into the excel spreadsheet, prior to analysis. Also, after the researcher conducted her own analysis, the research supervisor also conducted a separate analysis, results were compared and were identical, with no errors noted, thus, establishing the validity of results.

3.4 Ethical considerations

In ensuring issues of ethical concern are taken care of, the benchmarks on assessing ethical principles proposed by Emanuel, Wendler, Killen and Grady (2004, p. 931), were adhered to as they relate to this study. These are highlighted below:

- ***Social value***

The study intended to add value to the research and practice community as findings are envisaged to add to the body of knowledge on the science and practice of dynamic assessment and ultimately, assist psychologists in practice in making more informed decisions about adopting dynamic assessment methods.

- ***Scientific validity***

To ensure the scientific design of the research realizes its social value for the participants, who are the primary beneficiaries, results of both tests were disclosed to them, in an appropriate manner.

- ***Fair selection of study population***

To ensure validity of the research, the sample was chosen, bearing their suitability for the research aim in mind. No undue advantage to participants was involved in their selection, and they were informed of their right to voluntary participation or withdrawal.

- ***Favourable risk-benefit ratio***

Participation in the research did not constitute any risk to the participants, compared to the benefit of free access to cognitive evaluation which they enjoyed. The assessment results were presented in a manner that would not cause distress to the participants but give them feedback on areas for improvement.

- ***Independent review***

To ensure researcher accountability, a study proposal was developed, which was subjected to ethical review by the institutional ethics review board (HSSREC), and only after their clearance was the study embarked upon.

- ***Informed consent***

In addition to ethical clearance from HSSREC, permission was sought from the institution registrar to allow recruitment of students. Also, before consenting to be involved in the study, participants were briefed of the research aim and their participation, as well as their right to voluntary participation and withdrawal. This was documented, and their signatures appended.

- ***Ongoing respect for participants***

Confidentiality of participants was ensured and maintained throughout and beyond the period of study as no names or test results of individual participants was disclosed in any form of public report dissemination. All the study data was accessible only to the researcher and her supervisors and confidential documents will be kept under locks for a period of five years after the study has been completed.

Findings from the study are made available by placing a copy of research report in the institution's library, for the benefit of participants and the community it seeks to benefit. Participants were also encouraged to email the researcher for a brief research report of the overall findings, if they wish.

In addition to the outlined ethical considerations, relevant ethical issues pertaining to the use of psychological test and data, as highlighted in subsection 2.1.3, were adhered to. For instance, to protect the copyright of the tests used, a research agreement was entered with

the distributors of the SPM, detailing the conditions of the test use (see Appendix G). In the same vein, the use of the LPCAT for research purpose was acknowledged by the test developer (see email in Appendix H) and the researcher also underwent a training on the LPCAT (see Appendix I for certificate) to ensure proficiency of test administration. Furthermore, the entire process of test administration, scoring and analysis were overseen by the research supervisors, who are licenced psychologists.

Aside these, confidentiality of test results was ensured, and informed consent was obtained from each participant, prior to test administration, (see subsection 3.3.4 for specific details). However, test results were not yet disclosed to participants as at the time of compiling the research report, as during assessment, participants who were interested in test results were informed reports would only be communicated to them at the completion of the research.

3.5 Limitations of the Study

The major limitation of this study are the sample size and the non-culturally representative makeup of the sample, which might have affected the generalizability of the findings. Aside these, the psychological measures used in the study were not administered to each participant in a particular order, which might have influenced their test results. These limitations are further discussed in Chapter 6.

3.6 Summary

This chapter detailed the methodology and general design of the study. This covered the paradigmatic approach, sample and sampling method, data collection and analysis processes, as well as details of the validity, reliability and rigour of the study. The chapter concluded with a highlight of the ethical considerations within the study.

The following chapter presents the results of the data analysis process.

CHAPTER FOUR

RESULT OF DATA ANALYSIS

4.0 Introduction

The results of the data analysis carried out to test the research hypothesis itemised in Chapter 3 are presented in this chapter. The result of the descriptive analysis of the selected demographic variables is first detailed, followed by a presentation of the results of the inferential analysis of each of the hypothesis stated.

4.1 Descriptive analysis

The demographic details sheet was administered to the research participants, to gather information about selected variables including the participant's age at assessment, racial group, gender, high school quintile (representing SES) and English language proficiency (as reflected by their language scores on the LPCAT). A descriptive analysis was thus carried out to outline the profile of the sample and the variables. The results are presented below:

4.1.1 Statistical distribution of sample. Tables 1 through 5 present a summary of the demographic distribution of the sample.

Table 1

Age Distribution of Sample

Age (years)	Frequency	Percentage (%)	Valid Percentage (%)	Cumulative Percentage (%)
17 - 20	52	56.5	56.5	56.5
21 - 24	36	39.1	39.1	95.7
25 - 28	4	4.3	4.3	100.0
Total	92	100.0	100.0	

As indicated in Table 1 above, the participants' ages ranged from 17 to 28 years. While most (56.5%) were between the ages of 17 to 20 years, their mean age was 20.35years ($SD = 2.23$).

Table 2

Description of Sample by Gender

Gender	Frequency	Percentage	Valid Percent	Cumulative Percent
Female	59	64.1	64.1	64.1
Male	33	35.9	35.9	100.0
Total	92	100.0	100.0	

Table 2 above indicates that a larger percentage of the participants (64.1%) were females, with 35.9% being males.

Table 3

Racial Distribution of Sample

Race	Frequency	Percent	Valid Percent	Cumulative Percent
Black	85	92.4	92.4	92.4
Coloured	2	2.2	2.2	94.6
Indian	4	4.3	4.3	98.9
White	1	1.1	1.1	100.0
Total	92	100.0	100.0	

Based on Table 3 above, the sample consisted predominantly of Black students (92.4%), while participants who identified as White made up the least number of participants (1.1%).

Table 4

Socio-Economic Distribution (Indicated by School Quintile) of Sample

High school quintile	Frequency	Percent	Valid Percent	Cumulative Percent
1	5	5.7	5.7	5.7
2	18	20.5	20.5	26.1
3	17	19.3	19.3	45.5
4	11	12.5	12.5	58.0
5	37	42.0	42.0	100.0
Total	88	100.0	100.0	

Of the 92 participants, four completed their high school outside of South Africa, hence their social economic status was not identifiable, since school quintile number could not be assigned to their schools. Therefore, only 88 of the participants were included in this analysis, of which a larger proportion (42%) were from high SES status schools (Quintile 5), while the least number of participants (5%) were from very low SES homes (Quintile 1).

Table 5

English Language Proficiency of Sample

LPCAT Language Score	Frequency	Percent	Valid Percent	Cumulative Percent
75 and above (Good)	88	95.7	95.7	95.7
Less than 75 (Limited)	4	4.3	4.3	100.0
Total	92	100.0	100.0	

Most of the participants (95.7%) possessed good English language proficiency, as indicated by their language scores on the LPCAT.

4.1.2 Basic descriptive analysis of main variables tested. Table 6 details the profile of the scores (including the mean T score, standard deviation, range, minimum and maximum scores) on each of the LPCAT scores (pre-test, post-test and composite score) and the SPM.

Table 6

Descriptive Analysis of Scores on the LPCAT and the SPM

Variables	Mean	Standard deviation	Range	Minimum score	Maximum score
LPCAT Pre-test	53.59	7.178	42	34	76
LPCAT Post-test	53.95	6.701	40	37	77
LPCAT Composite score	53.52	6.777	32	35	67
SPM total score	45.90	7.613	33	25	58

The mean scores for the sample were 53.59 (pre-test), 53.95 (post-test), 53.52 (composite) and 45.90 (SPM). It is interesting to note that the minimum score on the LPCAT increased

from 34 (pre-test) to 37 (post-test), which is not unexpected, given the assumption of an improvement in performance, after relevant training within the LPCAT (Schaap & Luwes, 2013).

4.2 Inferential statistical analysis

Inferential statistical analysis was carried out to examine if a relationship exists, first between the LPCAT pre-test scores and the SPM, and then among the independent variables and the LPCAT pre-test, post-test and composite scores, as well as the nature of such relationship. These were done in line with the set research aim and hypotheses.

4.2.1 Results of correlation analysis. As noted in Chapter section the main research aim of this study was to determine if a statistically significant relationship exists between undergraduate students' performance on the LPCAT and the SPM. Consequently, the following null and alternate hypothesis were drawn:

- **H₀**: There is no statistically significant relationship between students' performance on the LPCAT and the SPM
- **H_a**: There is a statistically significant relationship between students' performance on the LPCAT and the SPM.

To test this hypothesis, a Pearson Product Moment Correlation (*r*) was computed between the LPCAT pre-test scores and the SPM total scores. A probability value (*p*-value) of .05 was used to determine the degree of significance of the relationship between the two variables. The results are presented in table 7 below;

Table 7

Correlation of LPCAT (Pre – Test) Scores with SPM Scores.

		LPCAT Pre – test	SPM Total score
LPCAT Pre- T	Pearson Correlation	1	.651**
	Sig. (2-tailed)		.000
	N	92	92
SPM Total score	Pearson Correlation	.651**	1
	Sig. (2-tailed)	.000	
	N	92	92

Correlation results indicate a strong positive correlation between undergraduate students' LPCAT pre-test and SPM scores, Pearson's $r(92) = .651, p < .05$. This therefore confirms the alternate hypothesis that a significant relationship exists between students' performance on the LPCAT and the SPM, thus providing empirical support for the construct validity of the LPCAT.

