University of Johannesburg

Faculty of Education

Assessing the culture fairness of an intelligence test by adjusting the test times and pictorial examples: a pilot study with grade 2 learners in four Johannesburg schools.

Samantha Leigh Mary Coppard

920407689

Dissertation submitted for the degree

Magister Educationis

In

Childhood Education

at the

University of Johannesburg

Supervisor: Professor Lara Ragpot

Co-supervisor: Professor Annemarie Fritz-Stratmann

Date: October 2018

DECLARATION

I declare that this dissertation is my own unaided work. It has not been submitted before for any other degree or examination at any other university.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Prof Ragpot, for encouraging me to embark on this academic journey and supporting and guiding me to its fruition. I would also like to thank my co-supervisor Prof Fritz-Stratmann for her guidance and support. I really appreciate it. Thanks to Antje Ehlert for assisting me with the statistical analysis and all your patience with my silly questions.

Thank you to Prof Ragpot and Manoj for offering their home as a quiet escape from the world so that I could focus and finish the study.

Patrick, my father in law, mentor, editor, and cheerleader. Thank you for all the hours, and hours of work you put into this. In the words of A.A Milne, “Piglet noticed that even though he had a very small heart. It could hold a rather large amount of gratitude”. I would not be here, if not for you.

Thank you to my Mom for reading through all 132 pages. As always, you were called upon by one of your daughters, for a ‘not so fun’ task, and you did it with a smile on your face and love in your heart. Thank you.

To my parents, Mom, Daddy, Popsie and Mom Gill, and my siblings, Pam and George. Thank you for encouraging me every step of the way. And for standing in the gap with my children when I wasn’t around.

To my beautiful children, Bella, Charlie and Quaid. You are my everything and I want to thank you for selflessly giving of your time, so I could do something for me.

Dylan, my husband, best friend, and constant support. I would never even have contemplated walking this journey without you by my side. Thank you for being example of where hard work and persistence can make impossible dreams a realty. I love you.

And finally, to my Jesus. I am only here because of Your Grace. And I cannot wait for what other adventures you have planned for my journey. All glory to You alone.

ABSTRACT

Keywords: *Culture fair test (CFT1-R), Cognitive test, Time limit in cognitive test, Language bias, Fluid intelligence*

The HPCSA's Policy on the Classification of psychometric measuring devices, instruments, methods and techniques, warns that it would be "unwise" (p.1) for the assessment profession to not pursue the adaption of existing and development of new, culturally fair tests. Even so, very few culturally relevant tests have been developed in South Africa (Foxcroft, Paterson, Le Roux, & Herbest, 2004). This is despite practitioners becoming more cognisant of the importance of using sound assessments, which maintain their validity across cultural groups (Paterson & Uys, 2005). There is an urgent need for the development and adaption of psychometric assessments in order to assure their validity in a multicultural South Africa.

A test is deemed culturally fair if the test is void of test items that are a source of potential bias. A culture fair intelligence test should therefore accurately test an individual's intelligence level regardless of their cultural or socio-economic background. Many psychologists believe that the idea that a test "can be completely absent of cultural bias" (Benson, 2003, p.1), that is culture free, is not possible. That being said, a culture fair, as opposed to culture free, test is a necessary and vital goal to strive towards should the assessment profession want to confirm to the regulations as outlined by the HPCSA and the employment Equity Act.

In this pilot research project, I argue that the presence of a time limit as well as a formal testing situation could increase test anxiety, and therefore hamper the learner's ability to supply answers that accurately reflect his/her intellectual ability. An adapted CFT1- R was administered to the sample. In order to establish the optimum time limit for each subtest, the following intervention was implemented during

administration. Once the official time for the subtest had lapsed, the administrators marked each child's progress at 30 second intervals. This continued until the child had finished the subtest, at which point the administrator made a note of the total time needed to complete the test.

The results of the data analysis indicated that South African learners may require more time, when completing the CFT1-R, than their German counterparts. Findings also indicated that mother tongue tuition versus second language tuition could possibly influence the child's ability to perform on a cognitive assessment. A further assumption that could be drawn from the findings is the effect of preschool education, both in terms of the child's access to a preschool education as well as the quality of education that was received. These assumptions require further research.

TABLE OF FIGURES

Figure 1. Item 22 from Picture Completion Subtest (Weschler, 2003a). 2
Figure 2. Human Brain Development (Nelson, 2000). 24
Figure 3. The Brain's plasticity across the human lifespan (Levitt, 2009)............................ 26
Figure 4. Fischer's four tiers of development, (Murphy, 2008)... 31
Figure 5. Development of Gc and Gf across the lifespan (Craik & Bialystok, 2006)............ 43
Figure 6.CFT1-R Subtest One, (Weiβ & Osterland, 2013).. 61
Figure 7.Examples from Animal Coding Subtest, WWPSI (Weschler, 2003a). 62
Figure 8. Labyrinth Three and Four from Subtest Two, CFT1-R, (Weiβ & Osterland, 2013).63
Figure 9. An item from Subtest Three, showing how the context of the image remains the same with only slight changes to the characteristics (Weiβ & Osterland, 2013). 64
Figure 10. An example from Subtest Four, CFT1-R, (Weiβ & Osterland, 2013). 64
Figure 11.An item from the Matrix reasoning subtest of the WISC, (Weschler, 2003b)........ 65
Figure 12. An example from Subtest Five, (Weiβ & Osterland, 2013)................................. 65
Figure 13. An item from the Matrix reasoning subtest of the WISC, (Weschler, 2003b)....... 66
Figure 14. An example from Subtest Six, (Weiβ & Osterland, 2013). 66
Figure 15. An item from the Raven's CPR, (Ravens, 1998).. 67
Figure 16.Subtest Six's example page, showing the original three examples and the additional two (bottom two). 69

LIST OF TABLES

Table 1. Allotted time limit per subtest of CFT1-R 68
Table 2. Subtests on CFT1-R........ 70
Table 3: Language and Gender Composition of Study Participants........ 72
Table 4: Gender and Age Composition of Study Participants based on home language 80
Table 5. Subtest One: Independent T-test results comparing the results from the unadapted version, and the version adaptedfor the purposes of this study 82
Table 6.Subtest Two: Independent T-test results comparing the results from the unadapted version, and the version adaptedfor the purposes of this study 82
Table 7. Subtest Three: Independent T-test results comparing the results from the unadapted version, and the version adaptedfor the purposes of this study 82
Table 8. Subtest Four: Independent T-test results comparing the results from the unadapted version, and the version adaptedfor the purposes of this study 83
Table 9.Subtest Five: Independent T-test results comparing the results from the unadapted version, and the version adaptedfor the purposes of this study 83
Table 10. Subtest Six: Independent T-test results comparing the results from the unadapted version, and the version adaptedfor the purposes of this study 83
Table 11. Optimum time limit for all participants based on t-test scores. 84
Table 12. Mean for Study Participants based on home language........ 84
Table 13. Subtest One: Dependant t test scores based on home language (comparing isiZulu, Sesotho and Afrikaans)........ 85
Table 14. Subtest Two: Dependent t test scores based on home language (comparing isiZulu, Sesotho and Afrikaans)........ 87
Table 15. Subtest Three: Dependant t test scores based on home language (comparing isiZulu, Sesotho and Afrikaans)........ 88
Table 16. Subtest Four: Dependent t test scores based on home language(comparing isiZulu, Sesotho and Afrikaans)........ 89
Table 17. Subtest Five: Dependant t test scores based on home language (comparing isiZulu, Sesotho and Afrikaans)........ 90
Table 18. Subtest Six: Dependent t test scores based on home language (comparing isiZulu, Sesotho and Afrikaans)........ 91
Table 19. Optimum time limit for participants of the CFT1-R based on home language (comparing isiZulu, Sesotho and Afrikaans) 92

TABLE OF CONTENTS

DECLARATION ..... **ii**

LETTER OF EDITING ..... **iii**

ACKNOWLEDGEMENTS ..... **iv**

ABSTRACT ..... **vi**

TABLE OF FIGURES ..... **viii**

LIST OF TABLES ..... **ix**

1. Chapter 1: Background and Rationale **1**

1.1. Introduction 1
1.2. Problem Statement 4
1.3. Aims and Objectives 6
1.4. Research Question 6
1.5. Theoretical Framework 7
1.6. Overview of Research Methodology 7
1.6.1. Research Design 8
1.6.2. Sample 8
1.6.3. Data Collection Methods 9
1.6.4. Data Processing and Analysis 9
1.7. Trustworthiness 10
1.8. Ethics 11
1.9. Demarcation of this Study 11
1.10. Conclusion 12

2. Chapter 2: Literature Review: Problematising intelligence **13**

2.1. Introduction to ideas of cognitive development 13
2.2. Philosophy of cognitive development 15
2.2.1. The philosophy of child. 17
2.3. The biology of cognitive development 20
2.3.1. Neuronal development in children. 23
2.3.2. Plasticity 25
2.4. Developmental progressions of intelligence 27
2.4.1. Jean Piaget's general theory of cognitive development. 27
2.4.2. Pascual-Leone's Theory of Constructive Operators 29
2.4.3. Fischer's Dynamic Skills Theory 30
2.4.4. Carey's Origins of Concepts 33
2.5. Culture and intelligence 33
2.5.1. Indigenous knowledge systems and intelligence 34
2.5.2. Language: the product and creator of culture 38
2.6. Psychology and intelligence 39
2.6.1. The history of intelligence and intelligence testing 39
2.6.2. Theories on Intelligence 41

2.6.2.1. Spearman's g Factor. 41
2.6.2.2. Guildford's structure of intellect. 41
2.6.2.3. Cattell's Theory on Intelligence. 42
2.6.2.3.1. Inductive Reasoning and Intelligence. 44
2.6.2.4. John Carroll's three stratum theory of intelligence 44
2.7. The quandary around intelligence testing 45
2.7.1. Psycho-educational assessments 46
2.8. Testing in a South African Context 47
2.8.1. Cognitive Assessments in South Africa 48
2.8.2. The oppressive history of psychometric testing in South Africa 50
2.8.3. South African legislation guiding the use of psychometrics in South Africa. 51
2.8.4. Problematising psychometrics in South Africa 51
2.8.5. Addressing the issue of bias in intelligence testing. 52
2.8.6. Time and test taking 54
2.9. Conclusion 57

3. Chapter 3: Research Design and Methods 59

3.1. Problem Statement 59
3.2. The Culture Fair Test Revision 5 60
3.2.1. Description of the measurement 60
3.2.2. Subtests of the CFT1-R 61
3.2.3. Adaptions made to the CFT1-R. 67
3.2.3.1. Independent variable: Time 68
3.2.3.2. Other Adaptions: Practice items. 69
3.3. Research Design 70
3.3.1. Independent variable (s). 70
3.3.2. Dependent variable (s). 70
3.3.3. Control of extraneous variable (s). 71
3.4. Sampling 71
3.4.1. Participants. 72
3.5. Data Collection Methods 73
3.6. Research Procedure 74
3.7. Statistical Techniques 75
3.8. Research objective(s) 75
3.9. Research Hypothesis 76
3.9.1. General Hypothesis One 76
3.9.1.1. Rationale. 76
3.9.2. General Hypothesis Two. 76
3.9.2.1. Rationale. 77
3.10. Ethical considerations 77
3.11. Conclusion 78

4. Chapter 4: Data Analysis 79

4.1. Introduction 79
4.2. Preliminary Analysis 79
4.2.1. Descriptive statistics 79
4.3. Research Question One 81
4.3.1 T test scores. 81
4.3.3. T test result. 85

5. Chapter 5: Findings and conclusions 94

5.1 Introduction 94
5.2 Research Question One 94
5.3 Research Question Two 96
5.4 Summary of findings 98
5.5 Implications of Research 99
5.6 Limitations 101
5.7 Considerations for Future Research 102
5.8 Summary of the study 103
5.9 Conclusion 105

6. Reference List 107

7. APPENDIX 1A: Example of consent letter sent to parents 119

8. APPENDIX 1B: Ethics clearance 120

9. APPENDIX 2: Original SPSS data outputs 121

9.1 Description of the sample 121
9.2 Description of home language groups too small for statistical analysis 123
9.3 Description of performance per subtest of entire sample 124
9.4 Development of performances depending on test time: Subtest One 125
9.5 Development of performances depending on test time: Subtest Two 126
9.6 Development of performances depending on test time: Subtest Three 127
9.7 Development of performances depending on test time: Subtest Four 128
9.8 Development of performances depending on test time: Subtest Five 129
9.9 Development of performances depending on test time: Subtest Six 130
9.10 Development of performances depending on test time and home language: Subtest One 131
9.11 Development of performances depending on test time and home language: Subtest Two 134
9.12 Development of performances depending on test time and home language: Subtest Three 137
9.13 Development of performances depending on test time and home language: Subtest Four 140
9.14 Development of performances depending on test time and home language: Subtest Five 143
9.15 Development of performances depending on test time and home language: Subtest Six 146

1. Chapter 1: Background and Rationale

1.1. Introduction

Albert Einstein said “Everybody is a genius. But if you judge a fish by its ability to climb a tree, it will live its whole life believing that it is stupid” (cited in Kelly, 2004, Everybody is a Genius section, para. 1). This one sentence encapsulates the dilemma of intelligence testing, not only in South Africa, but on a global scale. In order to obtain an accurate and valid score, it is imperative that we make use of assessment tools that will assess test takers fairly and not discriminate against them based on language, gender or culture.

A test is deemed culturally fair if the test is void of test items that are a source of potential bias (Foxcroft, 2004; Reynolds & Kamphaus, 2003), that is, should a test discriminate against an individual solely due to his/her culture[1], that test would be considered culturally unfair (Foxcroft, 2004; Reynolds & Kamphaus, 2003). A culture-fair intelligence test should therefore accurately test an individual’s intelligence level regardless of their cultural or socio-economic background (Foxcroft, 2011), with as few ‘culturally loaded’ items as possible; these would typically be items which include the testee to have knowledge about a specific cultural practice in order to answer the question successfully. An example of this is item 22 (see figure 1) of the Picture Completion subtest, Weschler Preschool and Primary Scale of Intelligence, WPPSI,

[1] Culture, as used in this study, refers to the community, society or context in which the child lives. Thus, children in South Africa will be exposed to a slightly different culture of beliefs and knowledge systems than children in Finland. Culture in the context of this study does not refer to race or imply any racial connotation.

(2003a). The WPPSI is on the HPCSA's list of classified and certified psychological tests (2017) and is used by psychologists to assess South African children (Foxcroft Paterson, Le Roux, & Herbest, 2004).

Figure 1. Item 22 from Picture Completion Subtest (Weschler, 2003a).

Figure 1 shows two ice skaters; the learner needs to identify that there is a blade missing from the one ice skate. However, ice skating is not as popular an activity in South Africa, as it is in the Northern hemisphere and is not easily identifiable by learners. This may distract learners from the task at hand and prohibit them from offering the correct answer. It speaks for itself that these types of items and tests would be discriminating against the testee on the basis of cultural knowledge and not ability to process information correctly or to problem solve. Similar issues were found by Mawila (2012), when she examined the quality of test items in the Junior South African Individual Scales (JSAIS). Mawila (2012) found that some of the items where outdated and developed specifically for use on English and Afrikaans test takers. She found that the Sesotho children could not relate to some of the images and could therefore not accurately answer the item. Many psychologists believe that the idea that a test "can be completely absent of cultural bias" (Benson, 2003a, p.1), that is culture free, is not possible as even non-verbal tests (thought to be more culturally fair than verbal tests), have verbal instructions (Benson, 2003a; Cole, 2009).

With that being said, it is important to note that both the 1998 South African Employment Equity Act, as well as the South African Health Profession's Council's (HPCSA) legislation, give a clear directive to South African practitioners, outlining their ethical and legal obligation to administer culturally fair tests (Van de Vijver & Rothmann, 2004). Even so, very few culturally relevant tests are available in South Africa (Donald, Thatcher, & Milner, 2014; Foxcroft et al., 2004). This is despite practitioners becoming more cognisant of the importance of using sound assessments which maintain their validity across cultural groups (Paterson & Uys, 2005). Foxcroft et al., (2004) noted that 65.8% of practitioners "indicated that they feel the tests that they use are only sometimes appropriate to use cross-culturally" (p. 20). The majority (58%) indicated that more culturally fair tests needed to be made available in South Africa. Based on these findings, there appears to be a justifiable demand from practitioners for tests that are deemed culturally fair in South Africa.

This study will contribute to the revision/adaptation of the Culture Fair Intelligence Test – Version 1 (CFT1-R) for a sample of culturally and linguistically diverse children in Gauteng, South Africa. The CFT1-R is a non-verbal, intelligence test designed to assess the fluid intelligence of children aged between four and nine years (Weiβ & Osterland, 2013). The CFT1-R is founded on Cattell's theory of intelligence (Carroll, 1984; Cattell, 1987; Kvist & Gustafsson, 2008). This specific test

(CFT1-R) was chosen for the purposes of this study, as it was initially used in the validity study of a South African sample of learners on the Mathematical and Arithmetical Competence Diagnostic test (MARKO-D) test (Fritz, Ricken & Balzer, 2009). The results showed that it could be a possible test for adaptation for South African children, as the CFT1-R tests fluid intelligence and there is very little language used in the test, in comparison to the outdated verbal IQ tests currently available in South Africa (see further discussions in paragraph 2.8.1, Cognitive Assessments in South Africa).

Language is thought to be the fault line along which culture varies and as such, has a considerable impact on test performance in a multilingual South Africa (Foxcroft & Aston, 2006). In a testing situation, language is therefore a source of potential bias. The development or revision of non-verbal test items has been postulated as a possible solution (Foxcroft & Aston, 2006; Nell, 1994; Owen, 1991). Simply translating a test into a mother tongue is in no way sufficient to ensure the validity of the test in a multi-cultural context. And in a multi-lingual society such as South Africa, the task of translating a test into all mother tongues is a complex one (Foxcroft, 2004; Fritz, et al., 2014). When developing new instruments or revising existing instruments, it is crucial to make allowances for cultural differences in a reliable and consistent manner.

1.2. Problem Statement

It is evident that there is an urgent need in the South African context for ethically and culturally sound assessments. This being said, adaptation and or development of new assessments for a multi-cultural, multi-lingual society is both costly and time consuming. It is possible that this is a major contributing factor as to why there are so few, if any, available tests to measure children’s cognitive functioning in South Africa.

Up until the early 1990s, the Human Sciences Research Council (HSRC) was the primary test developer and supplier (Foxcroft, 2004 et al.; Foxcroft & Roodt, 2006). However, it has gradually been receiving less government funding for this kind of test development and has been restructured (Foxcroft et al., 2004). According to Foxcroft et al., (2004), this has left practitioners unsure of what role the HSRC will play in future test development. Various small organisations have “sprung up to develop and supply tests, but there is no body or organization to coordinate test development activities” (Foxcroft et al., 2004, p. 1) and to validate whether tests are culturally fair and reliable. As of 2003, there has been no governing organisation with regards to test distribution, as the HSRC relinquished their role as the central test distributor in South Africa. Since then, the responsibility of test development and adaption has been left with private test distributors, academics and psychologists (de Beer, 2017).

As a psychometrist, I understand, first-hand, the challenges faced by the Education Psychology and Psychometrics fraternities. Practitioners try their best to put together batteries of tests that do not discriminate against children. For example, the Gauteng Education Department has brought out a basic screening assessment that does not make use of an intelligence test. However, a culture fair, as opposed to culture free, test is a necessary and vital goal to strive towards, should the South African assessment profession want to abide by the regulations as outlined by the HPCSA and the Employment Equity Act. The aim of this study is to get practitioners in South Africa one step closer to attaining that goal, by starting the process of validating an assessment tool for the South African population.

1.3. Aims and Objectives

The overall aim of the study is to establish how the cultural fairness of an intelligence test, can be notably advanced/adjusted by modifying the test times and the pictorial examples used in the test. To realise the aim[2] of the study, the following objectives are set:

- To establish, from the literature, what constitutes a culturally fair test.
- To determine what the optimal subtest time limits are on the CFT1-R for a small sample in Johannesburg, South Africa.
- To determine whether home language is an important variable when determining optimal test time limits.
- To determine whether the addition of pictorial examples is an important variable when determining cultural fairness.

1.4. Research Question

The above stated discussion of the identified problem has led me to the formulation of my research question.

Can the cultural fairness of an intelligence test be advanced/adjusted by modifying subtest times and providing additional pictorial examples?

---

[2] As this is a study that may be of a sensitive nature, I wanted to include this section. I have stated the aims and objectives above, but I want to demarcate very clearly what I am *not* trying to accomplish with this study. It is important to note that the main objective of this study is not to criticise the assessment fraternity and I am cognisant that they are doing the best they can under the circumstances. Furthermore, I am not attempting to find an underlying concept of intelligence that would discriminates against certain individuals.

1.5. Theoretical Framework

The theoretical field in which this study is located is chiefly that of the psychometric approach, which can be defined as:

> "A theoretical perspective that portrays intelligence as a trait (or set of traits) on which individuals differ; psychometric theorists are responsible for the development of standardised intelligence tests" (Benson, 2003a, p. 301).

Using the above as my focus, I will investigate the concept of intelligence and cognitive development within this broad lens. Moreover, I will focus more specifically on the work of Raymond Cattel as his intelligence theory is the theoretical foundation for the CFT1-R. The CFT1-R is a timed, cognitive assessment which measures fluid intelligence (Weiβ & Osterland, 2013). It is grounded in Cattell's Theory of Intelligence, which describes intelligence as having two traits, namely crystalized and fluid intelligence (Cattell, 1987; Weiβ & Osterland, 2013). A more in-depth discussion on intelligence theories, cognitive development theories and testing and culture can be found in Chapter Two (see paragraph 2.6, Psychology and intelligence).

This theoretical framework will give this research its intended focus, guiding the way that I investigate the literature and collect and analyse the data. (Pollard, Johnston, Dixon, 2007).

1.6. Overview of Research Methodology

This is a pilot study, intended to determine whether the adaptation of an intelligence test would affect it culture fairness. The following research design and methodology were selected.

1.6.1. Research Design

A quantitative research design was implemented by sampling participants in the same grade (grade 2) across different schools from similar low socio-economic backgrounds and different languages of instruction, in urban Johannesburg, over a two-week period. Taylor and Yu (2009) note that in South Africa socio-economic status (SES) has historically been dictated by race rather than merit. They go on to link a leaner's SES to the quality of education they receive. As a result, access to education does not necessarily transform social inequalities, but may replicate the model (Taylor and Yu, 2009).

The CFT1-R instructions had been translated from German into English, isiZulu, Sesotho and Afrikaans in initial studies (n=200) that accompanied the standardising procedures of another test that is studied in the Centre for Education Practise Research at the University of Johannesburg. As already mentioned, it is presupposed that simply translating the instructions is in no way adequate when revising a test. In addition to the translation (and for the purposes of this pilot study), this proposed research will concentrate on a further two aspects of the test to improve its culture fairness. These are:

Adding two more examples per subtest to further ensure the testee clearly understands what is required of him/her (Foxcroft, 2011).