4.2.2 Results of multivariate analysis (multiple regression analysis). The secondary aim of this study, as stated in Chapter 3 was to determine if a statistically significant relationship exists between selected demographic variables (race, SES, gender and language proficiency) and students' performance on the LPCAT. Four corresponding hypotheses were also outlined, as in Chapter 3. To test these hypotheses, a multiple regression analysis was completed, and the results are presented next.

4.2.2.1 Hypothesis 1: Relationship between multiple variables (race, SES, English language proficiency and gender) and LPCAT pre-test scores. Tables 8 to 10 present the results of the multiple regression that was run to predict LPCAT pre-test score, based on gender, race, SES (school quintile) and English language proficiency.

Table 8

Model Summary of Regression (Pre-Test)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.393 ^a	.154	.113	6.143

Note. a. Predictors: (Constant), LPCAT language, race, gender, high school quintile

Table 9

Results of ANOVA^a for Regression (pre-test)

Model	Sum of Squares	Df	Mean Square	F	Sig.
1 Regression	570.590	4	142.647	3.780	.007 ^b
Residual	3132.490	83	37.741		
Total	3703.080	87			

Note. a. Dependent Variable: LPCAT Pre-test

b. Predictors: (Constant), LPCAT language, race, gender, high school quintile

Table 10

Regression Coefficients^a (pre-test)

Model	Unstandardised Coefficients		Standardised Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	46.493	4.876		9.535	.000
Gender	-2.321	1.430	-.171	1.623	.108
Race	1.474	1.384	.113	1.065	.290
High School Quintile	.885	.523	.184	1.692	.094
LPCAT language	.063	.039	.174	1.614	.110

Note. a. Dependent Variable: LPCAT pre-test

Analysis of variance (as in Table 9) showed that there was no collective significant effect between the variables and LPCAT pre-test scores, $F(4, 83) = 3.780, p = .007$. Also, none of the four individual variables significantly predicted the LPCAT pre-test scores ($P > .05$), after controlling for each of the other variables in the model. Hence, we fail to reject the null hypothesis, meaning that no statistically significant relationship exists between the selected demographic variables and students' LPCAT pre – test scores.

4.2.2.2 Hypothesis 2: Relationship between Multiple Variables (race, SES, English language proficiency and gender) and LPCAT post-test scores. Tables 11 to 13 present the results of multiple regression that was run to predict LPCAT Post - test score, based on gender, race, SES (school quintile) and English language proficiency.

Table 11

Model Summary of Regression (post-test)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.406 ^a	.165	.125	6.101

Note. a. Predictors: (Constant), LPCAT Language score, race, gender, high school quintile

Table 12

Results of ANOVA^a for Regression (post-test)

Model	Sum of Squares	Df	Mean Square	F	Sig.
Regression	610.490	4	152.622	4.101	.004 ^b
Residual	3089.135	83	37.218		
Total	3699.625	87			

Note. a. Dependent Variable: LPCAT post – test score

b. Predictors: (Constant), LPCAT language score, race, gender, high school quintile

Table 13

Regression Coefficients^a (post – test)

Model	Unstandardised Coefficients		Standardised Coefficients	t	Sig.
	B	St d. Error	Beta		
(Constant)	46.962	4.842		9.699	.000
Gender	-2.619	1.420	-.193	1.845	.069
Race	1.014	1.375	.078	.737	.463
High School Quintile	.784	.519	.163	1.510	.135
LPCAT language score	.078	.039	.215	2.010	.048*

Note. a. Dependent Variable: LPCAT post-test

* statistically significant at .05

As indicated in Table 12, analysis of variance showed that there was no collective significant effect between the variables and LPCAT Post-test scores, $F(4, 83) = 4.101, p = .004$.

Therefore, we fail to reject the null hypothesis and conclude that no statistically significant relationship exists between selected demographic variables and LPCAT post-test scores.

However, when the individual predictors were examined further to determine their effect (Table 13), English language proficiency ($t = 2.01, p = .048$) was a significant predictor in the model. This implies that English language proficiency significantly predicted LPCAT post-test scores, when other variables are accommodated for.

4.2.2.3 Hypothesis 3: Relationship between Multiple Variables (race, SES, English language proficiency and gender) and LPCAT composite scores. Tables 14 to 16 present the results of the multiple regression that was run to predict LPCAT composite score, based on gender, race, SES (school quintile) and English language proficiency.

Table 14

Model Summary of Regression (Composite Score)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.382 ^a	.146	.105	6.049

Note. a. Predictors: (Constant), LPCAT language, race, gender, high school quintile

Table 15

Results of ANOVA^a for Regression (Composite Score)

Model	Sum of Squares	Df	Mean Square	F	Sig.
Regression	518.035	4	129.509	3.539	.010 ^b
Residual	3037.408	83	36.595		
Total	3555.443	87			

Note. a. Dependent Variable: Composite

b. Predictors: (Constant), LPCAT language, race, gender, high school quintile

Table 16

Regression Coefficients^a (Composite Score)

Model	Unstandardised Coefficients	Standardised Coefficients	T	Sig.
-------	-----------------------------	---------------------------	---	------

	B	Std. Error	Beta		
(Constant)	46.917	4.801		9.771	.000
Gender	-2.290	1.408	-.172	-	.108
Race	1.017	1.363	.079	.746	.458
High School Quintile	.856	.515	.182	1.663	.100
LPCAT language	.065	.039	.182	1.685	.096

Note. a. Dependent Variable: Composite

Analysis of variance presented in Table 15 showed that there was no collective significant effect between the variables and LPCAT composite score, $F(4, 83) = 3.539, p = .010$. Also, none of the four individual variables significantly predicted the LPCAT Pre-test scores ($P > .05$), after controlling for each of the other variables in the model. Hence, we fail to reject the null hypothesis, meaning that no statistically significant relationship exists between the selected demographic variables and students' LPCAT composite scores.

4.3 Summary

This chapter contained a description of the profile of the research sample and variables, as well as presentation of the inferential statistical analysis that guided the exploration of the research aims and hypotheses. The succeeding chapter will discuss and interpret the data herein presented.

CHAPTER 5

DISCUSSION OF FINDINGS

5.0 Introduction

Chapter 4 presented the results of the statistical analysis of the data collected to test the research hypotheses detailed in Chapter 1. In this chapter, the research findings will be discussed to provide answers to the research questions. The positioning of the research findings, relative to existing body of knowledge and theory, as reviewed in Chapter 2, as well as implications of the findings will also be presented in this chapter.

5.1 Discussion of results

The purpose of the present inquiry was to investigate the construct validity of the LPCAT. This was achieved by examining if a significant relationship existed between students' scores on the LPCAT and the SPM. To further explore the validity of the LPCAT, four secondary aims were drawn, to evaluate the effect of certain demographic variables on students' performance on the LPCAT. The following sections therefore discuss the research findings, guided by the research questions. The implications of the findings are also touched on.

5.1.1 Main research aim: construct validity of the LPCAT. The study's main research aim was fulfilled by examining the following null and alternative hypothesis;

- **H₀:** There is no statistically significant relationship between students' performance on the LPCAT and the SPM.
- **H_a:** There is a statistically significant relationship between students' performance on the LPCAT and the SPM.

As stated in Chapter 3, the pre-test score on the LPCAT is intended to measure a similar construct as that measured by static cognitive tests, such as the SPM. It is that which states the current level at which the individual is performing, prior to training, hence the basis for speculating a correlation with scores on the SPM. The results of the statistical analysis carried out in this study, as presented in Chapter 4, demonstrated that a highly significant positive association existed between students' scores on the LPCAT and their scores on the SPM, $r = 0.651$ ($P < .05$). Hence, the null hypothesis (H_0) was rejected and the alternate hypothesis (H_a) accepted, indicating that the LPCAT indeed measures a similar fluid reasoning construct as that measured by the SPM. This therefore supports the claim that the LPCAT is a measure of fluid intellectual ability (De Beer, 2006), in other words, it lends credence to the construct validity of the LPCAT. In addition, the approach to evaluating the construct validity of the LPCAT used in this study is in line with Downing's (2003) suggestion that data be collected to confirm or negate prior stated hypotheses. At the same time, it follows similar approach used by other construct validity investigations (De Beer, 2003; Maruff et al., 2009; Norris & Tate, 2000; Rushton et al., 2004).

Most studies on the LPCAT have focused on its predictive validity, using the post-test scores (Loggie, 2007; Mphokane, 2014; Schaap & Luwes, 2013; Schoeman, De Beer, & Visser, 2008), hence much is not known about the nature of the pre-test scores, therefore, this study's findings cannot be adequately compared to others, as it investigated a different aspect of the instrument. Nevertheless, this research finding is consistent with another study that have found a positive correlation between the LPCAT pre-test score and other cognitive measures (De Beer, 2003). In addition, considering evidence in support of the construct validity of the SPM (Owen, 1992, Pind, Gunnarsdóttir & Jóhannesson, 2003), a strong correlation with it is a good indicator of the validity of the LPCAT.

Therefore, it can be said that the current finding not only provides a robust authentication of the LPCAT, but is also of relevance to psychological assessment practitioners, considering that limited empirical evidence on the validity and reliability of measures are some of the limitations of the dynamic assessment approach (Caffrey et al., 2008; Elliot, 2000a; Haywood & Tzuriel 2002; Lidz, 2009).

5.1.2 Secondary research aim: effect of selected demographics on LPCAT pre-test, post-test and composite scores. As indicated in Chapter 3, three secondary research questions and corresponding hypotheses were raised, to better evaluate the results obtained from the analysis of the main research question. These investigated the effect of selected demographics (race, SES, gender and language proficiency) on three of the scores generated from the LPCAT (Pre-test, post-test and composite scores). The results of the analyses are discussed below and implications drawn.

5.1.2.1 Hypothesis 1: The effect of race, SES, English language proficiency and gender on LPCAT pre-test scores.

H₀: There is no statistically significant relationship between selected demographic variables and students' LPCAT pre-test scores.

H_a: There is a statistically significant relationship between selected demographic variables and students' LPCAT pre-test scores.

The results of relevant data analysis (Chapter 4: 4.3.2.1) showed that none of the independent variables had a statistically significant effect on the participants' LPCAT pre-test scores, $F(4, 83) = 3.780, P > 0.001$, therefore, the null hypothesis failed to be rejected. In other words, none of either race, gender, English language proficiency, nor SES of participants predicted their scores on the LPCAT. The implication of this empirical evidence therefore is that; pre-test scores on the LPCAT are not associated with race, gender, English language proficiency nor SES of the test taker, consistent with the culture-fair claim of the LPCAT (De Beer, 2003; 2006).

Meanwhile, although a comparison of the effect of the demographic variables on the LPCAT and SPM is beyond the purview of the present study, it is worthy of mention that when a similar regression was run using the SPM as a constant variable, race had a statistically significant effect on the mean score ($r = .024, P < .05$) after controlling for each of the other independent variables. Though the extent of a true prediction in this regard might be limited,

given the disproportionate composition of the study sample (92.4% Black), this still raises some concern about the culture-fairness of the SPM, since similar results could have been expected on the LPCAT. Additionally, when findings of previous studies on the SPM that report racial difference need to be considered (Owen, 1992; Rushton et al., 2004), further investigation in this regard appears to be warranted.

Over and above this, considering the controversies surrounding cognitive assessment, such as the susceptibility of most test measures to culture-bias (Brown, Reynolds & Whitaker, 1999; Seifert, 2011) and the significant effect of previous learning (Dickens, 2008), which make the measures amenable to prejudiced practices, (as seen in Brigham, 1923; Rushton & Jensen, 2005), it is imperative to have access to measures that are limited in terms of such bias. This research finding is therefore pertinent because it suggests that in addition to having a highly significant construct validity, the LPCAT pre-test scores can be conveniently utilised for university students of varying demographical background, without fear of prejudice against any group of students.

5.1.2.2 Hypothesis 2: The effect of race, SES, English language proficiency and gender on LPCAT Post-Test scores.

H₀: There is no statistically significant relationship between selected demographic variables and students' LPCAT post-test scores.

H_a: There is a statistically significant relationship between selected demographic variables and students' LPCAT post-test scores.

The result of the statistical analysis to determine if race, gender, English language proficiency and SES were associated with performance on the LPCAT post-test score, presented in 4.3.2.2, demonstrated that the variables generally do not significantly correlate with LPCAT post-test scores, $F= 4.101$, $P>0.001$. Hence, the study failed to reject the null hypothesis, meaning that no statistically significant relationship existed between the variables tested and LPCAT post-test score. Hence, this finding substantiates both the culture-fair claim of the LPCAT post-test and the argument for the assessment of learning potential, rather than IQ levels.

Further on this hypothesis however, when other variables were held constant, English language proficiency significantly correlated with the post-test score ($P<.05$). This translates to mean that individuals with higher English language proficiency might be more likely to

have higher learning potential scores. This therefore brings to fore, the possibility of language bias of the LPCAT post-test score when English language proficiency of the examinee is not taken into consideration.

As described in Chapter 2, the LPCAT post-test was designed to measure learning potential, rather than IQ, to buffer the documented inadequacies of static measures (De Beer, 2006; Lunt, 2001; Murphy & Maree, 2006; Tzuriel, 2001), specifically with regards to culture bias. Therefore, it is expected that performance on the LPCAT will not be connected to demographics such as SES, previous knowledge, cultural background and language proficiency (De Beer, 2005). True to this claim, a number of studies have confirmed that performance on the LPCAT is not influenced by demographics such as gender (composite score: Mphokane, 2014) and race (De Beer, 2010; Gilmore, 2009; Mphokane, 2014 (composite score); Schoeman et.al, 2008). Findings from hypothesis 1 of this study also confirm the culture-fairness of the LPCAT pre-test score. However, the same might not hold true for the LPCAT post-test, considering that this study failed to reject the current null hypothesis, based on the statistically significant relationship that was noticed between the post-test scores and language proficiency. Besides, other studies also found a significant relationship between the LPCAT post-test score and English language proficiency (Schoeman, De Beer, & Visser, 2008). The implication of this seems to be that individuals with no less than average English language proficiency are more likely to benefit more from the training included in the test, compared to those with limited English language proficiency when their English language proficiency is not taken into consideration before testing.

However, considering that the importance of language proficiency for academic achievement and progress, cognitive and educational development have been well established in the literature (Webb 2002; Zaimaan, Van de Flier and Thijz, 2000), it appears valid to assume that finding a relationship between English language proficiency and the post-test scores does not diminish the validity of the LPCAT as a measure, rather it establishes the importance of exploring the language proficiency of the examinee prior to test administration. The LPCAT instruction has been made available to be read out in the 11 official languages in South Africa (De Beer, 2006). Therefore, it might rather be worthwhile to re-investigate the effect of language proficiency on the LPCAT post-test

when instructions are given in the examinee's most proficient language, in order to shed more light in this regard. This could be the subject of future research.

In general, however, the results of the current hypothesis lend credence to the culture-fairness of the LPCAT post-test. On a broader level, it also supports the argument for the assessment of learning potential, rather than IQ levels, which is in line with Vygotsky's notion of the zone of proximal development, as expatiated in Chapter 2: 2.7.1 of this report. Lastly, this finding establishes the practicability of the LPCAT in a socio-culturally diverse context as in South Africa.

5.1.2.3 Hypothesis 3: Relationship between multiple variables (race, SES, English language proficiency and gender) and LPCAT composite scores.

H₀: There is no statistically significant relationship between selected demographic variables and students' LPCAT composite scores.

H_a: There is a statistically significant relationship between selected demographic variables and students' LPCAT composite scores.

The last pursuit of the current study was to evaluate the relationship between the selected demographic variables and LPCAT composite scores. Results of the multiple regression analysis, presented in Chapter 5: 5.1.2.3, suggest that no significant relationship exists between race, gender, SES and English language proficiency and LPCAT composite score, $F(4, 83) = 3.539, p > 0.001$. At the same time, none of the individual variables displayed a significant relationship with the composite score ($p > .05$) after controlling for others. This finding is not unexpected given the result of the earlier hypothesis tested, since the composite score is a reasoned aggregate of the pre-test, post-test and difference scores (De Beer, 2005).

According to the test developer (De Beer, 2005), the composite score considers the level of improvement that is shown, relative to current ability levels (i.e. pre-test scores). This is of greater advantage compared to using only the difference scores, given that improvement is not particularly shown when current ability levels are high, compared to when the current ability levels are low. In other words, a low difference score will not equate to low potential levels when the initial ability levels are high. Given the relevance of the composite score therefore, the current finding is of importance, as it serves to indicate that when decisions are made, using the global potential (composite score) levels of the individual, such

decisions are not likely to be compromised by the race, gender, SES or English language proficiency of the individual. Likewise, the results serve to confirm other research that has found no statistically significant relationship between the LPCAT composite score and race and gender (Gilmore, 2009; Mphokane, 2014).

5.2 Summary

This chapter discussed the research findings, in the light of the research questions the study aimed to answer and, in relation to the theoretical framework within which the study is positioned. The implications of the findings were also discussed, as well as how they compare with previous studies, in the process uncovering new knowledge about the LPCAT.

In summary, the current research found a strong correlation between the LPCAT and the SPM, thereby achieving its aim of validating scores generated on the LPCAT. In addition, the research findings confirm the non-amenable of the LPCAT pre-test, post-test and composite-test scores to demographic variables tested. Also, the study showed that the English language proficiency of examinees need to be taken into consideration before testing, to truly benefit from the culture-fairness of the LPCAT post-test.

The next and last chapter will summarise and conclude the study. Recommendations and suggestions for further research will also be presented.

CHAPTER 6

SUMMARY, CONCLUSION AND RECOMMENDATIONS

6.0 Introduction

This study investigated the construct validity of the LPCAT, by comparing scores on it to those obtained on the SPM, which is intended to measure a similar construct. In addition, the study investigated the culture-fair claim of the LPCAT, by evaluating if a significant relationship exists between certain demographics and the pre-test, post-test and composite scores generated on the LPCAT. This chapter therefore recaps the research process and findings, presents concluding statements and recommendations, as well as highlights the limitations of the study and areas that need further investigation.

6.1 Summary of research process and findings

The aim of this study was twofold; (i) to investigate the construct validity of the LPCAT and (ii) to determine if selected demographic variables, specifically; race, SES, gender and English language proficiency impact students' performance on the LPCAT. To achieve these, a quantitative approach was adopted, and data was collected using the LPCAT and SPM. A convenience sampling method was adopted, and the sample was made up of 92 undergraduate students (after data cleaning) currently studying at the UKZN, Pietermaritzburg campus (details of the sample and sampling method is contained in Chapter 3). A descriptive and inferential analysis of the research data was completed and presented in Chapter 4, followed by a discussion of the research findings (Chapter 5). The discussion of findings, guided by the research questions and hypotheses (as in Chapter 1), was completed in relation to relevant literature review and theory, earlier presented in Chapter 2. The results obtained from these endeavours are summarised below:

- A statistically significant positive relationship was found between the LPCAT pre-test scores and the SPM scores.
- No statistically significant relationship was found between race, gender, SES, and English language proficiency and LPCAT pre-test scores.
- No statistically significant relationship was found between race, gender, SES, and English language proficiency and LPCAT pre-test scores. However, when other

variables were accommodated for, English language proficiency showed a significant relationship with the post-test scores.