Adding extra time per subtest to establish what the optimum test time for a South African testee would be (Mandinach, Bridgeman, Cahalan-Laitusis, & Trapani, 2008; Portolese, Krause, & Bonner, 2016).

1.6.2. Sample

The population was grade two pupils attending schools in and around Johannesburg. The population was comprised of four homogenous subgroups based on their home language; specifically, isiZulu, Sesotho, English and Afrikaans. These subgroups were chosen so as to replicate the subgroups used during the initial administration of the original CFT1-R in the mentioned pilot of the MARKO-D test validation in 2014. One hundred and twenty grade 2 learners, 30 from each subgroup, were selected, using convenience sampling methods. It should be noted that this sampling method may cause the data to be less credible in terms of generalisability (Chaturvedi, n.d; Marshall, 1996 Miles, Huberman, & Huberman, 1994; Creswell & Clark, 1997).

(Miles, Huberman, & Huberman, 1994; Creswell & Clark, 1997).

1.6.3. Data Collection Methods

The adapted CFT1- R was administered to the sample in groups of ten learners per sitting. The researcher administered the CFT1- R to the English-speaking learners. The isiZulu, Sesotho and Afrikaans tests were administered by experienced test administrators who could speak the other three subgroup languages, thus decreasing the possibility of language bias. I was present at all testing opportunities and supervised all the test administration.

1.6.4. Data Processing and Analysis

Once the test was administered, it was scored, and the raw scores were inputted in an Excel spreadsheet. The normed scores were not determined as there are no South African norms available for the CFT1-R. This, in turn, was inputted into the Statistical Package for Social Sciences (SPSS) to run the statistical analysis.

With regards to the possibility of increasing the time limit, data was correlated to establish the mean test time for each subtest, thus giving an indication if more time needed to be allocated in the revision of the subtest. An independent *t-test* analysis was used to compare standard time limit between the original CFT1-R and that of the research sample.

A dependent t-test analysis was used to establish whether there is significant variation in optimal test time within the sample in terms of home language (General

Hypothesis Two). In order to achieve this the test times of each language group where compared with one another. I analysed the increased performance per 30 second interval for each of home language groups.

1.7. Trustworthiness

Trustworthiness refers to the value or worth of a research project. Andres (2012) notes that research is deemed trustworthy if it produces data that answer the research question and if the sample is an accurate reflection of the population at large. In quantitative research, the term validity and trustworthiness are used interchangeably (Andres, 2012; Efron & Ravid, 2013).

In order to ensure trustworthiness, the ethical guidelines regarding the use of psychological assessments were strongly adhered to. These include safeguarding the integrity of the assessments, strictly adhering to the administration instructions of the assessment and ensuring that rigorous attention was paid to the scoring of assessments (Foxcroft & Roodt, 2013). Furthermore, extraneous variables were identified and controlled to the best of the test administrators' abilities (see 3.4.3 Control of extraneous variables). This increased the internal validity of the research (Efron & Ravid, 2013). As this research project involved children in very diverse settings, some extraneous variables were not anticipated and will be discussed in paragraph 5.6, Limitations. In terms of external validity, convenience sampling methods (see paragraph 1.6.2, Sample) was selected. As the sample population was so large, the convenience sampling method means that the sample more accurately reflects the sample population, therefore increasing external validity (Andres, 2012; Efron & Ravid, 2013).

It has to be noted though that the sample is not a true representative sample of the entire South African population but is rather representative of the heterogenous languages groups I researched.

1.8. Ethics

I adhered to general guidelines of ethical research in this study. I have obtained ethical clearance from the Research Ethics Committee of the Faculty of Education at the University of Johannesburg (Appendix 1B).

In terms of ethical conduct with participants, since this study is located in the practice zone of psycho-educational assessment, the guidelines for ethical practice in this field will be strongly adhered to according to the guidelines set by the Health Professions Council of South Africa (HPCSA, 2006b). These guidelines include informed consent from the parents of minors, informed assent from the participants themselves and the maintenance of confidentiality (Fleet & Harcourt, 2018; HPCSA, 2006b; Sattler & Hoge, 2006).

A letter introducing the study was sent to the parents or legal guardians of potential participants, which needed to be signed and returned to me. The notion of voluntary participation was communicated to all participants and the right to withdraw at any time was clearly conveyed. This is in line with Fleet & Harcourt (2018), who argue that receiving consent from the child is a vital step in researching children. The principal of anonymity was maintained by using an identification code (viz. 001-120) for all other participants. All assessment results and other data will be locked in storage and no one will have access to the data except for my supervisor and me.

1.9. Demarcation of this Study

The study consists of the following chapters:

Chapter One serves as an introduction to the study intended to orientate the reader to the context and theory that underpins the research. It will also give a brief overview of the study.

Chapter Two contains a detailed literature review, presenting and synthesising the most recently theory around intelligence, cognitive development and problematising psychometrics in South Africa. A review of the theory that has led to the development of the assessment instrument in this study is also presented in detail.

Chapter Three presents the research design and methodology that was utilised in this study. Since the data collection in this study was done by using an assessment instrument, a detailed description of how the instrument works has been given.

Chapter Four presents the results of the assessment and an analysis of the results according to the data analysis methods described in chapter 3.

Chapter Five will conclude the study with a discussion of the results outlined in chapter 4, a summary, a discussion of limitations and will suggest recommendations for further research.

1.10. Conclusion

This chapter has described the background and rationale for this study and presented an overview of how the research will take place. The next chapter will present a detailed discussion of the theory that underpins this study by reviewing and synthesising literature in the field of intelligence and intelligence testing.

2. Chapter 2: Literature Review: Problematising intelligence

2.1. Introduction to ideas of cognitive development

For those of us who are involved in education and seek to offer appropriate interventions to learners, we need to have a clear understanding of where learners' difficulties lie. Psychometrics was founded on this premise. In 1900, a French professional group for child psychology, La Société Libre pour l'Etude Psychologique de l'Enfant, approached Alfred Binet to develop a measurement tool which would allow the educators to identify where a child has processing difficulties, so that the educators could facilitate the appropriate support (Shaffer, 2002). In order to develop a measurable intelligence quotient (IQ) which can be assessed in scholastic assessments, the concept of intelligence needs to be defined. However, quantifying an abstract concept, such as intelligence is problematic and has resulted in a large variety of definitions and theories on this topic (Shaffer, 2002). It is these kinds of shifting definitions that have contributed to the strong association between humanities and qualitative research (Bond & Fox, 2015).

The aim of this chapter is to explore whether there is information in literature about a more viable option to the cognitive measures we currently have available in South Africa, to assess children's cognitive functioning. To achieve this, I need firstly to examine what has been done historically as well as current global trends in terms of the assessment of cognition. I will then explore the above in a South African context, looking to answer questions such as 'what issues do we have in South Africa in terms of cognitive assessment', 'what can we do differently' and 'why do we need South African assessment tools'.

The desire to define, understand and quantify intelligence is not a new phenomenon, nor is it specific to a certain discipline (Cattell, 1987). A simple search on any Internet search engine reveals society's quest and curiosity about cognitive development. The Internet and social media are rife with quick quizzes and apps that 'assess' one's IQ (https://www.test-iq.org; https://iqtestnow.org; https://intellitest.me). Answers and explanations are a rich melting pot of both scientific theories and untested viewpoints borne out of their implicit theories and understandings. Cognition is not something that is easily defined, because any definition is dependent on one's viewpoint and ontological orientation. Intelligence cannot be defined in singularity, but rather it is made up of differentiated views (Shaffer, 2007). Everyone's view or understanding will be influenced by their personal philosophy as well as their exposure to a certain discipline (Mpofu, 2002).

For the purposes of this study I want to look at different disciplines and how they view childhood cognitive development. With all the contributing views on intelligence, I will argue that one cannot look at one view in isolation, but rather need to consider the perspective and influence of various aspects to gain a more holistic understanding. To answer the questions on why we need psycho-educational assessments, one must firstly look at how we define intelligence. I start with the perspective of broad philosophical viewpoints, followed by a hard science position, not only looking at philosophical views, but also neuro-scientific properties that influence intelligence. Thirdly I will look at the psychological viewpoints on intelligence as well as the history of how intelligence has been defined. Following on from this—and possibly the most important in terms of the purpose of this research—I will look at how culture influences our understanding of intelligence. Lastly, I will look at intelligence testing, on a global stage and then, more specifically, in the multi-cultural context of South Africa.

The philosophical view on cognitive development will be the starting point for this chapter. It was the early philosophers who first attempted to gain an understanding of cognitive development, and in so doing laid the foundation for the study of intelligence (Princiotta & Goldstein, 2015).

2.2. Philosophy of cognitive development

The desire to understand human intellect has fascinated scholars for centuries (Cattell, 1987). Philosophical scholars in early history are no exception, the most noteworthy of these, being the Greek philosopher, Plato. Plato's viewpoint of intelligence "set an integral foundation for our current conceptualizations of intelligence" (Princiotta & Goldstein, 2015, p.83).

His view of intelligence was said to have been inspired by looking up at the stars. Plato argued that it was not possible for all that beauty to have been achieved haphazardly but was rather because of the deliberate influence of an intellect or *nous* (Carpenter, 2010). Plato argued that the beauty of creation/being able to create implied the existence of a divine intellect who thought the world into creation (Carpenter, 2010). In *Timaeus*, Plato stated that "there can be no wisdom [*sophia*] or intelligence [*voũv*] without soul…It is impossible for intelligence to be [or arise] without soul" (as cited in Carpenter, 2010, p.40).

For Plato, there was an undeniable link between human intellect and the soul**.** He gave an analogy of the intellect as being the charioteer that must guide the soul by controlling two horses. A white horse representing rationality and moral impulses and a dark horse representing irrational impulses and urges. He believed humans were in a constant state of conflict between their “appetites and emotions” (Princiotta & Goldstein, 2015, p. 85). According to Plato the “supreme goal of life” (Princiotta & Goldstein, 2015, p. 85) is for a human to “free the soul as much as possible from the adulteration of the flesh” (Hergenham, cited in Princiotta & Goldestein, 2015, p.85) Plato viewed intellect as the means by which individuals would resolve the conflict within themselves and ultimately free their truest part, the soul, from the confines of their humanness (Carpenter, 2010; Cattell, 1987; Princiotta & Goldstein, 2015; Shaffer, 2007).

In terms of a more modern outlook on intelligence, one can find a view on the quantification of intelligence from the field of physics. As briefly outlined in the introduction, one of the major controversies with intelligence testing is the issue of quantifying an abstract concept. Theoretical physicist, Kaku addresses the need to quantify intelligence in all life forms. In *the Origins of Intelligence* (2010), Kaku sets out to quantify consciousness and intelligence. To claim that humans are the only form of intelligent life is ignorant and misguided (Kaku, 2014). Should one take the view that man’s technological advancements are indicative of our superior intelligence, one could argue that, for example, the fact that we can fly, that we have built flying machines, demonstrates our intellect as a species. However, should you look at the seed of a maple tree, it is the exact shape of an aircraft’s aerofoil. The maple tree has been producing aero dynamically perfect ‘wings’ long before the Wright brothers took to the sky. There is also an argument that there is a form of intelligence present in this case, albeit on a less obvious scale than seen in an animal species (Kaku, 2010).

Based on his argument, Kaku would see the maple tree has having roughly 10 units of consciousness, as it is able to sense temperature change and adjust to its environment etc. Kaku notes that the maple tree has evolved a seed that has a shape that will enhance its ability to fly. Therefore, the maple tree has a level of intelligence. Kaku goes on to state that according to his model, the most intelligent 'thing' we know of is the universe itself (Kaku, 2014). In very broad terms, perhaps intelligence is not limited to and for the exclusive use of the human species, but rather it is a universal concept, that we—as human beings—have managed to harness and/or develop to a greater extent than other species. While the viewpoint in this section is speculative, it is worth keeping in mind. It shows that, in terms of this study, that the concept of intelligence stretches beyond the realm of education. Adopting a single view on the subject will restrict our ability to measure it.

The philosophical view on intelligence is quite broad, looking at a universal rather than individual intelligence. This broad view forms a foundation from which to launch a more specific study into the variances in individual intelligence. And it is these variances that psychometrics sets out to measure. For the purpose of this study, which looks specifically at the ability to measure intelligence in children, I will now narrow my scope to look at the philosophy of child.

2.2.1. The philosophy of child.

The literature on intelligence, from a philosophical perspective, dates back centuries. However, there is a notable absence of discussion relating specifically to children (Gopnik, 2009). Alison Gopnik is a professor of psychology at the University of California and contends that, children are "both profound and puzzling and this combination is the classic territory of philosophy…Yet you could read 2,500 years of philosophy and find almost nothing about children" (2009, p. 5). Over the past 30 years the scientific world has seen what has been described as a scientific revolution in terms of our understanding of babies and children. This, in turn, has seen philosophy including children in their quest to understand how and why we are the way we are

(Gopnik, 2009). This trend will surely mean that, when it comes to cognitive development testing in children, test developers will have a more inclusive understanding of how a child thinks, from a philosophical perceptive.

A fundamental difference between the human species and other animals is our ability to change. Not only do we change our world, but we change ourselves and others around us. Carey, who has specialised in exploring issues around the development of the human mind, specifically in children, argues that while there are striking similarities between humans and certain animals when it comes to cognition, problem solving, rationality and intelligent thought, there are also vast differences (Carey 2009; Lupyan 2015; Spelke 2003). This is echoed by cognitive psychologist Elizabeth Spelke. Spelke (cited in Lupyan, 2015) states that:

"Although all animals find and recognize food, only humans developed the art and science of cooking. [A]ll animals need to understand something about the behaviour of the material world to avoid falling off cliffs…but only humans systematize their knowledge as science…all social animals need to organise their societies, but only humans create systems of laws and political institutions to interpret and enforce them" (p.277).

It is our knowledge that distinguishes us, as human. Understanding how we attain knowledge, manipulate it and use it to offer alternatives in our world is key to understanding how we can access it. Gopnik (2009) argues that by studying how children attain, manipulate and use knowledge, we can explain how and why the human species effect change. The study of children will assist philosophers and scientists alike in answering fundamental questions, including those pertaining to intelligence (Gopnik, 2009). And in so doing offer a more valid means to measure intelligence.

The question that has led philosophers to look at ideas dating back to Plato, has been 'How can we know so much about the world?' The study of children, how their minds are changed by the world and, in turn, how it changes the world, seems to be essential to answering that question. The scientific methods of experimentation and statistical analysis seem to be programmed into the minds of young babies. Gopnik found that the steps followed by infants when testing their understanding of the world are similar in nature to those a researcher would use. An infant will develop a hypothesis, test the hypothesis and then accept the hypothesis or develop a new hypothesis that will explain test results. Gopnik (2009, p.108) notes that "very young children unconsciously use these techniques to change their causal maps of the world. Those programs allow babies, and so the rest of us, to find the truth". Imaginative play allows children to learn not only how to create causal structure in their world, but importantly how to visualise the possibility of a new world or scenario. And it is this ability that allows them to effect change. The knowledge about our power to influence change is one that is developed when a child creates an imaginary world with imaginary people. Before a child can offer counterfactuals, they must understand the causal structure of that specific environment. Gopnik (2009) states that it is this knowledge that makes "creativity possible" (p.49). To make changes or offer alternatives one needs to know how events are connected. It is our knowledge about the world that underpins our ability to change it (Gopnik, 2009).

When looking at the evolution of intelligence from a philosophical perspective there are two main theories. One, that it is the understanding of the *physical* cause, which in turn has allowed us to develop and use complex tools. And the other being an understanding of the *psychological* cause, which has allowed us to grasp complex social networks and develop culture (Gopnik, 2009). Both are evident when looking at the evolution of intelligence and both are dependent on an individual's ability to understand causal structure—an ability fine-tuned in the mind of a child (Gopnik, 2009).

The philosophical view on intelligence shows that on a universal scale the desire to quantify intelligence exists even when looking at non-human intelligence. On an individual scale, understanding how a child develops cognitively is key. It offers insight that is vital in defining and quantifying the development of intelligence in children. Furthermore, a child's intelligence sets the foundation for adults' intelligence. This understanding allows us to more effectively define a measurable intelligence.

The cognitive development of children is a keystone of this research as I set out to problematise our ability to access cognitive development in children. The next step to developing a better understanding of the cognitive development in children is to take cognisance of the scientific view of cognition and cognitive development.

2.3. The biology of cognitive development

Cognitive development cannot be looked at in isolation as there are various mechanisms that can affect how it is developed and how it is expressed. I have briefly outlined how cognitive development can be explained from a philosophical lens. Next, I will take a biological perspective, in terms of what the brain looks like, brain functionality and in which way, if any, this affects intelligence. When trying to understand the concept of intelligence, one cannot ignore the biological effects. The brain is the organ that is primarily associated with cognitive development, with the pre-frontal cortex being described as the seat of intelligence (Shaffer, 2002).

To begin this process of understanding the biology that underlies intelligence, my first question is, 'Is there a link between the physical structure of an individual's brain and their cognitive skills?'. Einstein was arguably one of the most intelligent men of our time (Shanks et al.,2013) and upon his death, his brain was removed from his body, and subsequently stolen, to be studied by scientists to explain his superior intellect (Weiwei et al., 2014). Recently, photographs of his brain were discovered and examined by Professor of Anthropology at Florida State University, Dean Falk. Upon examining the pictures, Falk concluded that Einstein's corpus callosum was larger than compared with those of controls. While this cannot account solely for Einstein's high intelligence levels, Men et al., (2014), argue that it was most certainly had to be a contributing factor. Other differences thought to have influenced his increased visuospatial and mathematical ability as well as his predilection for thought experiments, was the fact that he had a higher than normal number of glial cells as well as a large prefrontal cortex (Men et al., 2014). This suggests a hard link between the structure of an individual's brain and their intelligence levels. What our brains look like influences the cognitive skills that cognitive tests set out to access. Variations in structure could account, to some extent, for variations we see in test scores. These variations have been supported by research on individuals with average intelligence. Andreasen et al., (1993) found that there was a significant, albeit moderate correlation between the size of the cerebral structure and intelligence. Jung & Haier, (2007) note that modern neuroimaging can explain the biology of intelligence.

John Duncan, the programme leader at the Medical Research Council Brain and Cognition Unit, Cambridge found that it would seem that intelligence has a neural basis (Duncan, 2001). More importantly, it appears that it is our frontal neurons, of which the human species has the highest proliferation when compared to other species, which enable intelligent behaviour. Furthermore, Duncan (2001) noted that these neurons were more flexible in functionality than those found in other parts of the cortex—"neurons in selected frontal regions adapt their properties to code information of relevance to current behaviour, pruning away all that is currently task-irrelevant" (Duncan, 2001, p.3) This indicates that there is a cluster of neurons that activate when intelligent behaviour is required, regardless of whether it is mathematical, reasoning skills, or literacy etc. Duncan (2001) suggests that individual differences in intelligence could be found in the differences of the structure and function (increased number of glial cells) of the prefrontal cortex. This would appear to support the findings of Falke (Men et al., 2014) that Einstein's intelligence could be partly attributed to the physical size of his prefrontal cortex.

When scoring a cognitive assessment, the probability of finding two identical results is highly unlikely. The aim of assessments is to quantify individual variances. As stated above, brain structure accounts for some of the individual variances found in intelligence tests. This is because no two brains will be identical. The reason for this is twofold: firstly, due to the "nonlinear process involved in neuronal morphogenesis" (Geake, 2008, p.2). As with any of our physical features, our brains do not develop as exact replicas of our parents. Our genetic code, which we inherent from our parents, influences everyone's neuronal morphogenesis. Secondly, no two individuals will have the same life experiences, and therefore will not have an identical neural development, not even paternal twins can share identical experiences. On the one hand our genetic code influences how, our brain will develop. On the other hand, our life experiences—especially in the early years of childhood—influence neural development processes such as neural pruning (Geake, 2008; Shaffer, 2007). Factors such as exposure to harmful chemicals in utero, malnutrition and exposure to physical traumas, to name a few, may negatively influence the neural development processes and ultimately intelligence levels (Geake, 2008; Grossberg, 2000; Neisser et al., 1996; Shaffer, 2007). It goes without saying that variations in our individual intelligence levels are rooted both in our biological as well as environmental variances (Shaffer, 2007).

To gain understanding of children's cognitive development it is important to not only be aware of the influence of biology in a broad sense, but to consider biology expressly in terms of a childhood.

2.3.1. Neuronal development in children.

When looking at assessing cognitive development and the problems faced when trying to assess it, I wanted to also look at how cognitive skills develop. During the seventh to the eighteenth weeks of gestation, neurogenesis—development of the neurons—is at its peak (Elliot, 1999). Synaptogenesis, the development of synapses, begins at around 28 weeks; however, this only peaks during the first two years of life. During this time, 83% of dendritic growth occurs. Another characteristic of postnatal synaptogenesis is the fact that the cerebral cortex produces twice as many synapses as is needed. This over-production leads to competition between the synapses (Gazzangia, Ivry, Mangun, 2002; Shaffer, 2007). Experiences with the environment solidify certain connections while those that are weak and have not been established or are no longer needed (i.e. the sucking reflex) undergo a process of neuron elimination, commonly referred to as pruning. The way in which our brains prune or discard connections that are weak and have not been established, points to the importance of stimulation and education in early childhood cognitive development. (Elliot, 1999; Gazzangia et al., 2002; Shaffer, 2007).

By birth, the baby has an almost anatomically complete brain, apart from myelination of the axons in the brain (Shaffer, 2007). Figure 2 shows how the first year of a child's postnatal development is key for the establishment of neural connections used in the sensory development as well as language development (Nelson, 2000). Once these have been established, there is a steep increase in the development of the neuronal connections required for cognitive functioning.