- No statistically significant relationship was indicated between race, gender, SES, and English language proficiency and LPCAT composite score.

6.2 Research conclusions and implications

This study was propelled by the identified need for the development and use of dynamic measures of learning potential that will supplement information gathered from static intelligence tests, following the realization that socio-demographic factors could influence performance on static measures, thereby disadvantaging already disadvantaged individuals. It is in this realisation that the LPCAT was developed, specifically for use within the South African context – an endeavour which is beyond necessary, considering the inequalities that pervade the nation. However, practitioners and policy makers are still sceptical about dynamic measures as the LPCAT, considering that they are relatively new and not as popular as the static measures and, therefore do not have as much empirical evidence. The current study, which found the LPCAT to be as valid as the long-standing SPM, in the measure of fluid reasoning ability therefore, reinforces the description of the LPCAT as a measure of fluid reasoning ability and as well puts the mind of test users at ease, in terms of the validity of the measure.

At the same time, the research findings reveal that performance on the test is not susceptible to demographic factors, particularly race, gender, SES or English language proficiency. In addition to this, the findings seem to buttress the need for assessment to be conducted in the individual's most proficient language, as it appears that the more proficient the individual is in the language of instruction, the better their performance. These findings therefore support the plausibility of the LPCAT for use within socio-culturally diverse settings, as that of the South African context. With regards to theory therefore, this aspect of the research findings lends credence to a dynamic assessment approach to learning potential, in addition to intelligence testing, on the premise that the shortcomings of the static method, particularly culture-bias, are more likely to be accounted for.

Aside from these, the research findings add to the body of evidence on the on the validity of the LPCAT, which implies that psychologists in practice have more reasons to be rest

assured, when using the instrument. In addition, the research finding identifies the need for practitioners to be cognisant of the language proficiency of their client, when deciding on the language of assessment, for more valid results, which is also in line with the ethics of assessment. Over and above these, findings from this study identify the need for practitioners to consider adopting dynamic approaches in their practice.

6.3 Limitations of the study

Certain limitations might have had significant impact on the findings and effective testing of the hypotheses in this study. The foremost of these being the size and racial composition of the sample. With regards to the sample size, participants were recruited from a small region of the country, hence, caution should be exercised when generalizing findings to other geographical parts of the country. In addition to this, although the recruitment poster used for the study invited students of all racial groups, respondents and participants were predominately Blacks (n= 92.4%), given that the convenience sampling method was used. Also, the difficulty with recruiting participants, due to the ongoing students' protest at the time, made it difficult to consider reviewing the sampling method, to allow for a more representative sample. This disproportion in the racial composition of the sample therefore, might have skewed the results of the analysis. Hence, future research can be designed, testing similar hypothesis, but using a purposive sampling method.

Aside these, the LPCAT and SPM used in this study were not administered to participants in any specific order, which brings about speculations on the possible effect of the training given on the LPCAT post-test on the SPM, for the participants who completed the LPCAT first. However, since the order in which each participant took the test was not noted, it was impossible to discount or prove this conjecture. Future research in this regard is perhaps necessitated.

6.4 Recommendations for future research

Just as this study has contributed to the literature available on the psychometric properties of a dynamic assessment measure, there is need for more research in this regard, going by the dearth of evidence on DA, noted in the review of literature (Chapter 2). This

will not only deal with the issue of reliability and validity of measures, but perhaps increase the popularity of DA within the psychological assessment community.

Meanwhile, going by the socio-culturally non-representative nature of the current research sample, mentioned in 6.3 above, it might be interesting to further test the relationship between the independent variables used in this study (particularly the effect of language of testing on the post-test score) and the LPCAT scores, this time using a purposive sampling method. This will perhaps yield a more generalizable summation on the culture-fairness of the LPCAT.

Likewise, based on the difficulty testing the possible effect that the training given within the LPCAT might have on the performance on the SPM, when both measures are administered in no particular order, as experienced within this study, a replication of the current investigation, with the SPM administered first to all participants, might be warranted. This will further reinforce the significant relationship found between both measures.

In addition, and as noted in the discussion chapter (see 5.1.1), most currently available researches on the LPCAT are based on its predictive validity, especially within the work environment. However, considering that results of static IQ tests are mostly relied upon in deciding the nature and extent of intervention learners will receive, despite the limitations in static assessment to reveal the learner's future ability or give information on the learner's learning processes and cognitive defects (details in Chapter 2), it seems imperative for research to be conducted on the practicability and gains of evaluating learning potential, in addition to current intelligence level, within the school setting.

6.5 Recommendations for practice

With regards to the practical utility of the LPCAT, practitioners need to be wary of the language of instruction they adopt in the administration of the test. Considering that the current research found English language proficiency to be associated with performance on the post-test, it might be necessary to ensure that the examinee is tested in their most proficient language, to fully benefit from the culture-fair attribute of the LPCAT. This

perhaps should not be burdensome contextually since the test instruction is available in the 11 official languages in South Africa.

On a broader level, given that findings from this and previous studies lend credence to the validity of the LPCAT and the dynamic approach, there is need for more to be done to promote the approach and increase its acceptance within the psychological assessment parlance. Therefore, based on the argument that DA requires more skills and expertise which are currently not available within most graduate psychology programs (Haywood & Lidz, 2005; Tzuriel, 2002), and for practice to be on par with evidence, there is need for the inclusion of the dynamic assessment approach in the curriculum of psychology graduate programs. This will attend to the issue of need for expertise and also go a long way in avoiding the challenge of having to help practitioners “unlearn” the static method (Haywood and Lidz, 2005) when they have become used to it.

6.6 Conclusion

This thesis focused on investigating the construct validity of the LPCAT and by extension, the effect of demographic factors on test scores. This chapter summed up the research process and findings and as well detailed the research limitations. Concluding statements and implications of the research findings were also highlighted and the chapter concluded with recommendations for practice and suggestions for future research.

REFERENCES

- Abdel-Khalek, A. M. (2005). Reliability and factorial validity of the standard progressive matrices among Kuwaiti children ages 8 to 15 years. *Perceptual and motor skills*, 101(2), 409-412.
- Aiken, L. R. (2003). *Psychological testing and assessment*. Boston: Allyn & Bacon.
- American Psychological Association. (2010). *Ethical principles of psychologists and code of conduct*. Retrieved from <http://www.apa.org/ethics/code/>
- Amod, Z., Heafield, D., & Seabi, J. (2018). Assessing a remedial intervention programme in developing the planning skills of Grade 4 and 5 learners. *International Journal of Disability, Development and Education*, 65(4), 428-441.
- Australian Psychological Society (2013). *Framework for the effective delivery of school psychological services*. Retrieved from <http://www.psychology.org.au/assets/files/framework-school-psychological-services.pdf>
- Benisz, M., Dumont, R., & Willis, J. O. (2015). From psychometric testing to clinical assessment: Personalities, ideas, and events that shaped David Wechsler's views of intelligence and its assessment. In S. Goldstein, D. Princiotta, & J. A. Naglieri (Eds.), *Handbook of Intelligence* (pp. 163-179). Springer, New York, NY.
- Bickman, L., & Rog, D. J. (Eds.). (2008). *The Sage Handbook of applied social research methods*. London: Sage Publications.
- Binet, A. and Simon, T. (1916). *The intelligence of the feeble-minded*. Baltimore: Williams & Wilkins.
- Blanche, M. T., Durrheim, K., & Painter, D. (2006). *Research in practice: Applied methods for the social sciences*. South Africa: Juta and Company Ltd.
- Boake, C. (2002). From the Binet–Simon to the Wechsler–Bellevue: Tracing the history of intelligence testing. *Journal of clinical and experimental neuropsychology*, 24(3), 383-405.
- Brigham, C. C. (1923). *A study of American intelligence*. Princeton: Princeton University Press.
- Brown, S., & Knight, P. (2012). *Assessing learners in higher education*. London: Routledge.
- Brown, R. T., Reynolds, C. R., & Whitaker, J. S. (1999). Bias in mental testing since Bias in Mental Testing. *School Psychology Quarterly*, 14(3), 208-238.

- Caffrey, E., Fuchs, D., & Fuchs, L. S. (2008). The predictive validity of dynamic assessment: A review. *The Journal of Special Education, 41*(4), 254-270.
- Campbell, D. J. (2008). Vygotsky's Zone of Proximal Development [Episode 4]. *The Learning Theory Podcast*. Podcast retrieved from http://www.danc_ampbell.us/podcast.html.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. New York: Cambridge University Press.
- Carroll, J.B. (1997). The three-stratum theory of cognitive abilities. In D.P. Flanagan, J.L. Genshaft, & P.L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 122-130). New York: The Guilford Press.
- Cattell, R. B. (1987). *Intelligence: Its structure, growth and action*. Amsterdam: Elsevier.
- Chaiklin, S. (2003). The zone of proximal development in Vygotsky's analysis of learning and instruction. In A. Kozulin, B. Gindis, V.S. Ageyev, & S.M. Miller (Eds.), *Vygotsky's educational theory in cultural context* (pp. 39–64). Cambridge, England: Cambridge University Press.
- Chambliss, D. F., & Schutt, R. K. (2012). *Making sense of the social world: Methods of investigation*. Sage.
- Cho, E., Compton, D. L., Gilbert, J. K., Steacy, L. M., Collins, A. A., & Lindström, E. R. (2015). Development of first-graders' word reading skills for whom can dynamic assessment tell us more? *Journal of Learning Disabilities, 1*-18.
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and Standardised assessment instruments in psychology. *Psychological assessment, 6*(4), 284.
- Cohen, R. J., & Swerdlik, M. E. (2010). *Psychological testing and assessment: An introduction to tests and measurement*. USA: McGraw-Hill.
- Cook, D. A., & Beckman, T. J. (2006). Current concepts in validity and reliability for psychometric instruments: theory and application. *The American journal of medicine, 119*(2), 166-e7.
- Creswell, J. W. (2012). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. United States of America: Pearson Education.
- De Beer, M. (2000a). *Learning Potential Computerised Adaptive Test (LPCAT): Technical Manual*. Pretoria: Production printers (UNISA).

- De Beer, M. (2000b). *The construction and evaluation of a dynamic, computerised adaptive test for the measurement of learning potential*. Unpublished D.Phil., thesis, University of South Africa, Pretoria.
- De Beer, M. (2002). *Utility of Learning Potential Computerised Adaptive Test (LPCAT) scores D predicting academic performance of bridging students: A comparison with other predictors*. Paper presented at the Fifth Annual Industrial Psychology Conference, Pretoria.
- De Beer, M. (2003). *A comparison of learning potential results at various educational levels*. Paper presented at the Sixth Annual Society for Industrial and Organisational Psychology of South Africa (SIOPSA) Conference, Johannesburg.
- De Beer, M. (2005). Development of the learning potential computerised adaptive test (LPCAT). *South African Journal of Psychology*, 35(4), 717-747.
- De Beer, M. (2006). Dynamic assessment: Practical solutions to some concerns. *South African Journal of Industrial Psychology*, 32(4), 1-7.
- De Beer, M. (2007). *Use of CAT in dynamic testing*. Poster presented at the GMAC Conference on Computerised Adaptive Testing, St Paul-Minneapolis, 7-8 June.
- De Beer, M. (2010). Longitudinal predictive validity of a learning potential test. *Journal of Psychology in Africa*, 20(2), 225-231.
- De Beer, M. (2016). *LPCAT training manual: interpretation and reports*. Pretoria: Mindset and Managing Initiatives.
- Denzin, N. K., & Lincoln, Y. S. (2000). Introduction: The discipline and practice of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 1–28). Thousand Oaks, CA: Sage.
- Dickens, W. T. (2008). Cognitive ability. In S.N. Durlauf (Ed.), *The New Palgrave: Dictionary of Economics* (pp. 1-896). Palgrave Macmillan UK.
- Domino, G., & Domino, M.L. (2006). *Psychological testing: An introduction* (2nd Ed.). USA: Cambridge University Press.
- Downing, S. M. (2003). Validity: on the meaningful interpretation of assessment data. *Medical education*, 37(9), 830-837.
- Du Plessis, G. A. (2008). Dynamic cognitive assessment: investigating learning potential within the context of an academic institution. (Doctoral dissertation), University of Johannesburg.

- Elleman, A. M., Compton, D. L., Fuchs, D., Fuchs, L. S., & Bouton, B. (2011). Exploring dynamic assessment as a means of identifying children at risk of developing comprehension difficulties. *Journal of learning disabilities, 44*(4), 348-357.
- Elliot, J.G. (2000a). The psychological assessment of children with learning difficulties. *British Journal of Special Education, 27*(2), 59-66.
- Elliot, J.G. (2000b). Dynamic Assessment in educational contexts: Purpose and Promise. In C.S. Lidz, & J.G. Elliot (Eds.), *Dynamic Assessment: Prevailing models and applications*. (pp.713-740). New York, USA: Elsevier Science Inc.
- Elliot, J.G. (2003). Dynamic Assessment in Educational Settings: Realizing Potential. *Educational Review, 55*(1), 15-32.
- Elliott, J. G., & Resing, W. (2015). Can intelligence testing inform educational intervention for children with reading disability? *Journal of Intelligence, 3*(4), 137-157.
- Emanuel, E. J., Wendler, D., Killen, J., & Grady, C. (2004). What makes clinical research in developing countries ethical? The benchmarks of ethical research. *Journal of Infectious Diseases, 189*(5), 930-937.
- Fahrmeier, E.D. (1975). The effect of school attendance on intellectual development in northern Nigeria. *Child Development, 46*, 281–285.
- Flanagan, D. P., & McGrew, K. S (2012). A cross-battery to assessing and interpreting cognitive abilities: Narrowing the gap between practice and cognitive science. In D. P Flanagan & P. L. Harrison, *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 314–325). New York: The Guilford Press.
- Flanagan, D. P., & Ortiz, S. O. (2001). *Essentials of cross-battery assessment*. NY: Wiley & Sons
- Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2007). *Essentials of cross-battery assessment* (2nd Ed.). New York: Wiley.
- Flanagan, D. P., Ortiz, S. O., & Alfonso, V.C. (2013). *Essentials of cross-battery assessment* (3rd Ed.). Hoboken: Wiley & Sons.
- Foster, S. L., & Cone, J. D. (1995). Validity issues in clinical assessment. *Psychological Assessment, 7*(3), 248.
- Foxcroft, C., & Roodt, G. (2009). *Introduction to psychological assessment in the South African context* (3rd Ed.). Cape Town: Oxford University Press.
- Frenkel, S., & Nobile, D. (2014). Learning to learn: Assessment of metacognitive competencies. *INTED2014 Proceedings, 3749-3758*.

- Gardner, H. (1993). *Frames of mind: Theory of multiple intelligences*. California: Fontana Press.
- Gilmore, N. (2009). *The relationship between learning potential and job performance* (Master's dissertation). University of South Africa.
- Glewwe, P., & Jacoby, H. (1992). *Estimating the determinants of cognitive achievement in low income countries*. Washington: World Bank.
- Gliner, J. A., Morgan, G. A., & Harmon, R. J. (2001). Measurement reliability. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40(4), 486-488.
- Goldstein, S., Princiotta, D., & Naglieri, J. A. (2015). *Handbook of intelligence. Evolutionary theory, historical perspective, and current concepts*. New York, NY: Springer.
- Gravetter, F. J., & Wallnau, L. B. (2016). *Statistics for the behavioral sciences*. Boston: Cengage Learning.
- Gredler M., & Shields, C. (2004). Does no one read Vygotsky's words? Commentary on Glassman. *Educational Researcher*, 33(2), 21 – 25.
- Green, C. D. (1999). *Classics in the history of psychology: An electronic resource*. Toronto: York University.
- Greenwood, J. D. (2015). Intelligence defined: Wundt, James, Cattell, Thorndike, Goddard, and Yerkes. In *Handbook of Intelligence* (pp. 123-135). New York, NY: Springer.
- Grigorenko, E. L. (2009). Dynamic assessment and response to intervention: Two sides of one coin. *Journal of Learning Disabilities*, 42(2), 111-132.
- Groth-Marnat, G. (2003). *The Handbook of psychological assessment* (4th Ed.). New Jersey: John Wiley and Sons.
- Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105-117).
- Guthke, J., & Stein, H. (1996). Are learning tests the better version of intelligence tests? *European Journal of Psychological Assessment*, 12(1), 1-13. Thousand Oaks, CA: Sage.
- Hamers, J. H. M., & Resing, W. C. M. (1993). Learning potential assessment: Introduction. In Hamers, JHM, Sijtsma, K., and Ruijsenaars, AJJM (Eds.). *Learning Potential Assessment; Theoretical, Methodological and Practical Issues* (pp.). Amsterdam/Lisse: Swets & Zeitlinger.
- Haywood, H. C., & Lidz, C. S. (2005). International survey of dynamic assessment trainers. *Journal of Cognitive Education and Psychology*, 5(2), 181-198.

- Haywood, H. C., & Tzuriel, D. (2002). Applications and challenges in dynamic assessment. *Peabody Journal of Education*, 77(2), 40-63.
- Heckman, J. J. (1995). Lessons from the bell curve. *Journal of Political Economy*, 103(5), 1091-1120.
- Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology*, 57, 253–270.
- Human-Vogel, S. (2004). Cognition and learning. In Eloff, I & Ebersöhn, L. (Eds.) *Keys to educational psychology* (pp. 13-30). Cape Town: UCT Press.
- Huysamen, G.K. (1996). *Psychological measurement: An introduction with South African examples*. (3rd Ed.). Pretoria: J.L. Van Schaik.
- Iftikhar, R. (2012). Ravens Progressive Matrices: Psychometric evidence, gender and social class differences in middle childhood. *Journal of Behavioural Sciences*, 22(3), 120.
- International Test Commission. (2013). International guidelines for test use. *International Journal of Testing*, 1(2), 93-114.
- Irwing, P., & Lynn, R. (2005). Sex differences in means and variability on the progressive matrices in university students: A meta-analysis. *British Journal of Psychology*, 96(4), 505-524.
- Kamin, L. J. (2006). African IQ and mental retardation. *South African Journal of Psychology*, 36(1), 1-9.
- Kanjee, A. (2001). Cross-cultural test adaptation and translation. In C. Foxcroft & G. Roodt (Eds.), *An introduction to psychological assessment in the South African context* (pp. 86-102). Cape Town: Oxford University Press.
- Karpov, Y. V., & Tzuriel, D. (2009). Dynamic assessment: Progress, problems, and prospects. *Journal of Cognitive Education and Psychology*, 8(3), 228-237.
- Kaufman, A. S. (2009). *IQ testing 101*. New York: Springer.
- Kaufman, A. S., & Kaufman, N. L. (2004). *Manual for Kaufman Assessment Battery for Children—Second Edition (KABC-II)—Comprehensive Form*. Circle Pines, MN: American Guidance Service.
- Kaufman, J. C., Kaufman, S. B., & Plucker, J. A. (2013). Contemporary theories of intelligence. *The oxford handbook of cognitive psychology*, 811-822.
- Khaleefa, O., Amer, Z., & Lynn, R. (2014). IQ differences between arts and science students at the University of Khartoum. *Mankind Quarterly*, 55(1/2), 136.
- Khumalo, N. (2014). *Some implications of the quintile school funding in South African public schools*. (Doctoral dissertation), Stellenbosch University.