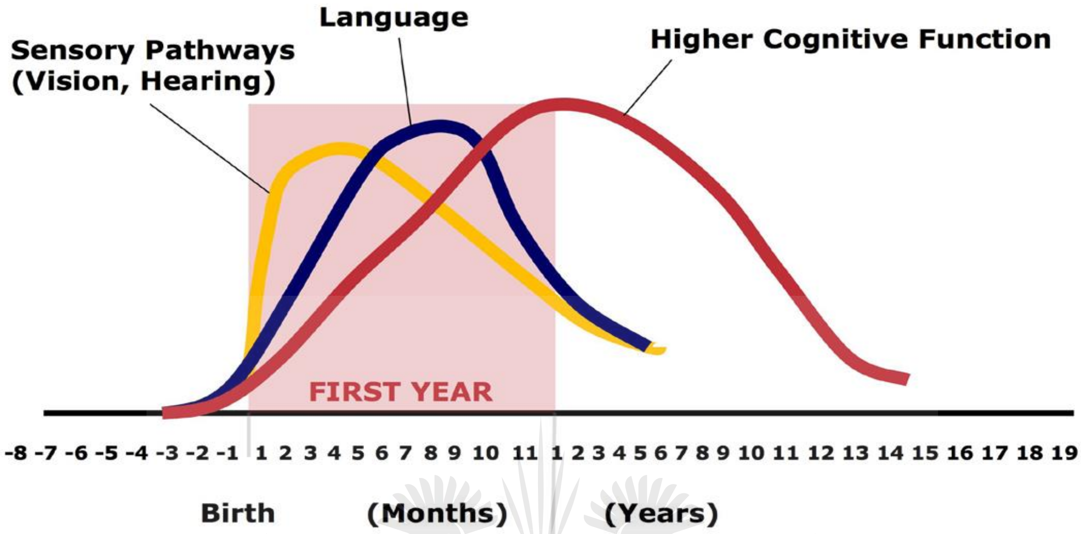

Figure 2. Human Brain Development (Nelson, 2000).

The sensitive period for this development falls between years one to six, before tapering off in middle adolescence (Elliot, 1999; Gazzangia et al., 2002; Shaffer, 2007). It is during these formative years that practitioners will start to make use of cognitive assessments. However, it is imperative that the assessment tools that we have at our disposal are valid and reliable. Any support that is needed will be more effective if it is delivered while the brain is still pliable.

By the age of six years old, a child's brain has increased significantly in volume (Gazzangia et al., 2002). This can be attributed to the myelination of neurons as well as the proliferation of glial cells. Research into this aspect of brain development has shown that white matter develops in a linear pattern and development does not differ across areas of the brain (Gazzangia et al., 2002; Haier, White & Alkire, 2003). On the other hand, grey matter development showed a preadolescent increase as well as a post adolescent decrease in development. Furthermore, unlike white matter, grey matter development is not the same across cortical regions. Haier et al. (2003), noted a high correlation between the volume of grey matter in the frontal cortex and intelligence scores across ages. He also noted that a high IQ predicated high functioning on non-IQ questions. Intelligent people think about everything differently. It has been proposed that this could be why gifted children enjoy socialising with other gifted children (Craik & Bialystok, 2006; Geake, 2005; Haier et al., 2003).

The early years of a child's life show a marked increase in both the volume of the brain, as well as the neuronal development that takes place. It goes without saying that these changes are influenced by the environment experienced by the child. In turn, these changes will influence the child's cognitive development and ultimately their intelligence levels. How these changes are made is referred to as plasticity (Fischer & Silvern, 1985). It is no coincidence that psycho-educational assessments are generally administered around six years of age. The process of plasticity, which is discussed below, is the reason that interventions can assist a child's cognitive skills. And the main purpose behind psycho-educational assessments is to distinguish what interventions are needed.

2.3.2. Plasticity

The information practitioners receive from a cognitive assessment highlight where a child is experiencing difficulty in terms of cognition. However, this information is worthless if it is not used to correctly identify interventions that will assist the child where difficulties are identified. Plasticity can be defined as the brain's ability to change its form or structure, usually because of experiences and learning (or interventions). As with intelligence, plasticity is not a uniquely human trait. Nonetheless, the brain of a human is thought to be more plastic than those of other species. Within the human species there is further variation in neuroplasticity, dependant on the age of the individual (Fischer & Silvern, 1985; Gazzangia, et al., 2002).

An adult's brain is fairly rigid in terms of how it can adapt to change. The brain of a foetus, on the other hand, is exceptionally malleable, allowing the brain to develop. While postnatal plasticity is limited in comparison to prenatal plasticity, it is evident from the figure below (Figure 3) that the brain of a child—in the early developmental years—exhibits greater plasticity and requires less effort to do so. Gazzangia et al. (2002), suggest that this can be seen when one looks at the time needed by a child to recover from brain trauma in comparison to that of an adult (Gazzangia et al., 2002; Gopnik, 2009).

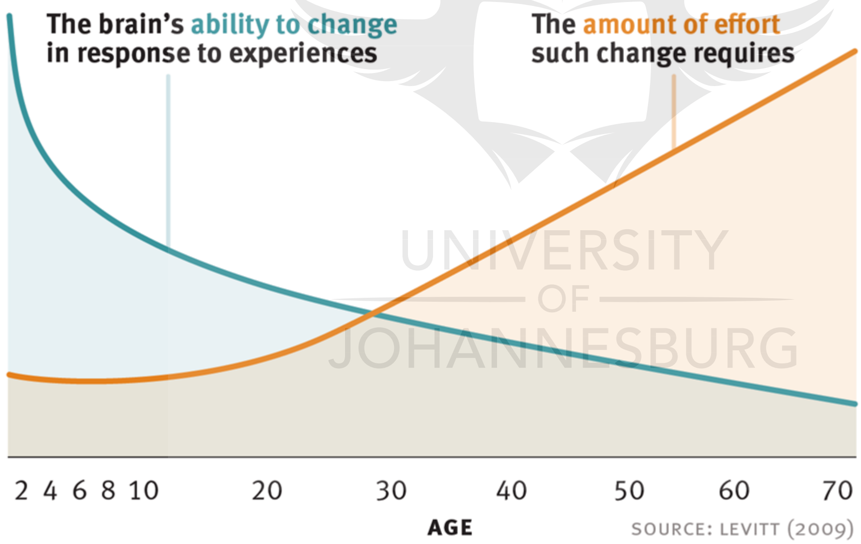

Figure 3. The Brain's plasticity across the human lifespan (Levitt, 2009).

The process of learning is key to plasticity. Gopnik (2009, p. 129) argues that if plasticity is indicative of an individual's ability to pay attention, or have a "vivid awareness", then one can infer that babies and toddlers are more open to learning experiences than an adult. Once again this points to the fact that early childhood is a critical time in cognitive development, and subsequently intellectual development in humans.

Plasticity, along with neuronal development, seems to be the key feature when understanding how children develop cognitively. It is through the process of plasticity that interventions can assist children with difficulties. And it is through psycho-educational assessments that practitioners can identify what interventions are needed. Furthermore, intelligence implies cognitive development, so to understand how intelligence develops in a child one must look at how intelligence's cognitive foundation is developed (Cockcroft, 2004b). This will be examined in the next section.

2.4. Developmental progressions of intelligence

As stated earlier, intelligence is interconnected with cognitive development. The reason quite simply being that intelligence, as a concept, encompasses the various cognitive abilities, including, for example, attention, memory and language. Cognitive development, especially in children, therefore, has a direct link to the subject of this research, which is cognitive assessments of children in South Africa. A basic understanding of the major cognitive developmental theories is key when attempting to develop a holistic view of intelligence, especially in terms of children (Cockcroft, 2004b).

2.4.1. Jean Piaget's general theory of cognitive development.

Piaget, the 'father of cognitive development' offers us the most well-known developmentally based conception of intelligence. Piaget's theory is considered a constructivist theory in the sense that he argues that a child constructs knowledge and meaning from their interactions with the environment, "through the continually shifting balance between the assimilation of new information into existing cognitive structures and the accommodation of those structures themselves to the new information" (Neisser et al., 1996). According to Piaget's theory, a child's cognitive development is stage specific, that is, each child develops through a set of specified stages. Piaget outlines four major stages that a child would develop through. Between the ages of zero to two, a child is in the Sensory Motor Stage (Shaffer, 2007). This stage is defined by a child developing the ability to differentiate themselves from other objects. The child will now see themselves as an agent of action, recognising that their actions influence objects, causing them to act intentionally. They also develop object permanence which enables them to understand that an object continues to exist even if it is out of sight. Between the ages of two and seven years old, a child is in the Pre-operational stage. This stage shows the development of a child's language abilities. They begin to be able to represent an object by using a word or an image (Cockcroft, 2004b; Piaget, 1952; Shaffer, 2007). Furthermore, a child can classify an object based on one salient feature. Egocentrism, the inability to understand a different perspective on something, other than your own, is still evident in this stage, albeit greatly improved (Cockcroft, 2004b, Piaget, 1952, Shaffer, 2007). The third stage of cognitive development is the Concrete Operational Stage. This stage is between the ages of seven to eleven. A child in this stage is characterised by the ability to engage in logical reasoning when faced with concrete tasks. Classification has also improved as a child is now able to classify an object based on various features (Cockcroft, 2004b; Piaget, 1952; Shaffer, 2007). The final stage of development is the Formal operational stage from eleven years through to adulthood. Cognitive reasoning skills now enable a child to think logically about abstract concepts, allowing them to think hypothetically, or ponder ideological questions (Cockcroft, 2004b; Piaget, 1952; Shaffer, 2007).

One of the most obvious criticisms of Piaget's theory is that it is domain general, that is, Piaget argued that cognitive maturation occurs concurrently across all domains of cognitive ability. Furthermore, Piaget paid little attention to individual variances in cognitive development as well as the effect of social interaction in cognitive development (Cockcroft, 2004b; Neisser et al., 1996; Santrock, 2005; Shaffer, 2007). These issues where addressed by neo-Piagetian theorists such as Pascual-Leone and Fischer. Neo-Piagetian theorists integrated the stages as outlined by Piaget with more recent concepts in cognitive development (Santrock, 2005). While these theories are stage specific, they are not domain specific and, as a result, address two major criticism of Piaget's work. Namely the issue of forwarding an explanation of individual variances in cognitive development as well as offering an alternative to Piaget's domain general theory (Santrock, 2005; Shaffer, 2007).

2.4.2. Pascual-Leone's Theory of Constructive Operators

Pascual-Leone's Theory of Constructive Operators explains cognitive processes as operating on two, hierarchical levels (Miller, Campbell & Juckes, 1987). The first level is defined by situational specific constructs. As with Piaget's schemas, these constructs are activated by an external cue, "when the features of reality correspond to the qualitative properties of the scheme it will apply" (Todor, 1979, p. 315).

The second level is comprised of situational-free meta-constructs. These constructs apply to the first level schemes allowing more probability to activate. The second-level meta-constructs work on the first level schemes themselves, but not on the initial input themselves. There are a few meta-constructs described by Pascual-Leone, but the one that is of significance to this research is the *M*-operator. This meta-construct "may be thought of as a quantification of an individual's cognitive capacity"

(Todor, 1979, p.316). *M* is described as the mental energy that the individual requires to activate or keep information 'open' in one's mind, which equates to working memory. While Pascual-Leone advocated a simple explanation of *M*, it is evident that this construct is complex in nature. The *M*-operator develops according to Piaget's stages of cognitive development. A child's *M* capacity is said to develop according to Piaget's stages of development and *M* is thought to play an active role in a child's ability to develop from one stage to the next. As *M* develops, it increases the number of mental units a child can hold in his/her mind at any given point. This will also increase the complexity of task that can be successfully completed. (Miller et al., 1987; Todor, 1979).

2.4.3. Fischer's Dynamic Skills Theory

Another prominent neo-Piagetian theorist is Fischer. Fischer's Dynamic Skills Theory is similar to Piaget's theory in the sense that he describes four stages (see figure 4) in cognitive development; namely the reflexes tier, the sensorimotor tier, the representational tier and the abstract tier (Fischer & Silvern, 1985). The reflexes tier deals primarily with the basic reflexes that are established in the first months of a child's life. The sensorimotor tier is like Piaget's sensorimotor stage in that it operates on a child's perceptions and actions. The representations tier is like Piaget's Concrete Operational stage. Both Piaget and Fischer describe their respective stages as operating on a child's ability to explain their reality. Again, Fischer follows Piaget by having a fourth stage which sees the reasoning skills developed in the previous stage being developed on a more abstract level (Fischer & Silvern, 1985; Murphy 2008).

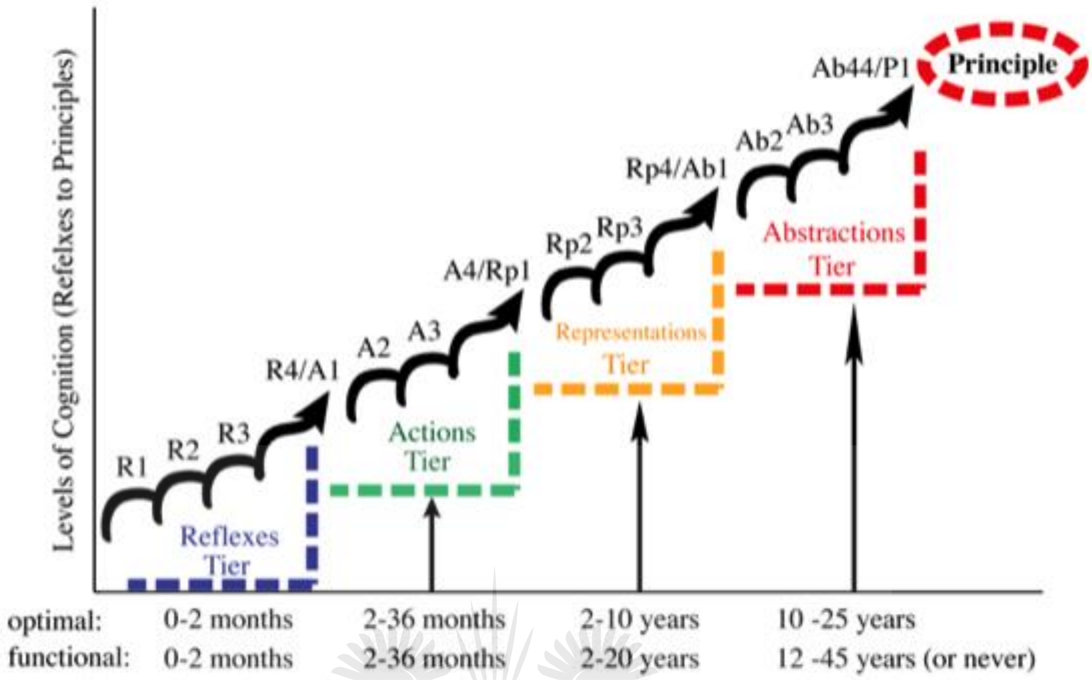

Figure 4. Fischer's four tiers of development, (Murphy, 2008).

Within each tier there are four levels or steps that an individual's needs to achieve to complete the specific tier. The first step, which overlaps with the last step of the previous tier, notes that the child can manipulate a single set of cognitive skills within that tier (Murphy, 2008). Once that has been accomplished, the child is able to move onto the second step which entails the ability to make connections between two sets or being able to identify the relationship between two sets. This process is described as mapping (Murphy, 2008). The third step involves creating a system of mappings by coordinating several mappings. In the fourth and final step, the child can link and co-ordinate systems with other systems, producing a system of systems. This system of system forms the foundation for the first step of the next tier (Murphy, 2008).

Fischer built onto Piaget's stage theory by including the role of the environment in his theory. He agreed with Piaget in the sense that a child would develop through certain, progressive stages; however, Fischer added that the environment within which a child exists will influence how a child progresses through each tier (Murphy, 2008). Reams (2014, p.139) describes Fischer's theory as "a very interactive model, taking biology, structure of the mind, social relationships and environmental influences all into account, to develop a general model of development that can be applied in any domain or context." In terms of this study, where I am looking at how culture bias can influence a child's ability to perform in a test, I have argued that one cannot explain cognitive development in an environmental vacuum. Rather for a cognitive development theory to be relevant, it must consider the effects of the environment on both a macro and a micro level. That is, the effects on the group at large that it wishes to explain, but also on how it can cause individual variances within that group (Murphy, 2008; Reams 2014).

Following the works of the neo-Piagetian theorists such as Pascual-Leone and Fischer, there has been an increase in post-Piagetian theorists, like Halford, Case and Siegler, to name a few (Morra, Gobbo, Marini, & Sheese, 2012). However, to have a richer understanding of cognitive development in children, it is important to discuss, albeit at a cursory level, some of the theories within the post-Piagetian approach. For the purposes of this research I will focus specifically on the works of Susan Carey. As I have argued, intelligence cannot be viewed in isolation as it is affected by various factors. The reason I choose to look at Susan Carey's theory is because, as stated in Chapter One, language is the fault line that divides cultures. If I am going to argue that a language free test is more culturally fair than a language loaded test, it seems only fitting that I should look at a cognitive development theory that places similar importance on language and the development of concepts.

2.4.4. Carey's Origins of Concepts

Another post-Piagetian theorist is cognitive developmental psychologist, Susan Carey. According to Carey, humans develop "rich" conceptual representations of our world (Carey, 2009). These start as a concept, or thought, which is developed into a belief, and further developed into a theory, the most complex of mental representations. When a child, for example, is presented with a situation that compliments their existing theory (referred to as the ancestor concept), the conceptual systems are reinforced as it holds true. However, when a child is confronted with a concept that is not only unexplained by the existing system, but is completely incoherent to it, incommensurability occurs. The ancestor concept may be held by the child for a time, despite contradicting concepts or thoughts, but after a time the child will grasp the new concept (bootstrapping) and develop a descendant concept (Carey, 2009). The descendant concept cannot be explained by the language of its ancestor concept and the ancestor concept ceases to exist. Furthermore, the child analyses the core principles at the foundation of the initial theory, discarding or altering them to match the beliefs of the new conceptual system. And this, according to Carey, is the fundamental of learning. While this occurs with high frequency in children, it is also evident in adults. The process of cognitive development, while predominantly seen in early childhood, is not exclusive to it (Carey, 2009).

2.5. Culture and intelligence

It is evident from the above that one cannot extricate the influence of culture and the environment in intelligence (Carey, 2009). Not only does it influence on a biological level, but also on a cognitive developmental level.

2.5.1. Indigenous knowledge systems and intelligence

When looking at assessments in the South Africa context, specifically in the education sphere, the use of intelligence testing is prevalent. They form an integral part of both the school readiness screenings, as well as special placement screenings. However, before intelligence testing in the South African context can be examined, the link between indigenous knowledge systems and intelligence needs to be examined (Mpofu, 2002).

The whole idea of defining and quantifying intelligence, as well as understanding it in different settings or societies across the world has been controversial for many years. One of the most controversial research projects was that of Flynn (2007). Flynn noted that the IQ score of humans were increasing by an average of 3 points per decade. He argues that the systematic increase in IQ scores is occurring too rapidly to be caused by evolution. While this is not proof against the role of genetics in intelligence, Flynn postulates that it can be taken as evidence of the effect of the environment on intelligence levels. This became known as the Flynn Effect. In addition, Flynn observed that the average IQ scores of blacks were increasing faster than the average score of whites. A possible reason for this was that in the past blacks had access to less favourable education environments, which may have negatively influenced their ability to perform well in an IQ test (Flynn, 2007; Shaffer, 2007). Nevertheless, as blacks are now gaining access to more favourable environments, their ability, as tested in an IQ test, has increased faster than compared to that of white. Flynn has used the above research to argue that IQ tests do not test for true intelligence per say, but rather an individual's access to favourable environments, or adaption to modernity (Flynn, 2007; Shaffer, 2007). However, in some circles, Flynn's work has been misunderstood as people brought in their own biases regarding what he said. This behaviour is common whenever the question of intelligence is involved. Nonetheless, despite the controversy surrounding dialogues pertaining to a specific society's view of intelligence, whether it be the USA, Australia or South Africa, one should still look at whether these dialogues or more specifically intelligence testing is relevant in a modern–day South Africa (Dawes & Biersketer, 2011; Flynn, 2007).

Dawes and Biersketer (2011) argue that, across societies, it is an individual's ability to problem solve which forms the basis of their implicit theory of intelligence and therefore that of their explicit theory. Nonetheless, while the ability to solve problems may be the universal foundation for defining intelligence, the problems that individuals face, differs across societies (Neisser et al., 1996). For example, Western society's view on intelligence has been described by Flynn (2013) as an adaption to modernity, rather than a definition of intelligence in its truest form. In more rural areas, on the other hand, an individual's memory or kinetic knowledge may be higher ranked over their ability to function in a technology driven society (Neisser et al., 1996).

Owusu-Ansah and Mji (2013) state that "African researchers need to persist in developing and using alternative methods of studying our reality and refrain from sticking to the research pathways mapped out by Western methodologies. Knowledge of science, and its methods of investigation, cannot be divorced from a people's history, cultural context and worldview" (pg.1-2). One cannot determine how to test intelligence levels in a specific society if the indigenous knowledge system of that society has not been researched. Serpell (2011) argues that the importance of research to not only meet academic criteria, but to "also resonate with indigenous understanding" (p.37).

Wober (cited in Mpofu, 2002) studied a Ugandan village in 1974 and noted that intelligence was defined in a socially orientated manner. The more an individual added value to the village as a collective, the more intelligent the individual was thought to be. This correlated with Irvine's research on the Shona's definition of intelligence carried out in 1970 (cited in Mpofu, 2002). However more recent research done by both Irvine and Mpofu on both the Shona and Ndebele tribes shows that the African perspective on intelligence is changing to include more Westernised conceptions. That is, while the view on intelligence still considers the collective, it now incorporates more individual characteristics such as level of schooling and economic wealth to name a few. Mpofu argues that as more remote African communities are encountering modern economies, so their conceptions of intelligence are changing. Their adaption to modernity, as Flynn (2007) describes, is becoming an integral part of the African view of intelligence. Nonetheless, perhaps the African view of intelligence is not becoming more focused on the individual, but rather they have come to see that the more "intelligent" and successful an individual is in modern terms, the more that individual has adapted to modernity so to speak, the more able that individual will be to assist his community.

Citing the research of Mpofu and Wober, Serpell (2011) concludes that "a distinction emerges between the notion of cognitive alacrity on the one hand and that of social responsibility on the other, with a highly valued personality trait defined as a combination of the two" (p.40). The research referred to by Mpofu (2002) is of societies primarily located in Zimbabwe and that means that the value of intelligence of a working age individual is predicated on whether they are in or out of the community concerned. The Zimbabwean in the diaspora (especially South Africa) is adding value (goods, money etc.) to the origin community in ways that are probably impossible for the resident worker. This is not just an African phenomenon, but it is the primary drive of diasporas across the globe.

While there is a growing demand for more research to be done in the literature on indigenous knowledge systems and intelligence, there is sadly not a huge amount of research on the topic. While psychology in general is spreading across the African continent it continues to be predominantly Eurocentric in nature, neglecting the epistemologies and values of the very people it aims to assist (Nsamenang, 2007). However, it is emerging that the African perspective of intelligence is more heterogeneous than its Western counterpart. Ngara (2007) agrees that not only is 'African Intelligence' an incorporation of the Western definition and the African community orientated definition, stating that one cannot deny the positive effect that the exposure to Western ways of knowing has had on the African. However, he expands on this by adding that the African way of knowing is spiritually centred, whereas the Western way of knowing is primarily driven by science. As a result, for many centuries the colonists dismissed the African way of knowing as being inferior. Nsamenang (2007) states that the Eurocentric knowledge system has dismissed the "tacit wisdom embedded in Africa's oral sources of knowledge like proverbs, folklore and practises" (p. 4). Nevertheless, Ngara (2007) cites various examples of where the spirit-centred African way of knowing comes to similar conclusions as that of the Western world.