- Kline, P. (2013). *Intelligence: The psychometric view*. USA: Routledge.
- Kozulin, A. (2011). Learning potential and cognitive modifiability. *Assessment in Education: Principles, Policy & Practice*, 18(2), 169-181.
- Lidz, C. S. (2009). Dynamic Assessment in School Psychology. *Communique*, 38(2), 16-18.
- Lidz, C. S. (2012). Dynamic assessment approaches. In DP Flanagan, JL Genshaft, & PL Harrison (eds.), Contemporary intellectual assessment. *Theories, tests, and issues*.
- Logie, M. (2007). *Learning potential and academic literacy tests as predictors of academic performance for engineering students*. (Masters dissertation), University of Pretoria.
- Lohman, D. F. (2005). The role of nonverbal ability test in identifying academically gifted students: an aptitude perspective. *Gifted Children Quarterly* 49(2): 111-38.
- Lubbe, C. (2004). Educational psychological assessment. In I. Eloff, & Ebersohn, L. (Ed.), *Keys to educational psychology* (pp. 13-30). Cape Town: UCT Press.
- Lunt, I. (2001). The practice of assessment. In H. Daniels (Ed.), *Charting the agenda: Educational activity after Vygotsky* (pp.145-170). London: Routledge.
- Lynn, R., Allik, J., & Irwing, P. (2004). Sex differences on three factors identified in Raven's Standard Progressive Matrices. *Intelligence*, 32(4), 411-424.
- MacCallum, R. C. (2003). 2001 presidential address: Working with imperfect models. *Multivariate Behavioral Research*, 38(1), 113-139.
- Maruff, P., Thomas, E., Cysique, L., Brew, B., Collie, A., Snyder, P., & Pietrzak, R. H. (2009). Validity of the CogState brief battery: relationship to Standardised tests and sensitivity to cognitive impairment in mild traumatic brain injury, schizophrenia, and AIDS dementia complex. *Archives of Clinical Neuropsychology*, 24(2), 165-178.
- Merrell, K. W., Ervin, R. A., & Peacock, G. G. (2011). *School psychology for the 21st century: Foundations and practices*. New York: Guilford Press.
- Mphokane, A. (2014). *The predictive validity of learning potential and English language proficiency for work performance of candidate engineers*. (Masters dissertation), University of South Africa.
- Neisser, U., G. Boodoo, T.J. Bouchard Jr., A.W. Boykin, N. Brody, S.J. Ceci, D.F. Halpern, J.C. Loehlin, R. Perloff, R.G. Sternberg, and S. Urbina (1996), Intelligence: Knowns and Unknowns, *American Psychologist*, 51 pp. 77-101.

- Norris, G., & Tate, R. L. (2000). The Behavioural Assessment of the Dysexecutive Syndrome (BADS): Ecological, concurrent and construct validity. *Neuropsychological Rehabilitation, 10*(1), 33-45.
- Ortiz, S., Ochoa, S. M., & Dynda, A. M. (2012). Testing with culturally diverse and linguistically diverse populations: Moving beyond the verbal-performance dichotomy into evidence-based practice. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 526–552). New York: The Guilford Press.
- Páchová, A. (2013). The problematic nature of culture-fair testing: training effect differences among Czech and Roma children. *Baltic Journal of Psychology, 14* (1&2), 79-91.
- Pase, M. P., & Stough, C. (2014). An evidence-based method for examining and reporting cognitive processes in nutrition research. *Nutrition research reviews, 27*(2), 232-241.
- Plomin R, Defries JC, McClearn GE, McGuffin P. (2001) *Behavioral Genetics*, 4th ed. New York: Worth Publishers.
- Poehner, M. E. (2010). *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development*. Berlin: Springer Science & Business Media.
- Ponterotto, J. G. (2005). Qualitative research in counselling psychology: A primer on research paradigms and philosophy of science. *Journal of Counselling Psychology, 52*(2), 126.
- Raven, J. (2000) The Raven's Progressive Matrices: Change and stability over culture and time. *Cognitive Psychology, 41*, 1–48.
- Raven, J., Raven, J.C. and Court, J.H. (1998): *Manual for Raven's Advanced Progressive Matrices*. Oxford, England: Oxford Psychologists Press.
- Resing, W. C. (2013). Dynamic testing and individualized instruction: Helpful in cognitive education? *Journal of Cognitive Education and Psychology, 12*(1), 81.
- Roid, G. (2003). *Stanford-Binet Fifth Edition*. Itasca, IL: Riverside.
- Rowe, S.M., & Wertsch, J.V. (202): Vygotsky's model of cognitive development. In U. Goswami (Ed.), *Blackwell handbook of child-hood development* (pp. 538 – 554). Oxford, England: Blackwell.
- Rushton, J.P., & Jensen, A.R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy and Law, 11*, 235–294.

- Rushton, J. P., Skuy, M., & Bons, T. A. (2004). Construct validity of Raven's advanced progressive matrices for African and non-African engineering students in South Africa. *International Journal of Selection and Assessment*, 12(3), 220-229.
- Rust, J., & Golombok, S. (2014). *Modern psychometrics: The science of psychological assessment*. Routledge.??
- Sattler, J. M. (2008). *Assessment of children: Cognitive foundations*. San Diego, CA: JM Sattler.
- Schaap, P., & Luwes, M. (2013). Learning potential and academic literacy tests as predictors of academic performance for engineering students. *Acta academica*, 45(3), 181-214.
- Schoeman, A., De Beer, M., & Visser, D. (2008). The relationship between learning potential, English language proficiency and work-related training test results. *South African Journal of Labor Relations*, 32(1), 47-62.
- Seifert, K. (2011). *Standardised and other formal assessments: Issues with Standardised tests*. Retrieved from <http://cnx.org/contents/703c6bc9-bbc6-4325-9e1c-60a9cd29591f@1>.
- Sellami, K., Infanzón, E., Lanzón, T., Díaz, A., & Lynn, R. (2010). Sex differences in means and variance of intelligence: Some data from Morocco. *Mankind Quarterly*, 51(1), 84.
- Skuy, M., Gewer, A., Osrin, Y., Khunou, D., Fridjhon, P., & Rushton, J.P. (2002). Effects of mediated learning experience on Raven's matrices scores of African and non-African university students in South Africa. *Intelligence*, 30, 221-232.
- Smit, M. (2010). *Educational psychologists' view of the relevance of dynamic assessment for their practice* (Masters dissertation), University of Stellenbosch.
- Snowman, J., & McCown, R. (2013). *Edpsych*. USA: Cenage learning.
- Spanoudis, G., Natsopoulos, D., & Lynn, R. (2016). Sex differences for 10 to 17-year olds on the Standard Progressive Matrices in Cyprus. *Mankind Quarterly*, 57(2), 283.
- Spearman, C. (1904). "General Intelligence," objectively determined and measured. *The American Journal of Psychology*, 15(2), 201-292.
- Stern, W. (1914). *The psychological methods of testing intelligence*. Baltimore, MD, US: Warwick & York.
- Sternberg, R. J., & Grigorenko, E. L. (2002). *Dynamic testing: The nature and measurement of learning potential*. Cambridge university press.

- Sternberg, R. J., Grigorenko, E. L., Ngorosho, D., Tantufuye, E., Mbise, A., Nokes, C., Jukes, M., & Bundy, D. A. (2002). Assessing intellectual potential in rural Tanzanian school children. *Intelligence*, *30*(2), 141-162.
- Terman, L. M. (1922). A new approach to the study of genius. *Psychological Review*, *29* (4), 310–318.
- Thorndike, R. M. (1997). The early history of intelligence testing. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 3-16). New York, NY, US: Guilford Press.
- Tzuriel, D. (2000). Dynamic assessment of young children: Educational and intervention perspectives. *Educational Psychology Review*, *12*(4), 385-435.
- Tzuriel, D. (2001). Dynamic assessment of learning potential: Ability, achievement, and behavior in children. In J.J.W. Andrews, D.H. Saklofske, & H.L. Janzen (Eds.), *Handbook of psychoeducational assessment* (pp. 451-496). San Diego; CA: Academic Press.
- Tzuriel, D. (2012). Dynamic assessment of learning potential. In *Self-directed learning-oriented assessments in the Asia-Pacific* (pp. 235-255). Dordrecht: Springer.
- Van De Vijver, F. J. (2002). Cross-Cultural Assessment: Value for Money? *Applied Psychology*, *51*(4), 545-566.
- Van Eeden, R. & De Beer, M. (2009). Assessment of cognitive functioning. In C. Foxcroft, & Roodt, (Ed.), *Introduction to psychological assessment in the South African context* (3rd Ed., pp. 127-147). Cape Town: Oxford University Press.
- Vereijken, B. (2010). The complexity of childhood development: variability in perspective. *Physical therapy*, *90*(12), 1850-1859.
- Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Webb, V. (2002). English as a second language in South Africa's tertiary institutions: a case study at the University of Pretoria. *World Englishes*, *21*(1): 49-61.
- Wechsler, D. (1939). *The measurement of adult intelligence*. Baltimore: Williams and Wilkins.
- Wechsler, D. (1944). *Measurement of adult intelligence*. Baltimore: Williams and Wilkins.
- Wechsler, D. (2008). *Wechsler adult intelligence scale -fourth edition (WAIS-IV)*. San Antonio: Pearson.
- Wechsler, D. (2012). *Wechsler preschool and primary scale of intelligence-fourth edition (WPPSI-IV)*. San Antonio: Pearson.

- Wechsler, D. (2014). *Wechsler intelligence scale for children-fifth edition: Technical and interpretive manual*. Bloomington: Pearson.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson Battery—Third Edition (WJ III)*. Itasca, IL: Riverside
- Zaimaan, H., Van der Flier, H., & Thijz, G.D. (2000). Selection as contract to teach at the student's level: experiences from a South African mathematics and science foundation year. *Higher Education* 40(1): 1-21.

APPENDIX A: UKZN HSSREC ETHICAL CLEARANCE (INITIAL COPY)



23 May 2016

Mrs Ganiyat O Zurakat 213570399
School of Applied Human Sciences
Pietermaritzburg Campus

Dear Mrs Zurakat

Protocol reference number: HSS/0577/016M

Project title: Dynamic assessment of cognitive ability: Investigating the construct validity of a learning potential measure among first year students of UKZN.

Expedited Approval

In response to your application dated 18 May 2016, the Humanities & Social Sciences Research Ethics Committee has considered the abovementioned application and the protocol have been granted **FULL APPROVAL**.

Any alteration/s to the approved research protocol i.e. Questionnaire/Interview Schedule, Informed Consent Form, Title of the Project, Location of the Study, Research Approach and Methods must be reviewed and approved through the amendment/modification prior to its implementation. In case you have further queries, please quote the above reference number. Please note: Research data should be securely stored in the discipline/department for a period of 5 years.

The ethical clearance certificate is only valid for a period of 3 years from the date of issue. Thereafter Recertification must be applied for on an annual basis.

I take this opportunity of wishing you everything of the best with your study.

Yours faithfully

Dr Shamila Naidoo (Deputy Chair)

/px

cc Supervisor: Ms C Mitchell & Dr N Munro
cc Academic Leader Research: Professor D Wassenaar
cc School Administrator: Ms Nondumiso Khanyile

Humanities & Social Sciences Research Ethics Committee

Dr Shenuka Singh (Chair)

Westville Campus, Govan Mbeki Building

Postal Address: Private Bag X54001, Durban 4000

Telephone: +27 (0) 31 260 3587/8350/4557 Facsimile: +27 (0) 31 260 4609 Email: ximbgo@ukzn.ac.za / snvmanm@ukzn.ac.za / mahump@ukzn.ac.za

Website: www.ukzn.ac.za

1910 - 2010
100 YEARS OF ACADEMIC EXCELLENCE

Founding Campuses: ■ Edgewood ■ Howard College ■ Medical School ■ Pietermaritzburg ■ Westville

APPENDIX B: UKZN HSSREC ETHICAL CLEARANCE (SUBSEQUENT COPY)



11 June 2018

Mrs Ganiyat O Zurakat (213570399)
School of Applied Human Sciences – Psychology
Pietermaritzburg Campus

Dear Mrs Zurakat,

Protocol reference number: HSS/0577/016M

Project title: Dynamic assessment of cognitive ability: Investigating the construct validity of the Learning Potential Computerised Adaptive Test (LPCAT) *within an academic context*

Approval Notification – Amendment Application

This letter serves to notify you that your application and request for an amendment received on 28 May 2018 has now been approved as follows:

- Change in Title
- Amendment to sample population

Any alterations to the approved research protocol i.e. Questionnaire/Interview Schedule, Informed Consent Form; Title of the Project, Location of the Study must be reviewed and approved through an amendment /modification prior to its implementation. In case you have further queries, please quote the above reference number.

PLEASE NOTE: Research data should be securely stored in the discipline/department for a period of 5 years.

The ethical clearance certificate is only valid for period of 3 years from the date of original issue. Thereafter Recertification must be applied for on an annual basis.

Best wishes for the successful completion of your research protocol.

Yours faithfully

Professor Shenuka Singh (Chair)

/ms

cc Supervisor: Ms Carol Mitchell & Dr N Munro
cc Academic Leader Research: Dr Maud Mthembu
cc School Administrator: Ms Nondumiso Khanyile

Humanities & Social Sciences Research Ethics Committee

Dr Shenuka Singh (Chair)

Westville Campus, Govan Mbeki Building

Postal Address: Private Bag X54001, Durban 4000

Telephone: +27 (0) 31 260 3587/8350/4667 Facsimile: +27 (0) 31 260 4609 Email: ximbap@ukzn.ac.za / snymam@ukzn.ac.za / mohunpo@ukzn.ac.za

Website: www.ukzn.ac.za



100 YEARS OF ACADEMIC EXCELLENCE

Founding Campuses: Edgewood Howard College Medical School Pietermaritzburg Westville

APPENDIX C: GATEKEEPER'S PERMISSION



20 April 2016

Mr Ganiyat Zurakat
School of Applied Human Sciences
College of Humanities
University of Kwa-Zulu Natal
Pietermaritzburg Campus

Email: 213570399@stu.ukzn.ac.za

Dear Mr Zurakat

RE: PERMISSION TO CONDUCT RESEARCH

Gatekeeper's permission is hereby granted for you to conduct research at the University of KwaZulu-Natal (UKZN), towards your postgraduate studies, provided Ethical clearance has been obtained. We note the title of your research project is:

"Dynamic assessment of cognitive ability: Investigating the construct validity of the Learning Potential Computerized Assessment Test (LPCAT) within an academic context".

It is noted that you will be constituting your sample by conducting a psychometric test with 1st year students in the 3 colleges on the Pietermaritzburg Campus, UKZN.

Please ensure that the following appears on your questionnaire/attached to your notice:

- Ethical clearance number;
- Research title and details of the research, the researcher and the supervisor;
- Consent form is attached to the notice/questionnaire and to be signed by user before he/she fills in questionnaire;
- gatekeepers approval by the Registrar.

Data collected must be treated with due confidentiality and anonymity.

Yours sincerely

MR SS MOKOENA
REGISTRAR

Office of the Registrar

Postal Address: Private Bag X54001, Durban, South Africa

Telephone: +27 (0) 31 260 8005/2206 Facsimile: +27 (0) 31 260 7824/2204 Email: registrar@ukzn.ac.za

Website: www.ukzn.ac.za



1910 - 2010
100 YEARS OF ACADEMIC EXCELLENCE

Founding Campuses: ■ Edgewood ■ Howard College ■ Medical School ■ Pietermaritzburg ■ Westville

APPENDIX D: RECRUITMENT POSTER



HEY.....

Did you know that some scholarships, bursaries and employers require you to complete valid learning potential assessments or IQ tests as part of their selection processes?

SO.....

Would you like an opportunity to get exposed to some of those assessments?

Would you like to know your potential for academic success

Are you an undergraduate student at UKZN?

Do you have at least an hour to spare?

WHY NOT VOLUNTEER FOR A STUDY TO TICK THOSE BOXES???

To volunteer

Please visit the Psychology Lab (psychology Building) on any of these dates:

Friday	Monday	Monday	Friday	Friday
September 30, 2016	October 3, 2016	October 3, 2016	October 7, 2016	October 7, 2016
1:00 pm	9:00 am	11:00 am	11:00 am	1:00 pm
-	-	-	-	-
3:00 pm	11:00 am	1:00 pm	1:00 pm	3:00 pm

For more information, please contact:

Email: [REDACTED]

WhatsApp: [REDACTED]

NB: R30 meal vouchers will be given to compensate for time!!!

APPENDIX E: DEMOGRAPHIC DETAILS SHEET

DEMOGRAPHIC DETAILS SHEET

Please note that these details will be used to describe the sample for this study and any information you provide here will be used in addition to the test scores, solely for the purpose of my research. Please fill in your details or tick the appropriate box:

1. NAME:

.....

2. COLLEGE:

.....

3. DEGREE:

.....

...

4. DATE OF BIRTH: AGE:

.....

5. GENDER: Male Female

6. RACE: Black Colored Indian White

Other (please specify):

7. Name of high school attended:

.....

8. Province and City in which high school is located:

.....

8. First Language:

.....

9. Have you ever completed a psychological assessment test? Yes No

10. If you answered yes above, please indicate the year and the name of the test(s).

.....

11. have you ever used a computer? yes no

APPENDIX F: INFORMED CONSENT FORM



4th August 2016.

Dear student,

INFORMATION SHEET AND CONSENT TO PARTICIPATE IN RESEARCH

I am Ganiyat Zurakat, an Educational Psychology Masters student of the University of KwaZulu-Natal, Pietermaritzburg campus. I am conducting a validation study of a cognitive assessment instrument (Learning Potential Computerised Adaptive Test (LPCAT) and would be comparing it with another Standardised measure (Ravens Standard Progressive Matrices (SPM)). You are being invited to consider participating in this study.

The study is expected to recruit about 100 to 200 participants among first year students, across the three colleges of the Pietermaritzburg campus of UKZN. Participation in this study requires you to complete two (2) psychological assessment instruments (computer and paper-and-pencil based). The duration of your participation, if you choose to enrol and remain in the study, is expected to be about 1.5 hours.

Aside the time it will take you to complete the tests, there are no potential harm or risks envisaged in your participation, but I hope the study will be beneficial in the aspect of providing you a feedback on your learning potential and cognitive ability levels (if you choose to know). It is my hope as well that the study will provide information on the validity of the LPCAT for use in academic settings and by psychologists in practice, in addition to adding to the body of knowledge currently available.