The fact is that despite the change that has occurred around it, the indigenous knowledge systems have remained the same to a certain extent. Changing on the periphery, as Western ideologies are included, the value systems, beliefs and knowledge have remained the same. This core foundation is now being researched and valued increasingly by the current generation (Barnhardt, 2005; Ngara, 2007).

And it is because of this that the indigenous knowledge system and its views on intelligence need to influence and guide the dialogue of intelligence testing in South Africa.

2.5.2. Language: the product and creator of culture

Understanding the value of language is essential to understanding why humans are intelligent. As a species, we are the only animals who create a “rich conceptual understanding” (Carey, 2009, p.3) of our world, but linguistic communication has enabled us to transfer this knowledge to future generations in a far more effective manner. Throughout our daily lives we rely heavily on the problem solving of previous generations. Life as we understand it has been shaped and enabled by the intelligence of those who came before us (Hall, 2013). Language offers us the ability to transfer not just innovations, but the rationale and idea behind these innovations and technologies onto future generations (Geary, 2015). This has meant that each generation does not need to start from the beginning but can rather build on and add to the knowledge that preceded it (Geary, 2015). Language is considered a “sociocultural resource” (Hall, 2013, p.7), not only grounded in the culture of its origin, but also a means with which both the culture and the individual within that culture can express themselves, sharing knowledge and growth. Hall (2013) states “language can only reflect cultural understandings; it cannot affect them (p. 16). While I agree that the primary role of language is to reflect, language is fundamental to learning, and I would argue that learning affects our cultural understandings. Therefore, language affects our cultural understandings, though perhaps in a more indirect manner. As Gopnik (2009) notes, it is our knowledge of the world that allows us to change it. (Geary, 2015; Hall, 2013; Lupyan, 2015).

2.6. Psychology and intelligence

The study of intelligence and subsequently, intelligence testing has its beginnings long before psychology was identified as a separate discipline (Cattell, 1987). Intelligence testing and psychological assessments in general, involves the collection of information which can be used to make a judgement or predication on the ability of either an individual or a group (Shaffer, 2007).

2.6.1. The history of intelligence and intelligence testing

The ability to make accurate predications with regards to success in academic achievement is considered of great importance (Beech & Singleton, 1997; Foxcroft, 1997). Saklofske, Van de Vijver, Oakland, Mpofu, (2015, p. 341) state "both the construct of intelligence and its measurement predate the establishment of psychology as a scientific discipline and most likely have a history as long as human civilisation". According to Kaufman (2009) and Saklofske, et al., (2015), in 2200 BC a Chinese Emperor administered proficiency tests to his officials once every three years. 1000 years later political candidates in the Chang dynasty were required to complete an ability test before taking office. Even more interesting was that in June 1763, an eight-year-old Mozart was presented at King George III's court, where his cognitive ability was tested by philosopher Daines Barrington. It has been proposed that Mozart was the subject of the first intelligence test report as we know them today. The beginnings of psychometrics has it root in the desire to quantify cognitive ability in children. This speaks to the value of this ability to practitioners and why it is vital that practitioners have access to valid and reliable assessments.

By the mid 1800's, the study of intelligence had evolved to such an extent that a taxonomy had been developed. This allowed individuals who were lacking in 'normal' levels of intelligence to be identified and receive, what was considered to be appropriate, intervention or care (Cattell, 1987; Kaufman, 2009).

Intelligence testing, as we understand it, has its origins in more recent history. In 1899 Binet was approached by the Free Society for the Psychological Study of the Child to develop a test that would give an indication of which children would benefit from remedial instruction (Shaffer, 2007). Not only did the *Binet-Simon test* allow psychologists the opportunity to assign children a standardised mental age, but they were able to make concise measurement statements (Cattell, 1987; Shaffer, 2007). However, Binet was criticised for not paying adequate attention to defining intelligence. It was argued that Binet did not have a concise idea of his own opinion on the subject (Cattell, 1987). This criticism stemmed from the fact that the *Binet-Simon test* adopted a multifocal perspective of the structure of intelligence. However, in some of his writings, Binet suggests a univocal perspective (Cattell, 1987; Kaufman, 2009; Neisser et al., 1996; Saklofske et al., 2015; Shaffer, 2007).

Towards the late 1800's and early 1900's the influx of immigrants to the United States of America produced a concern that the general intelligence of the population would decrease. This spurred an increase in the development and use of intelligence testing in the United States of America. In 1916, Lewis Terman Publishing House published a revised version of the *Binet-Simon test*, namely the *Stanford-Binet test*. The start of the First World War further propelled the industry to develop tools that would enable them to place soldiers in areas that would best suit their skills and abilities (Kaufman, 2009; Shaffer, 2007). From henceforth, the use of intelligence testing in psychology, the workplace, as well as the classroom became standard procedure. As IQ testing became more popular there was greater demand for more comprehensive theories (Benson, 2003a; Kaufman, 2009; Saklofske et al., 2015).

2.6.2. Theories on Intelligence

There are various intelligence tests available to practitioners and these tests are underpinned by intelligence theories. In order to fully comprehend intelligence tests, I will look at the various theories that underpin them. The increase in the development of the IQ test, led to an increase in the need to develop more comprehensive theories on intelligence, which, in turn, would influence the models used to develop measurement tools themselves (Saklofske et al., 2015). The most notable of these being Spearman, Guildford Carroll and Cattell's theories of intelligence.

2.6.2.1. Spearman's g Factor.

Following the popularity of the various versions of the *Binet test*, Spearman (1927) examined whether intelligence is a single power (univocal) or rather a bundle of unrelated abilities (multifocal). To answer this question Spearman developed factor analysis (Shaffer, 2007). Factor analysis gives an indication on whether specific abilities are related, unrelated or inversely related. He proposed the *g* factor theory. According to Spearman an individual's intellectual abilities rise from a foundation of general intelligence (*g)* (Cattell, 1987; Shaffer, 2007). As a result, if one could access an individual's general intelligence levels, one could predict the level of their specific abilities. Spearman set the stage for what continues to this day to be a back and forth debate among theorists. Each of them attempting to answer Spearman's question, and in so doing define intelligence (Cattell, 1987; Shaffer, 2007).

2.6.2.2. Guildford's structure of intellect.

Guildford was an American psychologist who proposed a multifocal theory of intelligence. Guildford rejected Spearman's univocal *g factor theory* and he proposed his own Structure of Intellect theory (SOI) (Guildford, 1967). SOI theory suggests that intelligence is multifocal, that is comprised of 180 factors or abilities. These abilities where arranged along a three-dimensional scheme, namely, content, operations and products (Guildford, 1967).

According to Guildford, an individual will apply their intellectual ability into four broad areas. These four areas are figural content, symbolic content, semantic content and lastly, behavioural content (Guildford, 1967). Next, there is the operations dimension. This includes the intellectual processes such as evaluation, convergent production, divergent production, memory and cognition (Guildford, 1967).

The last dimension in Guildford's SOI is the products dimension. The operations, discussed above, are applied to the various contents and result in products These products consists of either units, classes, relations, systems, transformations and implications (Guildford, 1967). Initially Guildford had 120 types of intelligence. However, he continued to revise his theory until his death. There are now 180 types of intelligence, which are made up of a specific operation being applied to a specific content, resulting in a specific product. Guildford's multivocal theory is a far cry from the univocal theories of Binet and Spearmen (Sternberg, 1982 & 2015).

2.6.2.3. Cattell's Theory on Intelligence.

The test that is the focus of this study is the Culture Free Test (CFT), which was developed by Raymond Cattell and is underpinned by his theory on intelligence. Cattell postulated that general intelligence is comprised of crystallised intelligence (G*c*) and fluid intelligence (G*f*). G*c* refers to one's ability to gain knowledge and apply that knowledge at a later stage. It involves the acquisitions of declarative knowledge and as Klauer & Willmes (2002) state it is "conceived of as a combined product of fluid intelligence and education" (p. 2). *Gc* refers to an individual's understanding of language, information and concepts of culture. G*c* is acquired through learning and experiences, it reflects learning experiences. It does not equate to memory, although long-term memory plays an integral part in *Gc*. (Cattell, 1987; Craik & Bialystok, 2006; Klauer & Willmes, 2002; Kvist, & Gustafsson, 2008).

Fluid intelligence (G*f*) refers to a capacity to solve novel or abstract problems. It involves concept formation, classification and includes inductive and deductive reasoning. Unlike crystallised intelligence, fluid intelligence is not a *learned* ability, but rather is determined by genetic and biological factors. According to research, the two types of intelligence develop along different trajectories across a lifespan. G*f* peaks in the mid-twenties, and then starts to decline, whereas G*c* continues to increase until the early seventies (see Figure 2.4), (Cattell, 1987; Craik & Bialystok, 2006; Klauer & Willmes, 2002; Kvist, & Gustafsson, 2008).

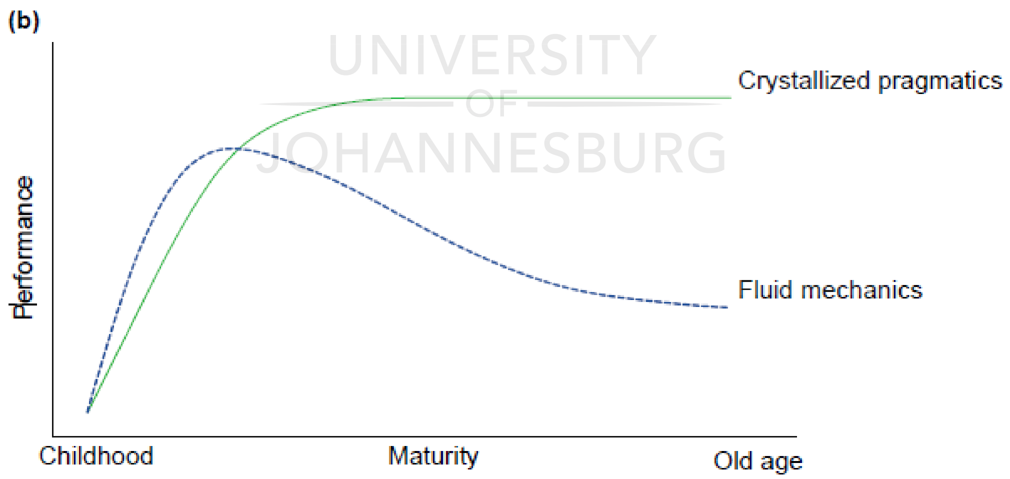

Figure 5. Development of Gc and Gf across the lifespan (Craik & Bialystok, 2006).

Bergman Nutley et al., (2011) note that "G*f* predicts performance on a wide range of cognitive activities, and low G*f* in children is a predictor of academic difficulties" (p.591). According to Geary (2015), human competence is underpinned by our ability to take advantage of our evolved brain and cognitive systems and to use these resources to create “evolutionarily novel abilities” (Geary, 2015, p.105). These abilities form the foundation of culture as we experience it and the cross-generational accumulation of these progressions has led to the development of non-evolved abilities, including, but not limited to, writing systems, literature, science and art. Geary (2015, p.105) asserts that fluid intelligence “is critical to the creation and learning of these non-evolved abilities.”

2.6.2.3.1. Inductive Reasoning and Intelligence.

The CFT1-R assesses inductive reasoning. The link between inductive reasoning and intelligence has been considered by four major intelligence theorists, namely Spearman, Thurstone, Cattell and Gustafsson (Klauer & Willmes, 2004). While their view on the extent of the influence of inductive reasoning on intelligence varies, they all concur that there is an undeniable link between the two. Cattell viewed both inductive and deductive reasoning as part of G*f*. It is useful, from the outset, to distinguish between the two. Shye (1998) describes inductive reasoning as establishing similarities or differences, whereas deductive as applying rules. Thus, inductive reasoning starts from the viewpoint of observations, from which a hypothesis can be made. Whereas deductive reasoning starts with a hypothesis, from which generalisations can be made. The CFT1-R assesses inductive reasoning.

2.6.2.4. John Carroll’s three stratum theory of intelligence

Over recent years, hierarchical models of intelligence have been favoured by psychometricians, the most popular being the Carrol-Horn-Cattell three stratum theory of intelligence (CHC). This theory hypothesises that intelligence is pyramid shaped, with *g* being the tip, followed by eight broad abilities on the second level (Shaffer, 2007).

Although theorists differ in the details of how they define intelligence, there is a commonality across all the theories. It is commonly accepted that intelligence signifies an individual's ability to foresee and predict discrepancies and innovations within their environment. Based on this information, he/she can "devise strategies to cope" (Geary, 2015, p.105). However, there is not agreement in how to define intelligence. This points to the problem with quantifying intelligence as outlined by Bond and Fox (2013): how can we confidently assess something that cannot be seen or conclusively defined? Having said this, why do we make use of cognitive assessments at all? Specifically, in the context of this study, there appears to be little current research on the western theories of intelligence looking at culture as a contributing factor. From an African perspective, research on indigenous knowledge systems and intelligence (discussed in 2.5.1. Indigenous knowledge systems and intelligence) seems to place more value on the role of intelligence.

2.7. The quandary around intelligence testing

There can be no denying that intelligence has distinguished us as a species. Nor can we deny that language has played an essential role in making that distinction. It has long since been a desire of humans to quantify an individual's intelligence levels. This desire has, in turn, led to the development of numerous IQ tests (Cattell, 1987; Lupyan, 2015). The use of IQ tests is prevalent in the field of education and is a common component in psycho-educational and school readiness batteries. The dilemma faced in South Africa is that psychometrics has a history of being used to oppress rather than liberate (HPCSA, 1997). Despite this history, we now find ourselves in a situation where we are bound by South African law to administer fair tests (Van de Vijver & Rothmann, 2004). Yet, particularly when it comes to psycho-educational assessments, the tools we have are not appropriately adapted or normed for use in South Africa. When administering a psycho-educational assessment most practitioners include a cognitive assessment (Foxcroft & Roodt, 2006). This despite the fact that the majority of practitioners in South Africa agree that the tests that are available are far from ideal. So why do we need psycho-educational assessment, especially if the cognitive assessment is not adapted for use in our population?

2.7.1. Psycho-educational assessments

Dr Linan-Thompson (2014), associate professor at the University of Oregon in the Department of Special Education and Clinical Sciences, states that one of the main objectives of any education system is to ensure that children succeed in their efforts to obtain an education (effectively to learn what previous generations have mastered). In so doing, they prepare themselves for the adult world (Linan-Thompson, 2014).

One of the ways this can be facilitated is by allowing the child the best opportunity to transition from informal schooling (Nursery and Pre-Primary School) to formal schooling (Primary School). Issues during this transition period can perpetuate into adulthood: “the level of success during transition to school or transfer between phases of education, both socially and academically, can be a critical factor in determining children’s future progress and development” (Vrinioti, Einarsdottir, & Broström, 2010, p. 5).

The use of tools such as school readiness assessments for example, allow educators the opportunity to identify children who are cognitively, behaviourally and emotionally ready for formal schooling (Foxcroft, 1996). Furthermore, it assists in identifying areas where a child might be having difficulty and requires additional support. Not only do psycho-educational assessments have therapeutic and diagnostic uses, but they allow educators to make informed decisions when discussing curriculum, progress evaluation as well as career development (Foxcroft, 1996; Foxcroft & Roodt, 2006). It is important to note here that the aim of psycho-educational assessments in general and intelligence testing more specifically, is not to label a child, but rather to identify children who need additional support. Research studies conducted in Brazil, Guatemala, Jamaica, Philippines as well as South Africa suggest that there is a strong correlation between early cognitive abilities and academic performance in primary school as well as into high school. Linan-Thompson (2014) states that the above-mentioned findings “support the use of screening measures to identify children who may need additional interventions or support to succeed in primary school” (p. 3). In addition, theoretically, an IQ test should enable the tester to identify individuals who are intellectually talented from all walks of life, thus opening education systems where admittance in the past has been restricted to race, gender, ethnicity, socio–economic background, or even one’s surname. However, this is an ideal that has yet to be fully realised and one of the vital steps that need to be taken to realising this is the development of new, up to date assessments (Benson, 2003a & 2003b). The fact is that even with our flawed assessments, practitioners are still able to derive an indication of where variations—whether they be positive or negative—lie and can make necessary recommendations based on these. The value of a cognitive assessment is clear, but South African psychologists, psychometrists and test developers and distributors are doing a disservice to our learners by not insisting that up to date, relevant tests are available.

2.8. Testing in a South African Context

Testing in a multicultural context, such as South Africa, is not a simple task, mostly due to the lack of culturally appropriate tests (Foxcroft & Roodt, 2006). There is a demand from practitioners for culture fair assessments tools that are not culturally biased (Foxcroft & Roodt, 2006). The lack of these unbiased tests can partly be attributed to the origins of South African psychometrics under the apartheid regime (Health Professionals Council of South Africa [HPCSA], 1997).

2.8.1. Cognitive Assessments in South Africa

The two biggest types of assessment in the educational psychology arena is psycho-educational assessment (94.5%) and school readiness assessment (80.9%) and a fundamental component of these assessment batteries is a cognitive assessment. This is evident when one looks at the top ten tests as used by educational psychologists. The top two in the list are cognitive tests, namely the Senior South African Individual Scale (Revised) and the Junior South African Individual Scales (JSAIS) (Foxcroft et al., 2004).

Based on my own personal experience as a psychometrist, as well as my interactions with various psychologists, particularly educational psychologists, I believe there are three cognitive assessments that are primarily used in psycho-educational assessments in South Africa. Namely the Junior South African Individual Scales (JSAIS), Raven's Colour Progressive Matrices and the Wechsler Pre-School and Primary Scale of Intelligence-Revised (VVPPSI-R) or the Wechsler Intelligence Scale for Children (WISC-III). All the above tests are listed on the HPCSA’s List of classified and certified psychological tests (2017) as “tests that have been classified, but not reviewed”.

As outlined in the literature, the advancement of intelligence/cognitive tests as well as the theoretical developments which underpin the tests have been nothing short of complex and at times contradictory. But a common theme, especially when one examines the indigenous knowledge system, is the awareness for the need to develop and adapt tests for use within the South African and African context. A test is as much influenced by the culture within which it was developed as the theory used as its foundation. It goes without saying then, that one cannot simply take a test from a Euro-American background, for example, and administer it unchanged or insufficiently changed and expect its validity to hold true. Unfortunately, in South Africa, very little has been done to adapt foreign tests to the local demographic (Foxcroft et. al., 2004; Van der Viver & Rothman 2004; Nsamenang 2007).

The JSAIS is one of the most popular tools for the assessment of cognitive ability in South African children (Foxcroft et al., 2004). It is an intelligence scale that gives an indication of future scholastic ability as well as possible areas of delay in intellectual functioning. The JSAIS is primarily used as a diagnostic tool for children with learning problems or disabilities (Madge, Van den Berg, Robinson, 1991). The JSAIS was first published in 1979, by the HSRC and is comprised of 22 subtests. It has been normed for children between the ages of three years to seven years and 11 months. However, these norms preclude most of the South African population. The tests had only been normed for use on White, Coloured or Indian children and was only available in English and Afrikaans (Huysamen, 1996). It was subsequently translated into isiZulu and Sesotho; however, it should be noted that a simple translation of a test will not necessarily remove cultural bias as the test items, wording or imagery may still be culturally loaded (Mduli, 2011; Oakland, 2005). Furthermore, practitioners believe some items in the various subtests are outdated and no longer relevant (Foxcroft et al., 2004). Both Naicker (2013) and Mawila (2012) agree that despite the translation of the JSAIS, both the isiZulu and Sesotho still have cultural bias in the test items. Despite the above concerns, the JSAIS is still widely used across South Africa, which points to one of the serious problems with intelligence testing in South Africa, there are very few tests available to practitioners and those that are available are seen as culturally biased (Foxcroft et al., 2004).

The Wechsler Intelligence Scale for Children (WISC) is another popular assessment tool in South Africa (Laher & Cockcroft, 2013) and is primarily used for diagnosing learning disabilities and attentional disorders. Currently it is available for use in English to children between the ages of six years to 16 years old. It contains 10 subtests, which provide a Full-Scale IQ score.

The WISC has not been adapted for use in South Africa, however, Shuttleworth-Edwards, Van der Merwe, Van Tonder & Radloff, (2013) conducted extensive research in order to develop norms for grade 7, South African learners. While these norms have assisted practitioners across South Africa, they only apply to grade 7 learners, meaning children who are not that age may still be mistakenly classified by the test results (Laher & Cockcroft, 2013).

The Ravens Colour Progressive Matrices Test (Ravens CPM) is a non-verbal intelligence test for use on children between the ages of five to 11 years of age. It is widely used in culture fair tests and is based on Spearman's theory of intelligence discussed earlier (Carlson & Jensen, 1980; Raven, Raven, & Court, 1998). As with the WISC-IV, the Raven's CPM is extensively use in South Africa, but there is still a need for norms for the non-adult South African population (Bass, 2000).

2.8.2. The oppressive history of psychometric testing in South Africa

The history of psychological assessments in South Africa has been "tainted by the legacy of segregation which influenced certain stereotypical attitudes and culturally insensitive and inappropriate interventions" (HPCSA, 1997, p.1). The purpose of the import of psychology—and therefore, by default, intelligence assessments—into Africa, was, as with other colonial imports, to uphold the colonial society, while discriminating against and supressing the indigenous people (Nsamenang, 2007). Westermann (cited in Nsamenang, 2007) poignantly comments that Eurocentric psychology in Africa had “largely lost sight of the soil out of which the existing [African] society has grown and the human values it has produced” (p.3).

2.8.3. South African legislation guiding the use of psychometrics in South Africa.

The 1998 Employment Equity Act, however, addressed the issue of fair, unbiased testing in South Africa’s multicultural context. The Act strictly prohibits the use of psychological assessments unless the tests “can be applied fairly to all employees and are not biased against any employee or group” (Employment Equity Act 55 of 1998, Section 8, Government Gazette, 1998). South African legislation and ethical codes are well developed and ahead of some countries within the Euro-American psychology fraternity (Wassenaar, 1998). Van de Vijver and Rothmann (2004) acknowledge that this law “is ahead of daily practice” (p.1) in South Africa. It is, however, the ethical (and legal) obligation of the assessment profession to bring “current practise in line with legal demands” (Van de Vijver & Rothmann 2004, p.1).

2.8.4. Problematising psychometrics in South Africa

The HPCSA’s Policy on the Classification of psychometric measuring devices, instruments, methods and techniques, warns that it would be “unwise” (p.1) for the assessment profession not to pursue the adaption of existing tests and the development of new, culturally fair tests. Even so, very few culturally relevant tests have been developed in South Africa (Foxcroft, Paterson, Le Roux, & Herbest, 2004). This is despite practitioners becoming more cognisant of the importance of using sound assessments which maintain their validity across cultural groups (Paterson & Uys, 2005). Foxcroft et al., (2004) noted that 65.8% of practitioners “indicated that they feel the tests that they use are only sometimes appropriate to use cross-culturally" (p. 20). Most practitioners surveyed (58%) indicated that more culturally fair tests are needed to be made available in South Africa. Based on these findings, there appears to be a justifiable demand from practitioners for tests that are deemed culturally fair in South Africa.

Most of the psychological tests, including, but not limited to intelligence tests, that are available in South Africa have been developed by the Euro-American psychology fraternity for use in their societies. It goes without saying that assessments are influenced by the socio-economic culture in which they have been developed. In addition, the purpose for which the assessment was developed will also affect the potential bias of the assessment. These tests have been influenced by Euro-American theories of psychology and have undergone validity and reliability testing which make them relevant and credible for use within those societies. It is beyond question that these tests are not suitable for use on an African child if they have not been adapted for that very purpose. An investigation into potential bias needs to be completed, and from there the assessment needs to be adapted and re-normed (Foxcroft & Roodt, 2006; Nsamenang, 2007).

2.8.5. Addressing the issue of bias in intelligence testing.

Language is thought to be the fault line along which culture varies and as such has a considerable impact on test performance in a multilingual South Africa. In a testing situation, language is therefore a source of potential bias. The development or revision of non-verbal tests has been postulated as a possible solution (Nell, 1994; Foxcroft & Aston, 2006; Owen, 1991). Simply translating a test into a mother tongue is in no way sufficient to ensure the validity of the test in a multicultural context (Oakland, 2005). When developing new instruments or revising existing instruments, it is crucial to make allowances for cultural differences and eliminate cultural bias by adjusting the test items thereby ensuring they match the culture in which the test will be administered. Only then can we ensure the validity of the test when used in multicultural groups (Van de Vijver & Rothmann, 2004; Foxcroft et al., 2004 & 2011; Benson, 2003). The most obvious way of minimising cultural bias is by removing language, creating a non-verbal test. In its simplest form, a non-verbal test does not require the test taker to have an understanding of a language to understand a question. It has been argued that verbal intelligence test does not only assess an individual's IQ levels, but more importantly the individual's language comprehension. By removing language from the test items, items will have less of a potential bias and will be more valid, in the sense that they are more likely to assess that which they set out to assess. (Foxcroft 2004 & 2011; Benson, 2003). Although language has been removed from the test items, it is still needed in the instruction. It is impossible to remove language all together from a test and there is research stating that symbols and icons can be more culturally loaded than words. (Benson, 2003).

While the use of a non-verbal assessment does negate the issue of language bias to a certain extent, an additional source of bias could be the timed assessments. Not only does the presence of a time limit affect test anxiety, but a child's test wiseness and exposure to a test environment can affect their ability to complete the test with speed and accuracy Oakland & Weilert (1971) define test wiseness as the ability to draw previous test taking experiences and use these test taking abilities to receive a score that accurately reflects their abilities. (Portolese, Krause, & Bonner, 2014; Hill & Wigfield, 1984; Samuels, 2015; Flippo, Appatova, & Wark, 2018).

2.8.6. Time and test taking

In this section we will be looking at time and what effect it may have on a test taking ability. Firstly, I will look at how a time constraint influences test anxiety, as well as the test taker's ability to perform at their optimal level. Secondly, we will look at the child's perceptual ability and exposure to testing environments and how this may be influenced by time constraints. The traditional intelligence theories discussed in section 2.6 of this chapter do not consider time allowance as a factor of one's ability to accurately represent intelligence. However, based on the below readings, there would appear to be an increase in research acknowledging the role of time in intelligence testing.

A test taker's experience of test anxiety can negatively affect their ability to perform optimally on a test (Portolese et al., 2014; Hill & Wigfield, 1984). According to Dusek (1980) test anxiety is a negative feeling or emotion that is experienced on a physiological and behavioural level. These feelings are elicited during a formal test or evaluation situation. Hill and Wigfield (1984) state that test anxiety is developed in young children as adult figures have high expectation for the child's performance. When the child performs poorly in a testing situation, the adult's negative reaction further enforces the child's test anxiety. Raufelder, Regner and Wood (2018) note that there is a correlation between emotionality (test anxiety) and learned helplessness in testing situations.

One of the factors that influence test anxiety and testing cognitive ability is the pressure of a timed test. The potential impact of extended time versus usual time allocation in a standardised test has been debated in various articles (Huesman & Frisbie, 2000; Mandinach et al., 2008; Zuriff, 2000). When researching the effects of time pressure Caviola, Carey, Mammarella, & Szucs, (2017), found that "the presence of a time constraint in any math or problem-solving situation can effect performance" (p 2). The addition of a time constraint can negatively influence the decision-making strategy chosen to complete the task. This is evident in both adults and children, who even at a young age will "adapt their strategy use to the external demands in terms of coping" (p.3, Caviola, et al., 2017). Various authors note that by removing the time limit all together, test takers are able to focus more effectively on important information, choosing the most suitable strategy selection. (Caviola, et al., 2017).

Originally, admission standardised test were not timed, however for ease of administration a time constraint was implemented (Evans, 1980). Whilst Caviola, et al., (2017) notes that an unlimited amount of time allows the test taker maximum opportunity to perform to the best of their ability, Portolese et al., (2016) and Mandinach et al., (2008) argue for extended time, rather than the removal of time constraints all together. When researching high school learners, Mandinach et al., (2008) found that while some learners performed better with more time, other learners performed worse when given too much time. His results found that:

- Children with low ability do not benefit from increased time as they do not have the skills. An extended time limit cannot improve their cognitive ability and *de facto* their performance.
- Both children with medium to high ability benefited from extended time, regardless of whether they had disabilities or not.
- Children with disabilities and low abilities where impeded by the extra time (Mandinach et al., 2008).

Literature shows that while the extension of a time limit can be beneficial, the removal of it all together can have detrimental effects on the test results (Mandinach et al., 2008; Portolese et al., 2017). A further implication of time on results could be the child's level of perceptual ability and test wiseness (Cohen, 2006). A limited time constraint may not allow sufficient time for the child to familiarise themselves with what may be a new and foreign environment that is the test taking situation.

In a previous administration of the CFT1-R on South African grade 1 learners it was informally reported by some of the test administrators that the perceptual skills of the learners inhibited their ability to do the assessment. Furthermore, it was observed that the grade 1 learners had little experience with a testing situation. One documented way of familiarising learners with a testing situation is the use of example items. Foxcroft (2011) noted that example items ensure that the learners understand what is required of them. This should assist in minimising testee anxiety in learners who are not experienced test takers.

Grade r in the South African school system is intended to lay the foundation for academic and perceptual ability (Samuels et al., 2015). Samuels, et al., (2015) found that by 2011, 80% of five-year olds in South Africa were attending grade r. Research shows that there is a positive correlation between grade r and future scholastic ability. However, the impact of grade r on scholastic ability in poorer areas was almost negligible due to instruction quality (Moletsane, 1996; Samuels et al., 2015). Moletsane (1996), found that the pre-school environment in poorer areas of South Africa was characterised by insufficient teaching materials (including little or no reading and writing instruments) and untrained teachers, some of whom cannot read or write themselves. The Department of Basic Education (2015) issued a statement that in 2013, 16 520 (78%) of grade r teachers in South Africa, did not have a diploma to teach. A child's socio-economic status directly affects the standard of education he or she will have access to (Van der Berg, 2007). This in turn affects their scholastic ability, such as the ability to work with speed and accuracy, as well as their writing ability due to lack of practise (Samuels et al., 2015; Truter, 2015; Van der Berg, 2007). This ties into test wiseness, as the lack of exposure to pen and paper type evaluations will affect their ability to perform with speed and accuracy (Flippo et al., 2018). It was this lack of exposure that lead me to assess grade 2 children, rather than grade 1, as these children would have had a minimum of a year's exposure to testing situations.

Based on the above research and for the purposes of this study, I will research what the optimum time extension is needed for South African children using the CFT1-R.

2.9. Conclusion

Worldwide, intelligence testing is used in scholastic applications in terms of school placement etc. However, in South Africa, children are currently being tested with tests that are outdated and/or are culturally loaded and as a result may discriminate against those being tested (Foxcroft et al., 2004). This literature review has provided a thorough exploration of the recent developments in the field of intelligence and cognitive development, both on a global and local stage. The reader was introduced to some of the salient theories of intelligence, specifically the theory of Cattell which forms the theoretical underpinning for the CFT1-R test. It is clear from the above review that one cannot define intelligence in isolation, but rather a multidisciplinary view needs to be taken. There is an urgent need for more research and development to take place, specifically in test development, as many of the tests that we currently have available have a strong cultural bias.

The above discussion of the identified problem has led me to the formulation of my research question.

How can the cultural fairness of an intelligence test be advanced/adjusted by modifying the test times and providing pictorial examples?

3. Chapter 3: Research Design and Methods

3.1. Problem Statement

Psychological tests are a vital part of assessment tools, especially for Educational Psychologists, who use them more frequently than their clinical and counselling counterparts (Foxcroft et al., 2004). Foxcroft et al., (2004) shows in applied practice with children, across the registration categories of psychologists, 22.2% sometimes use psychological or educational assessments, while 39.3 % frequently make use of psychological or educational assessments. When looking specifically at the category of Educational Psychologists, 94.1% use tests with children in applied practice. The percentage of clinical and counselling psychologists who use tests with children in applied practice is also high, at 60.5% and 69.5% respectively.

The two biggest types of assessment in the educational psychology arena are psycho-educational assessment (94.5%) and school readiness assessment (80.9%). A fundamental component of these assessment batteries is a cognitive assessment. This is evident when one looks at the top ten tests used by educational psychologists. The top two in the list are cognitive tests, namely the Senior South African Individual Scale – Revised and the Junior South African Individual Scales (JSAIS) (Foxcroft et al., 2004; Madge, van den Berg, Robinson, 1985).

It is clear from the above that psychological and educational tests are commonplace in applied practice with children. However, Foxcroft et al., (2004) and Van der Vijver & Rothman (2004) agree that one of the major issues faced by the assessment fraternity in South Africa is that of culture fairness in testing.

There are only a handful of intelligence assessment tools that have been developed in South Africa. And most of these tests were developed under a regime that had a political agenda to segregate the population. Cohen and Swerdlik (2002) note that “some potential problems related to test fairness are more political than psychometric in nature” (p. 20). In addition to the outdated South African intelligence tests (discussed in paragraph 2.8.1, Cognitive Assessments in South Africa), South African practitioners have access to tests that were developed in a Euro-American context and were normed to suit their respective demographics. To simply administer these tests in a South African context—with little to no adaption or re-norming—calls the validity of these tests results into question, as the tests are likely to have bias in terms of language and culture (Foxcroft et al., 2004).

As mentioned in chapter 2 the HPCSA has warned against the use of tests that have not been adapted for use in South Africa. However, they are unable to ban the use of these tests as this would leave Educational Psychologists with very few alternatives (The Professional Board for Psychology, 2009).

This chapter will introduce the Culture Fair Test and outline the research design and methods that have been employed to direct the empirical work of this study.

3.2. The Culture Fair Test Revision 5

3.2.1. Description of the measurement

The Culture Fair Intelligence Test (CFT1-R), was developed by R. H. Weiß & J. Osterland. It is a revised version of the CFT, which in turn was adapted from Cattell’s Culture Fair Test. The assessment is therefore based on Cattell’s conceptualisations of intelligence and assesses what Cattell defined as fluid intelligence (discussed in section, 2.6.2.3 Cattell’s Theory on Intelligence).

The CFT1-R is a speed test and is comprised of six subtests, which will be discussed in detail below. The first three subtests assess perceptual skills, attention, and visual motor processing speed. The last three subtests assess the child's inductive reasoning skills by making use of tasks that require the child to identify the rules of relationship between elements and then use these rules to complete a structure (Weiβ & Osterland, 2013). It is currently available in German and Polish. Due to the fact that the test is non-verbal and that it does not require pre-existing knowledge, the CFT1-R is considered to be culture fair in that the result can be considered as "independent from culturally–determined experiences" (http://www.en.practest.com.pl/CFT1-R -cattell-culture-fair-intelligence-test-version-1).

3.2.2. Subtests of the CFT1-R

Subtest One - Substitutions: The first subtest of the CFT1-R is the Substitutions test. In this subtest the child is presented with a page of symbols, namely a pencil, clock, house, pair of scissors, teacup and a knife, (Figure 6).

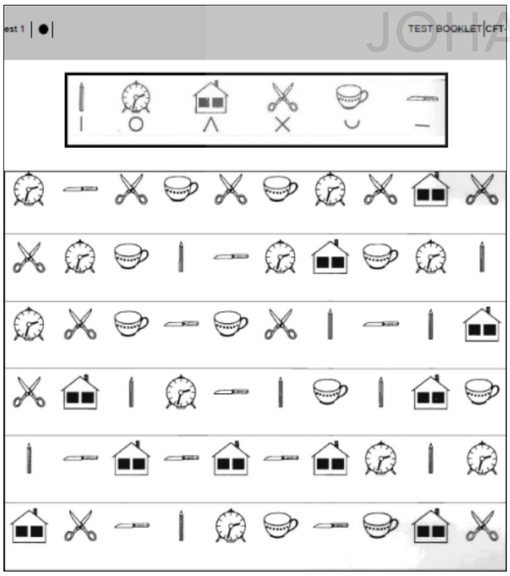

Figure 6.CFT1-R Subtest One, (Weiβ & Osterland, 2013).

Using the key at the top of the page, the child copies or codes each symbol into its corresponding simple geometric shape. This subtest not only gives an indication of visual-motor processing speed, but it also assesses a child's short-term memory skills (Sweet, 2013; Weiβ, & Osterland, 2013).

Subtest One is similar to the Animal Coding Subtest (Figure 7) in the Wechsler Preschool and Primary Scale of Intelligence IV (WPPSI), which also makes use of a key that allows the child to "mark shapes that correspond to pictured animals" (Hanson, 2012, p. 12).

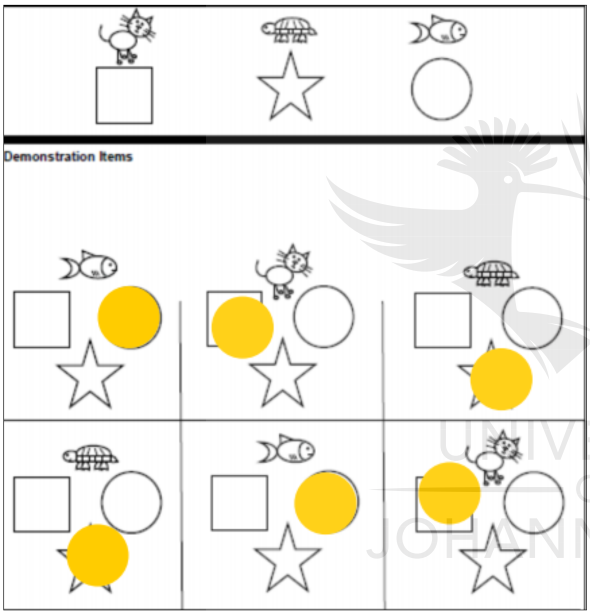

Figure 7.Examples from Animal Coding Subtest, WWPSI (Weschler, 2003a).

The Animal Coding Subtest was a new addition to the WPPSI-IV. In earlier editions, a coding subtest similar in design to the CFT1-R was used (Weschler, 2003a).

Subtest Two - Mazes: In the mazes subtest the child is presented with a five-labyrinth series, containing 15 mazes in total. As with Subtest One, visuo-motor skills are assessed as well as the child's processing speed. The use of mazes as a tool to assess intelligence is not new. In 1914, an Australian schoolteacher developed the Porteus Maze test. There have been numerous adaptions and revisions since then and the Porteus maze test is still used among practitioners today (Porteus, 1950 & Weiβ & Osterland, 2013).

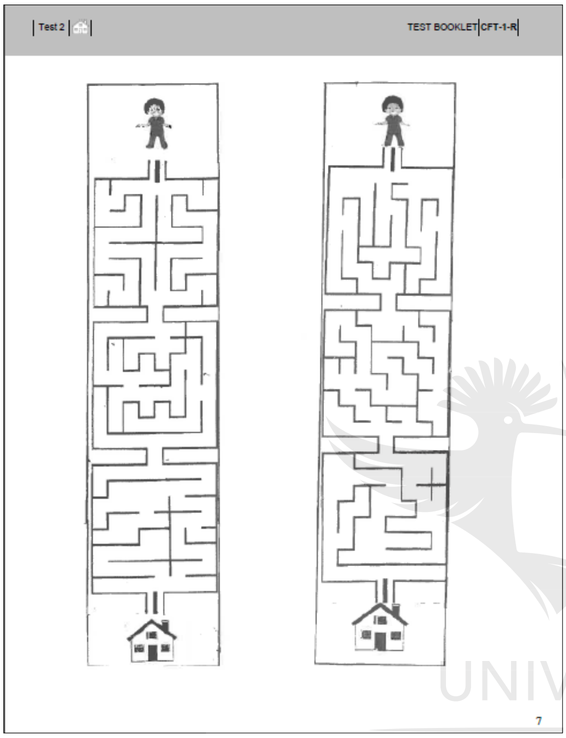

Figure 8. *Labyrinth Three and Four from Subtest Two, CFT1-R, (Weiβ & Osterland, 2013).*

Figure 8 shows the layout of item three and four of Subtest Two, CFT1-R. These two items are on an A4 size page.

Subtest Three – Similarities: In the third subtest of the CFT1-R, children are presented with an image and they are expected to find the corresponding image in the row of five images to the right. The images to the right have had slight changes made to their visual characteristics, however, the context of the image remains the same. This means the child needs to be able to discriminate between the context of the image and the characteristics of the image. The test assesses the child's visuospatial abilities as well as reasoning skills (Weiβ & Osterland, 2013). Figure 9 shows an example of a test item from Subtest Three.

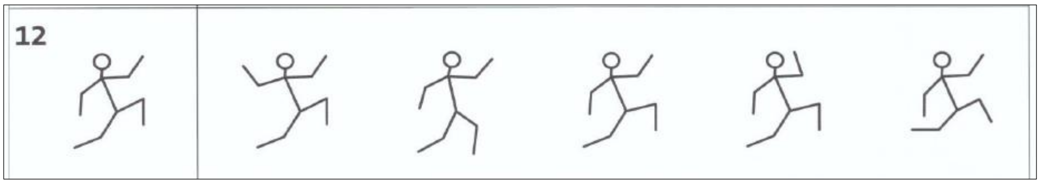

Figure 9. An item from Subtest Three, showing how the context of the image remains the same with only slight changes to the characteristics (Weiβ & Osterland, 2013).

There is a Similarities test in the WISC test battery, however, in this test the child is asked how two words are similar, for example, “How are whales and lions similar?”. The verbal similarities tests assesses verbal reasoning skills as well as verbal concept formation and forms a part of the verbal scale score provided by the WISC (Weschler, 2003b). As outlined in chapter 2, language is a major source of test bias, and the removal of language from an assessment could improve the cultural fairness and validity of said assessment (Van der Vijver & Rothman, 2004). It would therefore go without saying that for the purposes of a culture fair assessment, a visual similarities subtest is more appropriate than the verbal similarities subtest.

Subtest Four: Complete the row. In this subtest the child is presented with a sequence of three pictures.

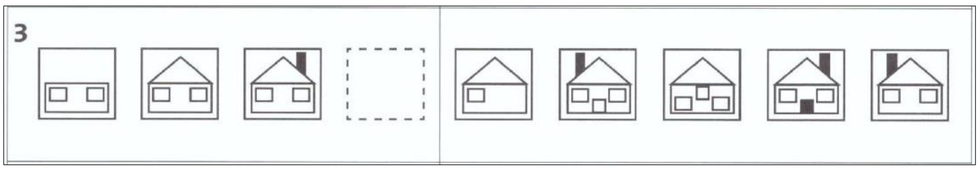

Figure 10. An example from Subtest Four, CFT1-R, (Weiβ & Osterland, 2013).

Test takers need to complete the sequence by selecting a picture that follows on from the rest (Figure 10).

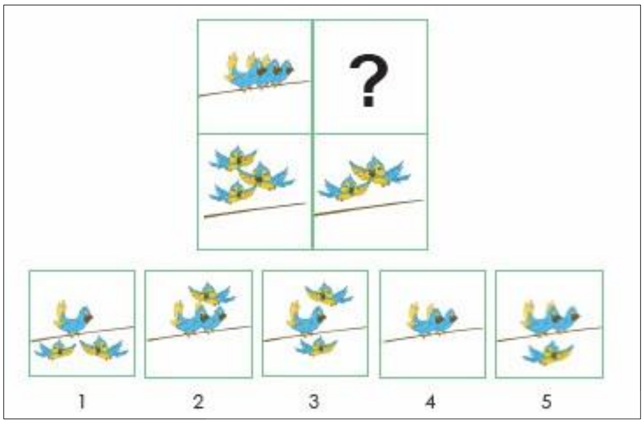

Figure 11.An item from the Matrix reasoning subtest of the WISC, (Weschler, 2003b).

As with the Matrix reasoning Subtest in the WISC (Figure 11), Subtest Four of the CFT1-R assesses visual processing as well as visual, abstract reasoning. (Weschler, 2003b).

Subtest Five: Classifications. This subtest is designed to assess a child's categorical, abstract reasoning, that is, how the child categorises images and views the relationship between them. The child is presented with five pictures, four of which are similar in some way. The child needs to identify the odd one out.

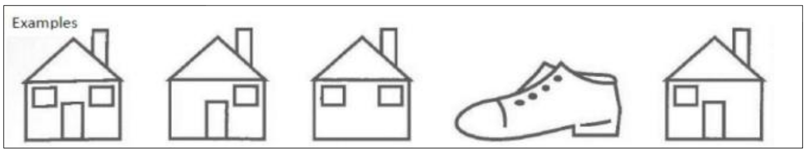

Figure 12. An example from Subtest Five, (Weiβ & Osterland, 2013).

Subtest Five in the CFT1-R (Figure 12) is comparable with the Picture Concepts subtest in the WISC (Figure 13).

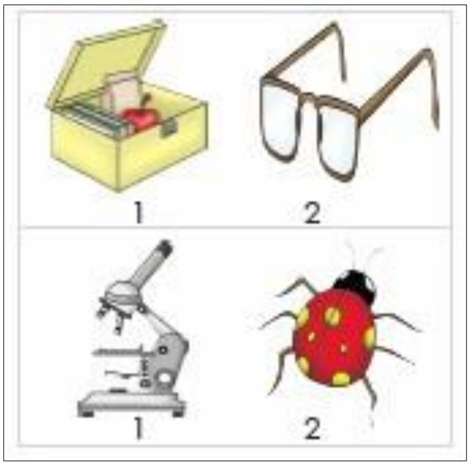

Figure 13. An item from the Matrix reasoning subtest of the WISC, (Weschler, 2003b).