This study has been ethically reviewed and approved by the UKZN Humanities and Social Sciences Research Ethics Committee (approval number), and provision has been made for you to be attended to at the student support centres across the three campuses, should you require any psychological support following your participation in this study.

Please note that your participation in this study is voluntary and that you can decide to withdraw such at any point without any accruing loss, penalty nor consequence. Also note that there are no financial benefits attached and no costs will be incurred by you either, should you decide to participate in the study.

To maintain confidentiality, any personal information gotten from you will only be assessable to me and my supervisors, will be kept under locks for a period of five years and destroyed via a shredder thereafter. Data generated from assessment will be used for research purposes only, without any indication of your name or identity.

Should you have any questions about this study or its procedures, now or in the future, please contact me, my supervisor or the research office at the following contact details:

RESEARCHER	SUPERVISORS	RESEARCH ETHICS COMMITTEE
Ganiyat Zurakat Email: ghanizurak@yahoo.com Tel: ██████████	Ms. Carol Mitchel Email: mitchelle@ukzn.ac.za Tel: ██████████ Dr. Nicholas Munro Email: munron@ukzn.ac.za	Humanities & Social Sciences Research Ethics Administration. Email: HSSREC@ukzn.ac.za Tel: 27 31 2604557 Fax: 27 31 2604609

Please sign and return the detachable part below and return to the researcher if you are willing to participate in the study

DECLARATION OF CONSENT

I (Full names of participant) have been informed about the study entitled ‘Dynamic assessment of cognitive ability: Investigating the construct validity of the Learning Potential Computerised Assessment Test (LPCAT) within an academic context’ by Ganiyat Zurakat.

I understand the purpose and procedures of the study and have been given an opportunity to ask questions about the study and have had answers to my satisfaction. I have also been informed about the benefit of participating and support I can get if any psychological discomfort occurs to me as a result of study-related procedures.

I declare that my participation in this study is entirely voluntary and that I may withdraw at any time without any direct consequences to my person.

If I have any further questions/concerns or queries related to the study I understand that I may contact the researcher at ghanizurak@yahoo.com or ██████████

If I have any questions or concerns about my rights as a study participant, or if I am concerned about an aspect of the study or the researcher then I may contact:

HUMANITIES & SOCIAL SCIENCES RESEARCH ETHICS ADMINISTRATION

Research Office, Westville Campus

Govan Mbeki Building

Private Bag X 54001

Durban

4000

KwaZulu-Natal, SOUTH AFRICA

Tel: 27 31 2604557 - Fax: 27 31 2604609

Email: HSSREC@ukzn.ac.za

Signature of Participant

Date

APPENDIX G: RESEARCH AGREEMENT (SPM)

JOHANNESBURG OFFICE

15 hunter avenue | ferndale | randburg | 2194
p.o. box 2560 | pinegowrie | 2123 | johannesburg | south africa
+27 11 781 3705/6/7 (tel) | +27 11 781 3703 (fax)
info@jvrafrica.co.za (email)

CAPE TOWN OFFICE

3rd floor | imperial bank terraces | tyger waterfront | carl cronje drive | bellville
p.o. box 5534 | tygervalley | 7536 | cape town | south africa
+27 861 333 916 (tel) | +27 861 004 225 (fax)
capeinfo@jvrafrica.co.za (email)

FREE STATE OFFICE

31 wessels street | eureka | bethlehem | 9700 | south africa
+27 58 303 3843 (tel) | freeestate-info@jvrafrica.co.za (email)

Co.Reg.No. 2001/015618/07 VAT Reg.No. 4300195064



RESEARCH AGREEMENT

As a psychological test provider, JvR Psychometrics (JvR) has encouraged research on the assessments we distribute since our inception in 1993. The company is committed to supporting research that contributes to the expansion of knowledge in the scientific community of Psychology in Southern and Sub Saharan Africa. Research on international and South African tests, distributed and published by JvR ensures that we offer products and services of the highest quality to our customers.

By conducting studies associated with the assessments that JvR distributes, research partners form an integral part of JvR's test development and validation processes. This agreement serves to delineate the roles and responsibilities of JvR and Ganiyat Zurakat (The Researcher) in the study Dynamic assessment of cognitive ability: Investigating the construct validity of a learning potential measure among first year students of UKZN that will run from 01 August 2016 until 01 August 2017.

List of assessments included in this agreement

- Ravens SPM

JvR PSYCHOMETRICS

www.jvrpsychometrics.co.za

www.psychologyafrica.com

Scope of the agreement:

1. This agreement covers the right to use the Ravens SPM according to these terms in the specified study only.
2. The Ravens SPM will be provided at 100% discount.
3. The Researcher will source the sample group and administer the Ravens SPM to the sample group.
4. The researcher will capture all the data and send this to JvR for scoring.
5. JvR will provide The Researcher with a scored dataset of results on the Ravens SPM for the above study.
6. JvR will provide The Researcher with the scale reliabilities for the Ravens SPM if and when requested.
7. The Researcher will conduct the statistical analysis of the data provided by JvR.
8. The Researcher will provide JvR with a copy of the final manuscript of the study within 2 years of the date of this agreement.
9. Failure to comply with the scope and roles and responsibilities stipulated in this agreement will result in The Researcher being invoiced for the full published price of Ravens SPM materials (price at the time of the infringement).
10. Item level data will not be provided to The Researcher.
11. The Researcher will be responsible for all couriering costs of relevant material.

Ganiyat Zurakat Roles and Responsibilities

1. The Researcher undertakes to ensure that she adheres to the Ethical Guidelines stipulated by the HPCSA, with special attention to the sections on research and psychometric assessment. The Researcher will also ensure that she follows best practice in psychological assessment, such as those described in the *International Test Commission's Test Users Guide*.
2. The Researcher undertakes to respect all copyright and legal rights held by JvR with regard to the security and ownership of test materials. No items of the Ravens SPM may appear in any article, dissertation, thesis, research report or any similar publication that may expose the content of the Ravens SPM items or compromise the security of the questionnaire.
3. Nothing contained in this agreement is intended or shall be construed as granting The Researcher any proprietary or intellectual property right, or any right, title to or interest in the Ravens SPM or any associated data. The Researcher agrees that she will not use data provided pursuant to this request for any purpose other than the analysis of that data as stipulated in this agreement. The Researcher will not

publish or adapt Ravens SPM items or scoring in any form. Violation of this acknowledgement will be a violation of South African copyright law.

4. The Researcher undertakes to ensure that the data provided is held securely and not made available to any third party without prior written consent by JvR.
5. The Researcher will not make any changes or adaptations to the items or appearance of the questionnaire.
6. The Researcher will ensure that acknowledgement to JvR for providing the psychological assessments is given in all research outputs resulting from this study.
7. The Researcher will provide JvR with copies of any research reports related to the use of the Ravens SPM in the study.
8. The Researcher undertakes to inform JvR in advance, should any of the research results related to the Ravens SPM be presented in any professional or public forum.
9. The Researcher will indemnify and hold JvR harmless from any claims, actions, suits, damages, costs and legal fees arising from any breach of The Researcher's obligations specified in this agreement.

JvR Roles and Responsibilities

1. JvR may include the research results in JvR publications, including placing the abstract or summary on JvR's website, in the magazine, or in any other published material. The author(s) will be acknowledged in the publication.
2. JvR undertakes to inform The Researcher in advance, should her research report or results related to the use of the Ravens in the study be presented in any professional or public forum.
3. JvR will be available, within reason, to The Researcher for advice on the appropriate use of the various assessments and interpretation of research results.
4. JvR may include the data in analyses that are performed to support the validity of the products. The data may be published, and/or combined with other datasets, to provide researchers and other professionals with South African information about the product. JvR will ensure that any identifying information is removed from the data before they are used in analyses. No individual person from the data set will ever be identified in publications. Acknowledgement for use of the data will be given to The Researcher.

APPENDIX H: ACKNOWLEDGEMENT OF LPCAT USE

From: Marie de Beer [mailto:marie@mminitiatives.com]

Sent: 08 March 2016 06:10 AM

To: Nicholas Munro; Lientjie Nel

Subject: Masters' student LPCAT training - Name and contact details required

Dear Nicholas

Mondi has agreed for the UKZN Masters' student to attend their training on the 20th of April. They have asked that you please provide the Name and Surname and contact details of the student.

Will you please send this to Lientjie - she will add your contact details and include you in all arrangements so that you are "in the loop".

Since the student will be using the LPCAT for his/her research, if we can arrange that we be given the results of the student's research, then he/she can attend the training without cost. If Mondi needs to charge from their side for refreshments etc. that can be paid by the student.

Kind regards

Marie

Dr Marié de Beer

Cel +27 (0) 82 781 4288

M & M Initiatives

26A Nicolson Street, Bailey's Muckleneuk, Pretoria, 0181

South Africa

Tel: +27 (0)12 460 9726 | Fax: 0866 001 491

APPENDIX I: LPCAT TRAINING CERTIFICATE



Mindset & Managing Initiatives

Members: Chris Marais, tel +27 (0)82 8973529; Dr Marié de Beer, tel +27 (0)82 781 4288; Registered cc CK94/09867/23
26A Nicolson Street, Bailey's Muckleneuk, 0181, tel: +27 (0) 12 460 97 26; e-mail: info@mminitiatives.com

This is to certify that

GANIYAT ZURAKAT

attended a

**Learning Potential Computerised Adaptive Test
(LPCAT)**

training session that was conducted on

20/04/2016

Dr Marié de Beer
Research Psychologist