As seen from the Figures 12 (CFT1-R) and 13 (WISC) the CFT images are 2D and were selected for their ability to be recognisable across cultures, whereas the WISC contains items with images that are culturally bound (such as the lunch box or microscope seen in Figure 13) and may lead to culture bias when used in a different cultural setting (Weschler, 2003b).

Subtest Six: Matrices. In this subtest the child is presented with a matrix with three images and the fourth image missing (Figure 14).

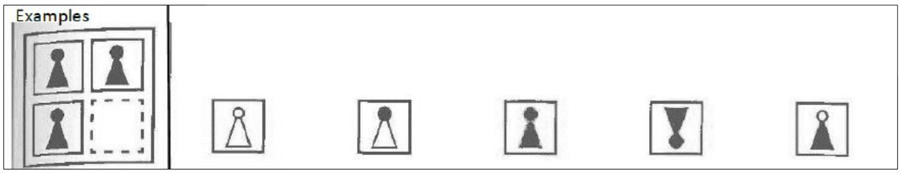

Figure 14. An example from Subtest Six, (Weiβ & Osterland, 2013).

Once this rule has been established, it is tested on the five options (Shye, 1998). The child needs to identify the rule and then apply that rule to identify which of the five options best follows the rule and completes the pattern in the matrices. This test assesses a child's inductive reasoning skills, as the child needs to analyse the three images to establish the rule.

Subtest Six is very similar to subtests in the Raven's Coloured Progressive Matrices (Ravens et al., 1998). An example of an item from the Raven's Coloured Progressive Matrices can been seen in Figure 15.

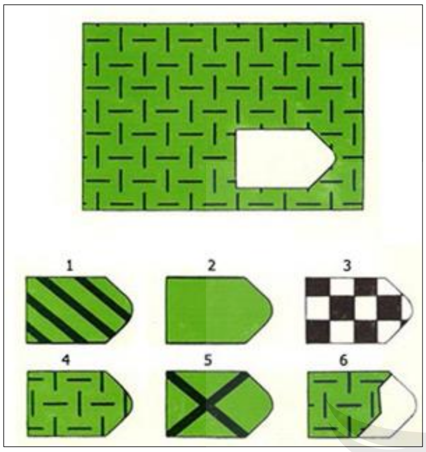

Figure 15. An item from the Raven's CPR, (Ravens, 1998).

There are many similarities between the Ravens CPR and Subtest Six of the CFT1-R. Figures 14 and 15 comprise just one example of the comparable items between these two tests.

3.2.3. Adaptions made to the CFT1-R.

The HPCSA's Policy on the Classification of psychometric measuring devices, instruments, methods and techniques, (2006) warns that it would be "unwise" (p.1) for the assessment profession to not pursue the adaption of existing and development of new, culturally fair tests. Even so, very few culturally relevant tests have been developed in South Africa (Foxcroft, Paterson, Le Roux, & Herbest, 2004). When developing new instruments or revising existing instruments, it is crucial to make allowances for cultural differences and eliminate/minimise cultural bias by adjusting the test items/times. In addition, it is important to ensure they align with the language and culture in which the test will be administered, with oral instruction. Only then can one begin to establish the validity of the test when used in multicultural groups (Van de Vijver & Rothmann, 2004; Foxcroft 2004 and 2011 and Benson, 2003). Prior to this study, an unchanged CFT1-R was administered to 300 grade 1 children. The results of this administration supported the findings cited in chapter 2, i.e., that the administration of tests in a culture outside of the one in which it was developed compromises the validity of the assessment as it is prone to bias.

It was decided that, for the purposes of this study, that the following adaptions would be made.

3.2.3.1. Independent variable: Time

In the original test time limits are outlined in the below table:

Table 1. Allotted time limit per subtest of CFT1-R

Subtest	Time limited (in seconds)
Subtest One	70
Subtest Two	90
Subtest Three	90
Subtest Four	180
Subtest Five	180
Subtest Six	180

Based on the administrator’s observations in the first administration of the CFT1-R it was decided that the time limits used in the German version of the test where incompatible in non-German cultures (discussed in 2.8.5, Time and test taking).

3.2.3.2. Other Adaptions: Practice items.

Originally, the CFT1-R had three examples per test. A further two examples per subtest where provided. The reason behind these additions was to further ensure the participants clearly understand what is required of them (Foxcroft, 2011). Examples allow the learners to familiarise themselves with the task prior to the start of the formal assessment. By providing the participants with an additional two examples, making the total number of examples, per subtest, five, the participants had greater opportunity to understand what was required. This would have decreased their testee anxiety and increased chances of receiving a more accurate reflection of their abilities (Samuels et al., 2015; Truter, 2015; Van der Berg, 2007; Flippo et al., 2018). See paragraph 2.8.6, Time and test taking for an in-depth discussion on this topic.

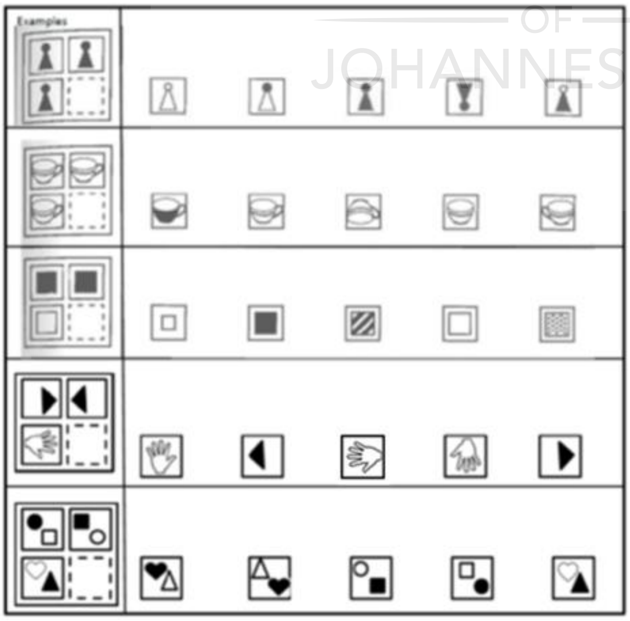

Figure 16.Subtest Six’s example page, showing the original three examples and the additional two (bottom two).

The inclusion of the additional two examples per subtest may have impacted on the results of this study and will be discussed in detail in chapter 5.

3.3. Research Design

The research design of this study is principally that of a psychometrics research project, positioned within a multi-cultural South African context.

3.3.1. Independent variable (s).

As the aim of this research project is to establish what the optimum test time is for South African learners from heterogeneous language groups. The primary independent variable is therefore the extension of the time limit to a maximum of five minutes per subtest. Minor independent variables are home language and gender.

3.3.2. Dependent variable (s).

Table 2 outlines the constructs of each of the CFT1-R subtests:

Table 2. Subtests on CFT1-R

Subtest	Construct
Subtest One: Substitutions	Visual-motor processing speed and short-term memory skills (Weiβ & Osterland, 2013).
Subtest Two: Mazes	Visuo-motor skills and processing speed (Weiβ & Osterland, 2013).
Subtest Three: Similarities	Visuospatial abilities and reasoning skills (Weiβ & Osterland, 2013).

Subtest Four: Complete the row	Visual processing and visual, abstract reasoning. (Weschler, 2003).
Subtest Five: Classifications	Categorical, abstract reasoning (Weiβ & Osterland, 2013).
Subtest Six: Matrices	Inductive reasoning skills (Weiβ & Osterland, 2013).

The scores from the above subtests of the CFT1-R will comprise the dependant variables. The tests will be scored as per the original CFT1-R test manual's instructions. These scores will be in raw score format and not normed scores, as the German norms may not be suitable for a South African context.

3.3.3. Control of extraneous variable (s).

In order to minimise the effect of extraneous variables on the test scores, the following measures were taken:

(1) The test was administered in an empty classroom or hall.

(2) The room was well lit.

(3) Tests were administered in the morning and, wherever possible, did not overlap with the children's break time.

(4) Verbal instructions where given by test administrators who were fluent in the language of assessment—in their home language.

3.4. Sampling

A convenience sampling method was adopted. Convenience sampling is "the least rigorous technique, involving the selection of the most accessible subjects"

(Marshall, 1996, p. 523). Marshall notes that this may cause the data to be less credible than a probable sampling method. However, convenience sampling is popular in pilot research projects because of the ease it allows the researcher (Chaturvedi, n.d; Marshall, 1996).

3.4.1. Participants.

Three schools situated in Johannesburg were selected, an isiZulu and Sesotho medium [3]school, an English medium school and an Afrikaans medium school. 30 learners in each language demographic were randomly selected for assessment. The total sample size was 120 participants. A breakdown of the sample can be seen in Table 3.

Table 3: Language and Gender Composition of Study Participants

Home Language	Total sample size	Male	Female
Sesotho	36	18 (50%)	18 (50%)
Isi-Zulu	37	12 (32%)	25 (68%)
Afrikaans	32	16 (50%)	16 (50%)
English	4	1 (25%)	3 (75%)
Tswana	6	4 (67%)	2 (33%)

[3] Medium refers to the language of tuition.

Venda	2	2 (100%)	0 (0%)
Sepedi	2	1 (50%)	1 (50%)
Xhosa	1	1 (100%)	0 (0%)
Total	120	55 (46%)	65 (54%)

Originally the plan was to assess 30 children from each of the four home language groups. However, upon analysing the data it was found that the children from the English medium school were not necessarily from English home language families. This dilution of the English sample meant that the English, Tswana, Venda, Sepedi and Xhosa sample were too small for statistical analyses. Ideally another sample of English home language children should have been identified and assessed, however, due to funding constraints this was not possible. This is one of the limitations on this research study that is discussed in chapter 5.

3.5. Data Collection Methods

The CFT1-R is a norm referenced assessment and is a German adaption of the American "Culture Fair Intelligence Tests - Scale 1" by R. B. Cattell. A normed intelligence assessment produces information on the testee's ability in relation to a comparison group. Should, however, the testee's context be incomparable to that of the demographic group, the normed score that the assessment produces may not be an accurate reflection of the individual's ability. (Cohen & Swerdlik, 2002; Weiβ & Osterland, 2013).

The CFT1-R can be administered as either an individual or a group test. It was administered as a group test in this study. The head administrator read the translated verbal instruction before each subtest. Then he/she would work through each of the examples, answering any questions that were asked. Once all the children had successfully completed the examples the test began (Weiβ & Osterland, 2013).

3.6. Research Procedure

After obtaining ethical clearance (Appendix 1b) as well as consent from the headmasters of all three schools, I sent out a letter (Appendix 1a) to 30 parents or guardians of the grade 2 children in the isiZulu, Sesotho, Afrikaans and English classes. As the children were all minors, consent from their parents or guardians was needed. The 30 children were divided into three groups of ten each and were assessed as a group, during school time. The administration procedures for each instrument was discussed in paragraph 3.5 of this chapter. The instruments was administered in the same order and the same instructions were given to all groups.

To establish the optimum time limit for each subtest, the following intervention was implemented during administration. Once the official time for the subtest had lapsed, the administrators will mark each child's progress at 30 second intervals. This will continue until the child has finished the subtest, at which point the administrator will make a note of the total time needed to complete the test.

The assessments were scored according to instructions, and the process of data analysis was begun (discussed in section 3.7, Statistical Techniques). Participants, together with their parents, were invited for individual feedback.

3.7. Statistical Techniques

The data was analysed using SPSS. An independent t-test analysis was used to indicate if there was any statistical variation between the standard time limit in the original CFT1-R and the entire sample. Field, (2009) describes an independent t-test as an effective statistical tool when establishing “whether two means collected from independent samples differ significantly” (p.787). The analysis from the above will either support or reject General Hypothesis One. The effect of test times will be examined for the entire sample. I analysed the increased performance per 30 second interval. An independent t-test was used to indicate what the best time would constitute an optimal time for age-appropriate performances. The data was analysed to establish the mean test time for each subtest, thus giving an indication if more time needs to be allocated in the revision of the subtest.

A dependent t-test analysis was used when comparing two means from the same sample (Field, 2009). To establish whether there was significant variation in optimal test time within the sample, the means of each home language group were compared with each other (General Hypothesis Two), the data was subjected to a dependent t-test. I analysed the increased performance per 30 second interval for each of the groups: General Hypothesis Two (viz. Isi-Zulu, Sesotho, Afrikaans).

The results of the dependant t-test would indicate variance (if any) in optimal time in terms of home language.

3.8. Research objective(s)

In view of the research questions and problem statements as outlined above, this research study aims to achieve the following objective (s):

1) To determine what the optimal subtest time limits are on the CFT1-R for a small sample in Johannesburg, South Africa.

2) To determine whether home language is an important variable when determining optimal test time limits.

3.9. Research Hypothesis

In accordance with the problem statement and the aims of this research study, a null hypothesis and alternative hypothesis for each of the independent variables of this study is provided below. The expected hypothesis is indicated with an asterix (*) on the left-hand side of the hypothesis. A rationale is provided for each of the general hypothesis.

3.9.1. General Hypothesis One

*Ho_1: A small sample of grade 2 South African children across heterogenous language groups required more time to complete the subtests in the CFT1-R effectively, than what has been allocated by the German version of the test.

Ha_1: A small sample of grade 2 South African children across heterogenous language groups did not require more time to complete the subtests in the CFT1-R effectively, than what has been allocated by the German version of the test.

3.9.1.1. Rationale.

It is expected that the null hypothesis (Ho_1) will not be rejected for most subtests with time as a variable. South African children are less accustomed to the time pressures in a test environment than German children. This can cause test anxiety, which could also influence their processing speed.

3.9.2. General Hypothesis Two.

*Ho_2: There is little to no statistically significant differences in the optimal test time of the sample with regards to the effect of home language (viz. Isi-Zulu, Sesotho, English, Afrikaans) on the optimal test time.

Ha_2: There is statistically significant differences in the optimal test time of the sample with regards to the effect of home language (viz. Isi-Zulu, Sesotho, English, Afrikaans) on the optimal test time.

3.9.2.1. Rationale.

It is expected that the null hypothesis (Ho_2) will not be rejected for most subtests with home language as a variable. The CFT1-R is a non-verbal test and the instructions do not adhere to a specific language. One may expect minor differences in the optimal time limit between some subtests of the groups tested.

3.10. Ethical considerations

There are several ethical considerations in this study. These have been described by De Roche & De Roche, (2010) and Henning et al., (2011). The considerations, as outlined by the above authors, correspond with the HPCSA's *General Ethical Guidelines for Health Researchers* (2008), specifically the section on the "Duties to the research participants". Below is a brief description of each ethical consideration as well as an explanation of how it will be addressed in this study:

1. *Do not harm participants.* As outlined in chapter 1, special care was taken to ensure that the participants were not subjected to any negative backlash for participating in the study.
2. *Obtain informed consent.* As outlined in section 3.6, informed consent was obtained from each participant's parent or legal guardian. Each potential

participant received a letter attached to the consent form, outlining the parameters of this study.

3. *Maintain privacy and anonymity.* Each participant's name was recorded on their assessment answer booklet. However, to maintain privacy and anonymity, when the assessment was scored, each participant received an identification number, i.e. 001. All data captured and analysed for a specific participant was done using this identification number. The participant's names will not be published.

4. *Give beneficial feedback.* Should the parents request, they will receive feedback on their child's results.

In addition to this, it must be noted that ethical clearance was obtained from the Research Ethics Committee of the Faculty of Education at the University of Johannesburg (Appendix 1b).

3.11. Conclusion

This chapter summarised the research design, sampling, data collection, research procedure, data analysis and ethical considerations pertaining to this study. It also contained a detailed description of the CFT1-R, as well as the adaptions made to the assessment for this study. In chapter 4, I will present and analyse the data that has been collected.

4. Chapter 4: Data Analysis

4.1. Introduction

In this chapter I will present the data collected from the administrations of the adapted CFT1-R. The objective of this chapter is to establish whether the data supports the hypotheses outlined in Chapter 3. My co-supervisor, Prof Fritz-Stratmann, assisted me with my statistical data and is from Germany. Therefore, I used the German version of SPSS for my statistical analyse. All SPSS data was printed in German and has been reproduced below in English. For confirmation of the below summaries, please see the comprehensive German tables in Appendix 2.

4.2. Preliminary Analysis

Each participant's test paper was assigned an ID code ranging from 001 to 120. The demographic information as well as ID code on each test was captured into an Excel spreadsheet. Both the home language and language of tuition was captured, as well as gender (F=female; M=male) and age of each participant. The CFT1-R test was then marked and the raw scores of each participant for each of the six subtests was captured into the Excel Spreadsheet and then entered into SPSS. A one indicated an incorrect answer and a zero indicated a correct answer. Where a participant did not answer, an S was entered and where two or more answers where given, a D was inputted. Descriptive statistics, such as mean and standard deviation were calculated using SPSS statistical software.

4.2.1. Descriptive statistics

120 grade 2 learners were randomly selected from three schools in the Johannesburg area. Initially my intention was that the population would comprise 30 participants from each of the four homogenous subgroups based on their home language, specifically, isiZulu, Sesotho, English and Afrikaans This was discussed at length (in paragraph 5.6, Limitations).

Sample characteristics are presented in Table 4 below. The sample has 46% male and 54% female participants. The average age of the entire sample was 7.1 years. isiZulu speakers made up 30.8% of the entire sample, followed by Sesotho (30%) and finally Afrikaans making up 26.6% of the entire sample. The remainder of the sample was made up of English, Tswana, Venda, Sepedi and Xhosa home language speakers.

Table 4: Gender and Age Composition of Study Participants based on home language

Home Language	Total sample size	Male	Female	Average age (yrs) at time of assessment
Sesotho	36 (30%)	18 (50%)	18 (50%)	7.0
Isi-Zulu	37 (30.8%)	12 (32%)	25 (68%)	7.12
Afrikaans	32 (26.6%)	16 (50%)	16 (50%)	7.35
English[4]	4 (3.3%)	1 (25%)	3 (75%)	7.0
Tswana	6 (5%)	4 (67%)	2 (33%)	7.4
Venda	2 (1.6%)	2 (100%)	0 (0%)	7.0
Sepedi	2 (1.6%)	1 (50%)	1 (50%)	7.5
Xhosa	1 (0.8%)	1 (100%)	0 (0%)	8.0
Total	120	55 (46%)	65 (54%)	7.1

4.3. Research Question One

The initial research questions examined what the optimal time limit is for the CFT1-R subtests, when applied to a small sample of grade 2 students in Johannesburg, South Africa. All the t scores according to time intervals across the subtests will be discussed together, as a similar pattern was noticed across all subtests. Where there are variations, these will be discussed individually.

4.3.1 T test scores.

An independent t-test analysis on correct answers at each time interval, for the entire sample of 120 participants[5], produced the following data. The data for Subtest One is presented in Table 5.

[4] See Paragraph 3.4.1. The 30 English participants were comprised of various home language speakers, not just English.
[5] 120 participants included the participants from the English medium school. This group was not included in the analysis of home language groups (see Paragraph 3.4.1 Participants).

Table 5. Subtest One: Independent T-test results comparing the results from the unadapted version, and the version adaptedfor the purposes of this study

Time Interval	Mean	Std Deviation	Sig (2-tailed)
70-100s	-12,50833	4,55959	0.000
100s-130s	-10,70192	4,85494	0.000
130s-160s	-10,08974	3,11755	0.000
160s-190s	-9,66667	4,14284	0.000
190s-220s	-8,54545	2,85736	0.000
220s-250s	-,6,64286	3,07864	0.000
250s-280s	-,7,00000	4,43471	0.006
280s-290s*	-4,75000	2,50000	0.032

**all participants had completed the test within 290s*

The T scores based on time intervals for Subtest One, Subtest Two, Subtest Three and Subtest Four showed significant differences in correct answers at each time interval. These scores suggest that the participant's performed better the more time that they received to complete the task.

Table 6.Subtest Two: Independent T-test results comparing the results from the unadapted version, and the version adaptedfor the purposes of this study

Time Interval	Mean	Std Deviation	Sig (2-tailed)
90-120s	-1,59664	,93264	0.000
120s-150s	-1,46729	,79287	0.000
150s-180s	-1,27835	,70330	0.000
180s-210s	-1,09211	,76903	0.000
210s-240s	-73333	,68561	0.000
240s-270s	-,78571	,68202	0.000
270s-300s	-,90625	,81752	0.000

In the first four subtests (see Tables 4, 5, 6 and 7) the participants required the entire five minutes to complete the task and giving them less time significantly hampered their ability to complete the tasks effectively.

Table 7. Subtest Three: Independent T-test results comparing the results from the unadapted version, and the version adaptedfor the purposes of this study

Time Interval	Mean	Std Deviation	Sig (2-tailed)
90-120s	-1,20354	,85731	0.000
120s-150s	-1,17045	,83352	0.000
150s-180s	-,89231	,83147	0.000
180s-210s	-,97222	,81015	0.000

210s-240s	-,84211	,83421	0.000
240s-270s	-,42857	,53452	0.078
270s-300s	-1,2000	,83666	0.033

Table 8. Subtest Four: Independent T-test results comparing the results from the unadapted version, and the version adaptedfor the purposes of this study

Time Interval	Mean	Std Deviation	Sig (2-tailed)
180s-210s	-,66667	,81650	0.000
210s-240s	-,40678	,59069	0.000
240s-270s	-,32500	,47434	0.000
270s-300s	-,37037	,68770	0.010

In Subtest Five (Table 9) the t scores show significant differences in scores up until 240 seconds. From 240s – 300s there was little improvement in the test scores. This indicates that the optimum time required to complete Subtest Five is 240s.

Table 9.Subtest Five: Independent T-test results comparing the results from the unadapted version, and the version adaptedfor the purposes of this study

Time Interval	Mean	Subtest Five Std Deviation	Subtest Six Sig (2-tailed)
180s-210s	-,56098	,80774	0.000
210s-240s	-,22727	,42893	0.021
240s-270s	-,41667	,90034	0.137
270s-300s	-,25000	,46291	0.170

In Subtest Six (Table 10) the t scores show significant differences in scores up until 240 seconds. From 240s – 300s there was little improvement in the test scores. This indicates that the optimum time required to complete Subtest Six is 240s.

Table 10. Subtest Six: Independent T-test results comparing the results from the unadapted version, and the version adaptedfor the purposes of this study

Time Interval	Mean	Std Deviation	Sig (2-tailed)
180s-210s	-1,27778	1,07406	0.000
210s-240s	-,50000	,53452	0.033
240s-270s	-,40000	,54772	0.178
270s-300s	*	*	*

*all participants had completed the test within 270s

Based on the results from the independent t tests, the participants in the sample required more time than was allocated in the original CFT1-R. A summary of optimum times based on the above data is presented in Table 11. Please refer to Appendix B for comprehensive SPSS tables.

Table 11. Optimum time limit for all participants based on t-test scores.

Subtest (original time limit)	*Optimum Time Limit in seconds*
Subtest One (70)	*290*
Subtest Two (90)	*300*
Subtest Three (90)	*300*
Subtest Four (180)	*300*
Subtest Five (180)	*240*
Subtest Six (180)	*240*

Based on the above result, when looking at the sample in its entirety, participants benefited from the additional time in all the subtests. In Subtest One, all participants had completed the test within 290 seconds. The full 300 seconds was beneficial for Subtests Two, Three and Four. However, in Subtest Five and Six, the improvement in test performance after 240 seconds was insignificant.

Research Question Two: The aim of research question two was to determine whether home language is an important variable when determining optimal test time limits. The mean for each home language group's performance on each subtest is outlined in Table 12 below.

Table 12. Mean for Study Participants based on home language

	IsiZulu	Sesotho	Afrikaans
Subtest One	73,4865	73,8056	74,6563
Subtest Two	14,1892	13,0000	14,1250
Subtest Three	10,3243	9,9167	10,2500
Subtest Four	5,5946	5,2500	7,0938
Subtest Five	9,1892	9,0556	9,1429
Subtest Six	7,7838	7,3056	8,1250

For the most part there was not significant differences in the mean score of the language groups per subtest. The only meaningful differences occurred on Subtest Two, where Sesotho participants scored lower than isiZulu. The difference was not significant. In Subtest Four, where Afrikaans participants scored significantly higher than the Sesotho participants, a possible reason for the differences could be the participants' attendance of pre-school. This is something that should be addressed in future studies. The t-values are outlined in the below section.

A dependant t test was used to establish the optimum test time specific to each home language group. A dependent t-test analysis on correct answers at each time interval produced the following data.

4.3.3. T test result.

The results of the t test analysis for Subtest One, based on home language are outlined in Table 13 below:

Table 13. Subtest One: Dependant t test scores based on home language (comparing isiZulu, Sesotho and Afrikaans)

		IsiZulu	
Time Interval	Mean	Std Deviation	Sig (2-tailed)
70-100s	-11,97297	4,57946	0.000
100-130s	-10,51515	3,77592	0.000
130-160s	-9,26923	3,44741	0.000
160-190s	-9,21053	4,72086	0.000
190-220s	-7,60000	3,30656	0.000
220-250s	-4,83333	1,47196	0.000
250-280s	-6,60000	4,03733	0.022
280-290s	-5,66667	2,08167	0.042
		Sesotho	
Time Interval	Mean	Std Deviation	Sig (2-tailed)

70-100s	-12,08333	3,84986	0.000
100-130s	-10,17241	7,27063	0.000
130-160s	-10,18182	2,57527	0.000
160-190s	-9,27273	2,32770	0.000
190-220s	-10,40000	1,14018	0.000
220-250s	-8,75000	2,06155	0.003
250-280s	-8,00000	7,07107	0.356
280-290s	*	*	*

		Afrikaans	
Time Interval	Mean	Std Deviation	Sig (2-tailed)
70-100s	-13,65625	5,25240	0.000
100-130s	-11,22222	3,28556	0.000
130-160s	-10,88889	3,49603	0.000
160-190s	-11,14286	5,49025	0.002
190-220s	*	*	*
220-250s	*	*	*
250-280s	*	*	*
280-290s	*	*	*

*all participants had completed the test by this time.

Based on the above results we see a variation among all home language groups for Subtest One. The isiZulu participants' scores improved significantly for the entire 280s, whilst the Sesotho participants only improved up until 250s. The Afrikaans participants had all completed the test by 190s, with their scores improving significantly up until this point.

The results of the t test analysis for Subtest Two, based on home language are outlined in Table 14 below:

Table 14. Subtest Two: Dependent t test scores based on home language (comparing isiZulu, Sesotho and Afrikaans)

	IsiZulu		
Time Interval	Mean	Std Deviation	Sig (2-tailed)
90-120s	-1,78378	,82108	0.000
120-150s	-1,22222	,76012	0.000
150-180s	-1,28571	,71007	0.000
180-210s	-1,08824	,75348	0.000
210-240s	-,77778	,69798	0.000
240-270s	-1,05263	,77986	0.000
270-300s	-,83333	,57735	0.000
		Sesotho	
Time Interval	Mean	Std Deviation	Sig (2-tailed)
90-120s	-1,47222	,97060	0.000
120-150s	-1,55882	,70458	0.000
150-180s	-1,21212	,69631	0.000
180-210s	-1,07692	,89098	0.000
210-240s	-0,72727	,070250	0.000
240-270s	-0,55556	0,51131	0.000
270-300s	-0,94118	0,96635	0.001
		Afrikaans	
Time Interval	Mean	Std Deviation	Sig (2-tailed)
90-120s	-1,71875	1,05446	0.000
120-150s	-1,60000	,76376	0.000
150-180s	-1,45000	,60481	0.000
180-210s	-1,16667	,57735	0.000
210-240s	-,55556	,72648	0.051
240-270s	-,66667	,57735	0.184
270-300s	-1,00000	1,41421	0.500

Based on the above results we see some variation among the home language groups for Subtest Two. The isiZulu and Sesotho participants showed significant improvement throughout the additional time. However, the Afrikaans participants improved significantly up until 210s – 240s. Thereafter improvement was insignificant.

The results of the t test analysis for Subtest Three, based on home language are outlined in Table 15 below:

Table 15. Subtest Three: Dependant t test scores based on home language (comparing isiZulu, Sesotho and Afrikaans)

		IsiZulu	
Time Interval	Mean	Std Deviation	Sig (2-tailed)
90-120s	-1,40541	,76229	0.000
120-150s	-1,19355	,94585	0.000
150-180s	-,72000	,54160	0.000
180-210s	-,91667	1,08362	0.014
210-240s	-1,20000	,83666	0.033
240-270s	*	*	*
270-300s	*	*	*
		Sesotho	
Time Interval	Mean	Std Deviation	Sig (2-tailed)
90-120s	-,97059	,16069	0.000
120-150s	-1,20690	,77364	0.000
150-180s	-1,00000	1,14018	0.001
180-210s	-,84615	,68874	0.001
210-240s	-,88889	,92796	0.021
240-270s	-,66667	,57735	0.184
270-300s	*	*	*
		Afrikaans	
Time Interval	Mean	Std Deviation	Sig (2-tailed)
90-120s	-1,34483	,85673	0.000
120-150s	-1,25000	,78640	0.000
150-180s	-1,23077	,72501	0.000
180-210s	-1,25000	,70711	0.002
210-240s	-,33333	,57735	0.423

240-270s	*	*	*
270-300s	*	*	*

*all participants had completed the test by this time

Based on the above results we see some variation among the home language groups for Subtest Three. The isiZulu and Sesotho participants showed significant improvement up until 240s. The Afrikaans participants improved significantly up until 210s. Thereafter improvement was insignificant.

The results of the t test analysis for Subtest Four, based on home language are outlined in Table 16 below:

Table 16. Subtest Four: Dependent t test scores based on home language(comparing isiZulu, Sesotho and Afrikaans)

		isiZulu	
Time Interval	Mean	Std Deviation	Sig (2-tailed)
180-210s	-,66667	,73380	0.000
210-240s	-,73684	,65338	0.000
240-270s	-,41667	,51493	0.017
270-300s	-,40000	,51640	0.037
		Sesotho	
Time Interval	Mean	Std Deviation	Sig (2-tailed)
180-210s	-,84000	1,06771	0.001
210-240s	-,23810	,43644	0.021
240-270s	-,23529	,43724	0.041
270-300s	-,18182	,40452	0.167
		Afrikaans	
Time Interval	Mean	Std Deviation	Sig (2-tailed)
180-210s	-,61111	,60768	0.001

210-240s	-,38462	,18040	0.054
240-270s	-,44444	,52705	0.035
270-300s	-,80000	1,30384	0.242

Based on the above results we see some variation among the home language groups for Subtest Four. The isiZulu participants showed significant improvement throughout the additional time. However, the Sesotho and Afrikaans participants improved significantly up until 270s. Thereafter improvement was insignificant.

The results of the t test analysis for Subtest Five, based on home language are outlined in Table 17 below:

Table 17. Subtest Five: Dependant t test scores based on home language (comparing isiZulu, Sesotho and Afrikaans)

		isiZulu	
Time Interval	Mean	Std Deviation	Sig (2-tailed)
180-210s	-,61538	,86972	0.025
210-240s	-,25000	,46291	0.170
240-270s	-,33333	,57735	0.423
270-300s	*	*	*
		Sesotho	
Time Interval	Mean	Std Deviation	Sig (2-tailed)
180-210s	-.92308	,95407	0.004
210-240s	-,28571	,48795	0.172
240-270s	-,66667	1,21106	0.235

270-300s	*	*	*

		Afrikaans	
Time Interval	Mean	Std Deviation	Sig (2-tailed)
180-210s	-,25000	.50000	0.391
210-240s	*	*	*
240-270s	*	*	*
270-300s	*	*	*

*all participants had completed the test by this time

Based on the above results we see some variation among the home language groups for Subtest Five. The Sesotho and isiZulu participants improved significantly up until 210s. Thereafter improvement was insignificant. Most of the Afrikaans participants had completed the test in the original test time of 180s.

The results of the t test analysis for Subtest Six, based on home language are outlined in Table 18 below:

Table 18. Subtest Six: Dependent t test scores based on home language (comparing isiZulu, Sesotho and Afrikaans)

		isiZulu	
Time Interval	Mean	Std Deviation	Sig (2-tailed)
180-210s	-1,57143	1,27242	0.017
210-240s	-,600000	,54772	0.070
240-270s	-,33333	,57735	0.423
270-300s	*	*	*
		Sesotho	
Time Interval	Mean	Std Deviation	Sig (2-tailed)
180-210s	-1,28571	,95119	0.012
210-240s	-,50000	,7-711	0.500
240-270s	*	*	*
270-300s	*	*	*
		Afrikaans	
Time Interval	Mean	Std Deviation	Sig (2-tailed)

180-210s	*	*	*
210-240s	*	*	*
240-270s	*	*	*
270-300s	*	*	*

*all participants had completed the test by this time

Based on the above results we see some variation among the home language groups for Subtest Six. The Sesotho and isiZulu participants improved significantly up until 210s. Thereafter improvement was insignificant. All the Afrikaans participants had completed the test in the original test time of 180s so no, t test could be applied as there was no time interval data.

Table 19. Optimum time limit for participants of the CFT1-R based on home language (comparing isiZulu, Sesotho and Afrikaans)

Subtest (optimal time limit for all language groups[6])	Optimum time in seconds		
	IsiZulu	Sesotho	Afrikaans
Subtest One (290)	280	250	190
Subtest Two (300)	300	300	240
Subtest Three (300)	240	240	210
Subtest Four (300)	300	270	270
Subtest Five (240)	210	210	180
Subtest Six (240)	210	210	180

Based on the above t value scores, there is some variation in optimal time,

[6] 120 participants included the participants from the English medium school. This group was not included in the analysis of home language groups (see Paragraph 3.4.1 Participants) as each group was too small for statistical analysis.

based on home language. Table 19 depicts a summary of the optimum time limit for participants, based on home language.

Based on the above results, we can see that there is variation in optimal time limit between home languages, particularly in Subtest One. This variation could be attributed to the quality of pre-school education received by previously disadvantaged groups, as discussed in chapter 2. As the pre-school history of the participants is not known, we cannot conclusively attribute the variation in times to this. However, this trend is one that has been researched (see discussion in chapter 2) and may well have influenced the results of this study. This trend and its impact on cognitive assessments may warrant further study in the future.

Although there are statistically significant variations among the language groups (viz. isiZulu, Sesotho and Afrikaans), all three groups are on par or below the optimal time limit for each subtest based on the entire sample (viz. isiZulu, Sesotho, Afrikaans, English, Tswana, Venda, Pedi and Xhosa). This will be discussed in paragraph 5.3.

5. Chapter 5: Findings and conclusions

5.1 Introduction

This study focused on the adaptation of a timed, culture fair test to improve its efficacy for use on grade 2 learners in the greater Johannesburg region. The test primarily sought to adapt time limits by allowing the participants an extended time limit. In addition, two pictorial examples were added to each subtest to give the participants more practice (see paragraph 3.2.3.2, Other Adaptions: Practise items). Foxcroft (2011) notes that multiple examples assist the leaner in understanding what is required of them to complete the task. The data analysis primarily focused on the entire heterogenous sample. In addition, the data specific to the three major home languages groups (viz. isiZulu, Sesotho, Afrikaans) were also analysed. The data were analysed using both independent and dependent t-tests where applicable

5.2 Research Question One

The first question in this study examined what the optimal time limit was, per subtest of the CFT1-R, based on a small sample of heterogenous language groups, from Johannesburg, South Africa. It was hypothesised that the grade two South African children, across heterogenous language groups, would require more time to effectively complete the subtests in the CFT1-R, than the time that had been allocated by the German version of the test. The data showed statistically significant improvement with extended time across all subtests. However, the improvement was not statically significant for the full 300 seconds in all the subtests. In Subtest One, all participants had completed the test within 290 seconds. In Subtests Two, Three and Four, participants showed significant improvement for the full 300 seconds. The improvement in Subtests Five and Six, were only significant up to 240 seconds and thereafter the improvement was insignificant. This supports the findings of Mandinach et al (2008), that too much time can be as detrimental to the participant as too little (see paragraph 2.8.5, Time and Test Taking).

The above results are supported by the literature reviewed in Chapter Two (see paragraph 2.8, Testing in a South African Context). It is unethical to take an assessment that has been developed in (and for) one culture and administer it, unchanged, in another culture. Despite the fact that the CFT1-R is classified as a test that focuses on culture fairness, it still seems to be inadvertently culturally biased. The presence of language in the instruction, as well as the fact that it is a paper and pencil timed test could be a source of bias (see paragraphs 2.8.4 Problematising psychometrics in South Africa, 2.8.5, Bias and paragraph 2.8.6, Time and Test Taking). This once again proves that the probability of a truly culture fair/unbiased intelligence test is a misnomer. This is reiterated at length in the literature by a plethora of theorists and researchers (Benson, 2003; Foxcroft, 2004 & 2011) and this study, yet again, substantiates their claims.

There was general consent, in the literature, that the use of an un-adapted Euro-American intelligence test in a South African context is both un-ethical and un-constitutional (Benson, 2003; Foxcroft 2004 & 2011; Foxcroft et. al., 2004; Nsamenang 2007; Van der Viver & Rothman 2004). The literature reviewed in Chapter Two formed the foundation for Hypothesis Ho_1 (see paragraph 3.9.1, General Hypothesis One). The data collected from this research study supported this hypothesis, South African children require more time to complete the CFT1-R effectively and with scores that more accurately reflect their abilities.

5.3 Research Question Two

The second question in this study examined if home language had an influence on optimal test time for a small sample from Johannesburg, South Africa.

It was hypothesised that there would be little to no statistically significant differences in the optimal test time of the sample with regards to the effect of home language (viz. Isi-Zulu, Sesotho, Afrikaans) on the optimal test time. However, statistically significant variations between the home language groups optimal time limit was found. This would appear to imply that home language does influence the optimal time needed to complete the CFT1-R assessment. However, home language, per se, might not be the cause of discrepancies (as instructions were given in the children's home languages), but rather the culture of schooling that differs among language groups.

Based on the literature discussed in Chapter Two (see paragraph 2.8.6, Time and Test Taking), there is a trend in South Africa where the pre-school education in previously disadvantaged areas is not at the same level as pre-schools in previously advantaged areas. The impact here is twofold. Firstly, learners in previously disadvantaged areas have limited access to pre-schooling. Secondly, the level of pre-school education received in the rural areas is often characterised by inadequate resources and poor teacher training. In other words, children from more disadvantaged areas, assessed in this study, would be less likely to attend pre-school and would therefore have had no introduction to the culture of schooling, as those who would have attended pre-school. Pre-school in this sense also does not imply just attending a formal pre-school before the grade r (kindergarten) year, but also includes the kindergarten year itself. This then means that the first time some children encounter any form of schooling (including sitting in a classroom, manipulating writing utensils, paper and books), would be in grade 1. Other research conducted in South Africa comments on the impact of this on children's readiness for school and learning in the entire foundation phase (Samuels et al., 2015, Truter, 2015, & Van der Berg, 2007). In contrast there are many privileged children who attend excellent pre-schools and also a formal grade r (kindergarten) programme. The different sub-samples from this study seemed to have been children from both these backgrounds and that could have accounted for the discrepancy in performance across language groups.

The exposure to schooling also impacts on the child's familiarity with a formal testing situation, which Cohen (2006) refers to as test wiseness (see paragraph 2.8.6, Time and Test Taking) and in turn impacts on their capacity to perform at a level that accurately reflects their abilities. It is participants from the previously disadvantaged home language groups (viz. Isi-Zulu and Sesotho) that required more time in comparison to the other language group (viz. Afrikaans). While General Hypothesis Ho_2 (see paragraph 3.9.2, General Hypothesis Two) was rejected by the data, the explanation that this can be attributed to level of exposure to pre-school education rather than home language, is interesting and warrants further study.

There is also another very important factor to consider when one investigates the discrepancy among results from the various language groups and this could indeed be attributed to home language itself, but also coincides with the difficulty around attending pre-schooling as described before. What I am referring to is the tendency in South Africa that children from a different language background attend an English medium of instruction school. Added to this it might be the child's first time in a formal classroom. Thus, you have a scenario of a child that comes from a home language background (other than English), being taught in English for the first time in grade 1 and it is the first time that the child enters a formal schooling situation. Thus, there are many things impacting on the child's exposure to schooling and learning in a new language (Chikiwa & Schäfer, 2018; Grobler, 2018; Selati, Adler, Reed, & Bapoo, 2002).

Upon examination of the optimal time limits of the three home language groups (viz. Isizulu, Sesotho, Afrikaans) in comparison to the optimal times based on the data collected from the entire sample (see Table 19), these groups required the same or less than the optimal time for 120 participants. It is possible that the participants from the English medium school required more time than the isiZulu, Sesotho and Afrikaans children, as they received their education (and test instructions for this study) in English, which is not their home language. This is further expatiated by the fact that it is highly probable that had any of these participants received pre-schooling (even though this was not formally included in the data and is a limitation of the study – see paragraph 6, Limitations), it would have been in their home language and not English (Moletsane, 1996). Stein (2017) and Stoop (2017) note the importance of mother tongue (home language) education, especially in the formative years. However, the prevalence of non-English home language speakers in English medium schools can be attributed to: firstly, the lack of home language medium schools due to practicality, including but not limited to lack of funding and lack of resources (Stoop, 2017); and secondly the widespread use of English in higher education and the job market in South Africa (Stein, 2017).

5.4 Summary of findings

The aim of this study was to determine whether the adaption of an intelligence test, could improve its culture fairness for use in a multicultural South African context.

General Hypothesis Ho_1 was proven, as the data revealed that Grade Two children in South Africa require more time to complete each subtest, than that allowed for in the original CFT1-R.

Upon examination of the data, General Hypothesis Ho_2 was rejected, as it was found that there were significant variations in optimum time amongst the home language groups (viz. isiZulu, Sesotho, Afrikaans). Despite this, valuable conclusions can still be drawn from the results of Research Question Two. The effect that pre-schooling and home language tuition may have on intelligence tests were highlighted as possible conclusions (see paragraphs 5.3, Research Question Two and 5.5, Implications of Research).

Both research question one and two were answered in part, however there are some limitations which will be discussed in paragraph 5.6, Limitations.

5.5 Implications of Research

One of the purposes of this study was to fill the void found in research with regards to psychometric test adaptation in South Africa. Especially in terms of intelligence testing, there is a great need for not only more test development, but also to find suitable tests to develop and norm for the South African population. Based on the data results it is evident that the adaption of time limits has improved the cultural fairness of the CFT1-R when used with grade 2 learners in Johannesburg, South Africa.

As this is a pilot study, a major implication of this study is that it can form a foundation for future studies to not only look at larger samples, but also, to make further adaptions to the CFT1-R, to identify and adapt other culturally loaded aspects of the assessment. These could include culturally loaded images in the test items or test instructions.

Lastly, what this study highlights is the misnomer of a culture fair intelligence test. I argue that once again, this study confirms that culture bias cannot be totally avoided in an intelligence test. This is primarily due to the fact that no test can be truly language free as the instructions, whether delivered through spoken or written word, are delivered through language (Benson, 2003). Geary, (2015) and Hall, (2013) note that language is founded in the culture of its origin. Foxcroft & Aston (2006) note that language is the main distinguisher between cultures and its presence has an undeniable impact on an intelligence test.

Furthermore, each culture has their own understanding of what defines intelligence. And one cannot extricate culture or environment when examining what intelligence means to different people (Carey, 2009; Mpofu, 2002). It, therefore, goes without saying that an intelligence test developed in a Euro-American culture will assess a Euro-American understanding of intelligence. Therefore, it cannot be successfully used in an African culture, and vis versa (Nsamenang, 2007). This was discussed at length (in paragraph 2.5.1 Indigenous knowledge systems and intelligence).

Based on the literature discussed in Chapter Two, I would argue that that the idea of a 'culture free' intelligence test is unattainable for the two reasons discussed: 1) Language is culturally loaded, and will always be present, even in a language free test. 2) The culture within which a test is developed will impact on the test development, even if it is indirectly. A more realistic goal would be that of a culture fair test, where the culturally loaded aspects have been adapted, whether through the adaption of time limits, inclusion of examples, adapted norms and or removal or adaptions of biased items (Cole, 2009).

5.6 Limitations

This study did have some limitations. The first limitation was the small sample size. But as this was a pilot study, it was never intended to have a large sample, but rather to start a discussion on the adaption of the CFT1-R for South African use.

Another limitation was the composition of the English school students. The 30 participants were comprised of not only English home language speakers, but also isiZulu, Sesotho, Venda, Pedi, Xhosa, Tswana and Afrikaans. The sample size for each of these home languages was too small for statistical analysis and therefore had to be removed from the data for Research Question Two, thus decreasing my sample size even further. Despite being assured by the principals that the participants were all English home language speakers, once I started analysing the data it was found not to be the case. By the time this had been discovered, I could not go and assess another group of English participants, as the time frame for assessment had expired. However, as this is a pilot study, it is intended to highlight issues in research such as that in a culturally diverse context, such as the one found in South Africa, you cannot be assured that the language of tuition and home language will be the same.

Another limitation was the lack of information pertaining to the level of pre-school education received by the participants. Had this been established it would have allowed me to say conclusively whether the level of pre-school education did in fact affect the results.

The use of convenience sampling, whilst a popular choice for pilot studies, is a limitation of this study. This sampling method has impacted the study in terms of generalisability, and the results is unlikely to be representative of the population (Marshall, 1996).

A final limitation of this study was that the data collection was unable to determine whether or not the additional pictorial examples improved the culture fairness of the CFT1-R. Based on Foxcroft (2011), a presumption can be made that the additional examples may have improved the learners' understanding of the task. But this cannot be conclusively stated.

The above limitations highlight the reasons pilot studies are conducted in the first place. Leon, Davis & Kramer (2011) note that the purpose of a pilot study is to "inform feasibility, which in turn, is instructive in that it points to modifications needed in the planning and design of a larger efficacy trial" (p. 5). Pilot studies leave a margin for error to investigate feasibility of the research in the field. Limitations in a pilot study are not only expected but need to be embraced. They are a great asset, as they inform future studies, increasing their efficacy (Leon, Davis & Kramer, 2011).

5.7 Considerations for Future Research

There are various areas to consider for future research. One area would be to identify culturally loaded items that may result in bias in the CFT1-R. Research shows that the culture of a population will affect how that population will define, and express their intelligence (Flynn, 2007; Mpofu, 2002). Consequently, by administering an assessment in a culture that the said tool have not been developed for, undermines the validity of the test because of cultural bias (Foxcroft et. al., 2004; Van der Viver & Rothman 2004; & Nsamenang 2007). To improve the culture fairness of the CFT1-R, research needs to be done to identify any cultural bias and make the necessary adaptions (HPCSA, 2006). While this pilot study has addressed the issue of adapting the time limit, the issue of culturally loaded items or the development of South African norms fell outside of my scope. Future research addressing these areas will only add to increasing the cultural fairness of the CFT1-R for use in South Africa.

Another area to examine for future research is how home language vs. secondary language tuition would play a role in children's ability to perform in an intelligence assessment. Stein (2017) and Stoop (2017) agree that the language of tuition impacts a child's ability to receive knowledge as well as their ability to express that knowledge. I want to add that it also influences the way in which language is processed in between the reception and expression stage. Not only in receptive and expressive vocabulary of language usage, can multi-lingual tuition become an issue. But also, in a child's ability and speed in processing output, where tuition language and home language differ (Stein, 2017; Stoop, 2017). Further research in this area will only benefit the discourse on this topic, which is currently taking place in South Africa, as current research is inconclusive.

Thirdly, it would be interesting to examine whether the quality of pre-school education received by children impacts on results. Samuels et al, (2015) shows that there while there is a positive correlation between grade r and future scholastic ability, in poorer areas this is almost negligible due to instruction quality, infrastructure and access to learning materials.

A final area of research would be to extend this study to a larger sample. This would enhance the generalisability of the study. It could also potentially lead to findings that have more significant results.

5.8 Summary of the study

In summary, the field of psychometric assessments in South Africa, has its foundation in the oppressive history of our country. Tests were developed with the intention to discriminate against people from certain cultures and races (HPCSA, 1997; Nsamenang, 2007). The mandate given to the South African psychology fraternity, by the South African Constitution, is to adapt, translate and develop measures that do not discriminate against an individual based on race, gender or language (Van de Vijver & Rothmann, 2004).

It was with the above in mind that this pilot study was developed. With the hope that it would form the foundation for, future, larger studies into the adaptation of the CFT-1R for use in the multi-cultural context, that is South Africa.

The data generated from this study supported the literature that the adaptation of a culture fair assessment improves the culture fairness of the assessment. It is evident from the data that South African grade 2 learners require more time than had been allocated in the original CFT1-R. In addition, there were significant variations in optimal test time among the home languages. The conclusions drawn from these results was the possible impact of the quality of pre-school education received in previously disadvantaged areas, thus affecting the learners' exposure to formal education and testing situations. A further possible conclusion is the influence of receiving tuition in a language other than the leaner's home language.

The limitations discussed in paragraph 5.6, Limitations, were the small sample size and lack of information on pre-school education and language of tuition. Had this research examined whether participants received quality pre-school education, and in which language this was provided, may have resulted in more conclusive results.

However, as these limitations have been identified in the pilot study, they serve to improve the efficacy of future, larger studies.

It is evident from the literature reviewed in chapter 2, as well as in the results of this study, that a truly culture free test is not attainable for various reasons (see paragraph 5.5, Implications of research). A culture fair test is a more realistic goal, however the time, cost and effort required to adapt existing tools means this is a long road for test developers, academics and psychologists to walk (Cole, 2009; de Beer, 2017; Foxcroft et al., 2004). Nonetheless, as practitioners, we are legally and ethically bound to strive to reach this goal, even if it is never fully realised (Van de Vijver & Rothmann, 2004). The effort at arriving at a solution can be summed up in the words of the South African author, JM Coetzee (1987), “to plant a sign or marker in the ground where I stand, so that in my future wanderings I shall have something to return to, and not get worse lost than I am” (p. 135–136). This pilot study is such a marker that strives to unpack the elements of a culture fair test and the ways that it might be realised in the future.

5.9 Conclusion

The purpose of this study was to start the dialogue on whether an adapted version of the CFT1-R would be feasible in a multi-cultural, South African context. Chapter One elaborates on the importance of the study, orientating the reader to the context that underpins the study. Chapter Two contains a literature review, outlining the various theories and limitations that influence the discussion on intelligence, cognitive ability, intelligence testing, and psychometric testing in South Africa. Chapter Three presents a discussion on the CFT1-R as well as outlining the reason, design methodology and how the research will be statically analysed. Chapter Four presents the results of the assessment after statistical analysis has been performed. Chapter Five concludes the study with a discussion of the results presented in Chapter Four. It also included a discussion on the limitations and implications of the study, suggestions for future research and a summary of the study.

6. Reference List

Andres, L. (2012). *Designing and doing survey research*. Sage.

Andreasen, N. C., Flaum, M., Swayze, V., O'Leary, D. S., Alliger, R., Cohen, G., ... & Yuh, W. T. (1993). Intelligence and brain structure in normal individuals. *American Journal of Psychiatry*, *150*, 130-130.

Bass, N. (2000). The Raven's Coloured Progressive Matrices Test: A Pilot Study for the Establishment of Normative Data for Xhosa-speaking Primary School Pupils in the Grahamstown Region. M.A. Thesis, Department of Psychology, Rhodes University, Grahamstown, South Africa

Barnhardt, R. (2005). Indigenous knowledge systems and Alaska Native ways of knowing. Anthropology & education quarterly, 36(1), 8-23.

Beech, J.R. and Singleton, C., (1997). The psychological assessment of reading. Psychology Press.

Benson, E. (2003a). Intelligence across cultures: Research in Africa, Asia and Latin America is showing how culture and intelligence interact. *Monitor on Psychology,* 34, 56.

Benson, E. (2003b). Intelligent intelligence testing. Monitor on Psychology, 34(2), 48-58.

Bond, T. G., & Fox, C. M. (2013). *Applying the Rasch model: Fundamental measurement in the human sciences*. Psychology Press.

Carpenter, A. D. (2010). "*Embodied Intelligent: Animals and Us*" Phronesis 55: 281-303

Carey, S. (2009). *Origins of Concepts*. Oxford. Oxford University Press.

Carroll, J. B. (1984). Raymond B. Cattell's contributions to the theory of cognitive abilities. *Multivariate Behavioral Research*, *19*(2-3), 300-306.

Carlson, J. S., & Jensen, C. M. (1980). The factorial structure of the Raven Coloured Progressive Matrices test: A reanalysis. *Educational and Psychological Measurement*, *40*(4), 1111-1116.

Cattell, R. (1987). *Intelligence: Its Structure, Growth and Action: Its Structure, Growth and Action.* Amsterdam. Elsevier Science Publishing Company, Inc.

Caviola, S., Carey, E., Mammarella, I. C., & Szucs, D. (2017). Stress, time pressure, strategy selection and math anxiety in mathematics: A review of the literature. *Frontiers in psychology*, *8*, 1488

Chaturvedi, K. (n.d.) *Sampling Methods*.

Chikiwa, C., & Schäfer, M. (2018). Promoting Critical Thinking in Multilingual Mathematics Classes through Questioning. *EURASIA Journal of Mathematics, Science and Technology Education*, *14*, 8.

Cockcroft, K. (2004a). Theories of cognitive development: Piaget, Vygotsky, and information-processing theory. In D Hook, J Watts & K Cockcroft. (Eds.), *Developmental Psychology* (pp. 176-199). Landsdowne: UCT Press.

Cockcroft, K. (2004b). Intellectual Development. In D Hook, J Watts & K Cockcroft. (Eds.), *Developmental Psychology* (pp. 201-217). Landsdowne: UCT Press.

Coetzee, J.M. *Foe*. New York: Penguin, 1987.

Cohen, R. J., & Swerdlik, M. E. (2002). Psychological testing and assessment: An introduction to tests and measurement (5th ed.). McGraw-Hill Companies, Inc

Cohen, J. (2006). Social, emotional, ethical, and academic education: Creating a climate for learning, participation in democracy, and well-being. *harvard educational Review*, *76*(2), 201-237.

Cole, M. (2009). The illusion of culture-free intelligence testing.

Craik, F. I., & Bialystok, E. (2006). Cognition through the lifespan: mechanisms of change. *Trends in cognitive sciences*, *10*(3), 131-138.

Creswell, J. W., & Clark, V. L. P. (2007). Designing and conducting mixed methods research.

Creswell, J. W., Klassen, A. C., Plano Clark, V. L., & Smith, K. C. (2011). Best practices for mixed methods research in the health sciences. Bethesda, MD: *National Institutes of Health*, 10.

Creswell, J. W. (2013). Steps in conducting a scholarly mixed methods study.

Dawes, A, & Biersyeker, L. (2011). Early Childhood Development. Nsamenang, A and Tchombe, T. (Eds.), *Handbook of Educational Theories and Practices: A Generative Teacher Education Curriculum* (pg 111-123). Bamenda, North West Region: Human Development Resources Centre

Department of Basic Education, 2015. *Basic Education on improving qualifications for Grade R teachers*. Retrieved from https://www.gov.za/speeches/department-basic-education-working-hard-improve-qualifications-grade-r-teachers-4-aug-

2015 on 21 August, 2018.

De Beer, J. (2017, January 25**)**. History of Psychometric Testing in South Africa [Blog post]. Retrieved from https://jvrafricagroup.co.za/history-psychometric-testing-south-africa/

De Roche, J. E., & de Roche, C. (2010). Ethics. In A. J. Mills, G. Durepos & E. Wiebe (Eds.), Encyclopedia of case study research. (pp. 337-345)

Donald, F., Thatcher, A., & Milner, K. (2014). Psychological assessment for redress in South African organisations: is it just?. *South African Journal of Psychology*, *44*(3), 333-349.

Duncan, J. (2001). An adaptive coding model of neural function in prefrontal cortex. *Nature Reviews Neuroscience*, *2*(11). Efron, S. E., & Ravid, R. (2013). *Action research in education: A practical guide*. Guilford Press.

Dusek, J. B. (1980). The development of test anxiety in children. In Irwin, B.S (Ed.), Test Anxiety: theory, research and applications. (p.88). Hillsdale, NJ: Lawrance Erlbaum Associates, Inc.

Elliot, L. (1999). What's going on in there. *How the brain and mind develop in the first five years of life*, 294-295.

Efron, S. E., & Ravid, R. (2013). *Action research in education: A practical guide*. Guilford Press.

Evans, F. R. (1980). A study of the relationship among speed and power aptitude test scores, and ethnic identity. College Board Research and Development Report, RDR 80–81, No. 2. Princeton, NJ: ETS.

Field, A. P. (2009). Discovering statistics using SPSS:(and sex and drugs and rock 'n'roll). Los Angeles [ie Thousand Oaks, Calif.].

Fischer, K. W., & Silvern, L. (1985). Stages and individual differences in cognitive development. *Annual Review of Psychology*, *36*(1), 613-648.

Fleet A., Harcourt D. (2018) (Co)-Researching with Children. In: Fleer M., van Oers B. (eds) International Handbook of Early Childhood Education. Springer International Handbooks of Education. Springer, Dordrech

Flippo, R. F., Appatova, V., & Wark, D. M. (2018). test Preparation and test taking. *Handbook of College Reading and Study Strategy Research*, 241.

Foxcroft, C. (2004). Planning a psychological test in the multicultural South African context. *SA Journal of Industrial Psychology*, 30 (4), 8 – 15.

Foxcroft, C., Paterson, H., Le Roux, N. & Herbest, D. (2004). Psychological assessment in South Africa: A needs analysis. Pretoria: Human Sciences Research Council.

Foxcroft, C. D., & Aston, S. (2006). Critically examining language bias in the South African adaptation of the WAIS-III. *SA Journal of Industrial Psychology*, 32(4).

Foxcroft, C. and Roodt, G. eds., 2006. An introduction to psychological assessment in the South African context. London, UK: Oxford University Press

Foxcroft, C. D. (2011). Ethical Issues Related to Psychological Testing in Africa: What I Have Learned (So Far). Online Readings in Psychology and Culture, 2(2). http://dx.doi.org/10.9707/ 2307-0919.1022.

Flynn, J. (2007). *What Is Intelligence?: Beyond the Flynn Effect*. Cambridge: Cambridge University Press

Fritz, A., Ricken, G., & Balzer, L. (2009). Warum fällt manchen Kindern das Rechnen schwer? Entwicklung arithmetischer Kompetenzen im Vor-und frühen Grundschulalter. *Fördernder Mathematikunterricht in der Sekundarstufe I. Rechenschwierigkeiten erkennen und überwinden*, 12-28.

Gazzaniga, M., Ivry, R., Mangun, G. (2002). *Cognitive Neuroscience: The Biology of the Mind.* NewYork: W.W. Norton and Company.

Geary, D.C. (2015). Social Competition and the Evolution of Fluid Intelligence. *Handbook of Intelligence: Evolutionary Theory, Historical Perspective and Current Concepts.* (pp. 105-120). Springer New York

Geake, J. G. (2008). The Neurological Basis of Intelligence: A Contrast with 'Brain-based' Education. British Educational Research Association Conference, University of Glamorgan.

Gopnik, A. (2009). *The philosophical baby: What children's minds tell us about truth, love & the meaning of life*. Random House

Government Gazette, 1998, Republic of South Africa, Vol. 400, no. 19370, Cape Town, 19 October 1998.

Grobler, R. (2018). Students' Perceptions of Code-Switching in Natural Sciences Classrooms: A South African Perspective. *Africa Education Review*, *15*(1), 38-51.

Grossberg, S. (2000). *Linking mind to brain: the mathematics of biological intelligence*. Boston University, Center for Adaptive Systems and Department of Cognitive and Neural Systems.

Guildford, J. P. (1967). Creativity: Yesterday, today and tomorrow. *The Journal of Creative Behavior*, *1*(1), 3-14.

Haier, R. J., White, N. S., & Alkire, M. T. (2003). Individual differences in general intelligence correlate with brain function during nonreasoning tasks. *Intelligence*, *31*(5), 429-441.

Hall, J. K. (2013). *Teaching and researching: Language and culture*. Routledge.

Hanson, J. (2012). *WPPSI-IV: Overview* [PowerPoint presentation]. Retrieved from https://d3gqux9sl0z33u.cloudfront.net/AA/AN/maspwebcom/downloads/249233/WPPSI_IV_3_hour_Overview_Handout.ppt Hanson WPPSI

Health Professions Council of South Africa. (2006a). *Form 208. Policy on the Classification of Psychometric measuring devices, instruments, methods and techniques.* Retrieved February 23, 2015 from http://www.hpcsa.co.za/Uploads/editor/UserFiles/downloads/psych/psycho_policy/form_208.pdf

Health Professions Council of South Africa. (2006b). *Form 223. Ethical Rules of Conduct for Practitioners Registered under The Health Professions Act, 1974.* Retrieved September 25, 2017 from http://www.hpcsa.co.za/downloads/conduct_ethics/rules/ethical_rules_psychology.pdf

Health Professions Council of South Africa. (2017). *Board Notice 155 of 2017. List of Classified and Certified Psychological Tests under The Health Professions Act, 1974.* Retrieved September 25, 2017 from

http://www.hpcsa.co.za/Uploads/editor/UserFiles/downloads/psych/List_of_Classified_tests_Board_Notice_155_of_2017.pdf

Hill, K. T., & Wigfield, A. (1984). Test anxiety: A major educational problem and what can be done about it. *The Elementary School Journal*, *85*(1), 105-126.

Huesman, R. L. and Frisbie, D. (2000). The validity of ITBS reading comprehension test scores for learning disabled and non learning-disabled students under extended-time conditions. Paper presented at the Annual Meeting of the National Council on Measurement in Education, New Orleans, Los Angeles, 24–28 April.

Huysamen, G.K. (1996). Psychological measurement. An introduction with South African examples. Pretoria: J.L. van Schaik.

Jung, R. E., & Haier, R. J. (2007). The Parieto-Frontal Integration Theory (P-FIT) of intelligence: converging neuroimaging evidence. *Behavioral and Brain Sciences*, *30*(2), 135-154.

Kaku, M. (2010). *The Origin of Intelligence.* Available online at: http://bigthink.com/big-think-tv/the-origin-of-intelligence.

Kaku, M. (2014). *The Future of the Mind: The Scientific Quest to Understand, Enhance, and Empower the Mind*. New York. Doubleday.

Kaufman, A. (2009). *IQ Testing 101*. New York. Springer Publishing.

Kelly, M. (2004). *The rhythm of life: Living every day with passion and purpose*. Simon and Schuster.

Klauer, K.J., & Willmes, K. (2002). Inducing inductive reasoning: does it transfer to fluid intelligence? *Contemporary Educational Psychology*, 27, 1–25.

Kvist, A & Gustafsson, J. (2008). The relation between fluid intelligence and the general factor as a function of cultural background: A test of Cattell's Investment theory. *Intelligence,* 36, 422–436.

Laher, S., & Cockcroft, K. (2013). Contextualising psychological assessment in South Africa. *Psychological assessment in South Africa: Research and applications*, 1-16.

Levitt, P. (2009). Brain architecture. Retrieved from http://developingchild.harvard.edu/science/key-concepts/brain-architecture/

Leon, A. C., Davis, L. L., & Kraemer, H. C. (2011). The Role and Interpretation of Pilot Studies in Clinical Research. *Journal of Psychiatric Research*, *45*(5), 626–629. http://doi.org/10.1016/j.jpsychires.2010.10.008

Linan-Thompson, S. (2014). Screening of School Readiness Skills: A review of the literature. United States Agency for International Development.

Lupyan, G. (2015). The centrality of language in human cognition. Language Learning.

Madge, E. M., Van den Berg, A. R., & Robinson, M. (1985). Manual for the junior South African individual scales (JSAIS). *Pretoria: Human Science Research Council*.

Marshall, M. N. (1996). Sampling for qualitative research. *Family practice*, 13(6), 522-526.

Mandinach, E. B., Bridgeman, B., Cahalan-Laitusis, C., Trapani, C., College, B., & Educational Testing, S. (2005). The Impact of Extended Time on SAT® Test Performance. Research Report No. 2005-8. ETS RR-05-20. *College Board* https://files.eric.ed.gov/fulltext/ED563027.pdf (Retrieved on 2018/08/14).

Mawila, D. (2012). *An explorative investigation of the quality of items of the performance scales on the translated Sesotho version of the junior South African individual scales JSAIS (giq-8)* (Doctoral dissertation, University of Johannesburg).

Mdluli, P. O. (2011). IsiZulu Translation of the Junior South African Individual Scales Answer Booklet (HSRC, 2004).

Men, W., Falk, D., Sun, T., Chen, W., Li, J., Yin, D., Zang, L. and Fan, M., 2014. The corpus callosum of Albert Einstein's brain: another clue to his high intelligence?. *Brain*, *137*(4), pp.e268-e268.

Miles, M. B., Huberman, A. M., Huberman, M. A., & Huberman, M. (1994). *Qualitative data analysis: An expanded sourcebook*. sage.

Miller, R., Pascual-Leone, J., Campbell, C., & Juckes, T. (1989). Cross-cultural similarities and differences on two neo-Piagetian cognitive tasks. *International Journal of Psychology*, *24*(1-5), 293-313.

Moletsane, M. E. (1996). Literacy education for pre-primary children in an informal daycare centre in Mamelodi.

Morra, S., Gobbo, C., Marini, Z., & Sheese, R. (2012). *Cognitive development: neo-Piagetian perspectives*. Psychology Press.

Mpofu, E. (2002). Indigenization of the psychology of human intelligence in Sub-Saharan Africa. In W. J. Lonner, D. L.

Murphy, R. (2008). *Models for EFL theory and methodology derived from a sir-based pilot study on Japanese cognitive development* (Master's thesis, University of Birmingham, Birmingham, England). Retrieved from http://www.birmingham.ac.uk/Documents/college-artslaw/cels/essays/matefltesldissertations/Thesis-R-Murphy.pdf

Naicker, A, (2013) An investigation of the quality of the test items on the numerical scales of the isiZulu translated version of the junior South African intelligence scales (JSAIS) (G-IQ8). M.ed. Psych (Educational Psychology). [Unpublished]: University of Johannesburg. Retrieved from http://www.ujdidispace.uj.ac.za (Accessed: 2018/10/11).

Neisser, U., Boodoo, G., Bouchard Jr, T. J., Boykin, A. W., Brody, N., Ceci, S. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. American psychologist, 51(2), 77.

Ngara, C. (2007). African ways of knowing and pedagogy revisited. *Journal of Genetic Psychology*, 155(3), 341-354.

Nsamenang, A. B. (2007). Origins and development of scientific psychology in Afrique Noire. Psychology: IUPsyS global resource (Edition 2007). International Journal of Psychology, 42.

Nell, V. (1994). Interpretation and misinterpretation of the South African Wechsler-Bellevue Adult Intelligence Scale: A history and prospectus. *South African Journal for Psychology*, 27, 43-49.

Nelson, C. A. (2000). Neural plasticity and human development: The role of early experience in sculpting memory systems. *Developmental Science*, *3*(2), 115-136.

Oakland, T., & Weilert, E. (1971). The effects of test-wiseness materials on standardized test performance of preschool disadvantaged children.

Oakland, T. (2005). Selected ethical issues relevant to test adaptions: In Hambleton, R. K., Merenda, P. F., & Spielberger, C. D. (Eds.). *Adapting educational and psychological tests for cross-cultural assessment*. Psychology Press.

Owen, K. (1991). Test bias: The validity of the Junior Aptitude Test for various population groups in South Africa regarding constructs measured. *South African Journal of Psychology*, 21, 112-118.

Owusu-Ansah, F. E., & Mji, G. (2013). African indigenous knowledge and research. African Journal of Disability, 2(1), 5-pages.

Paterson, H., & Uys, K. (2005). Critical issues in psychological test use in the South African workplace*. SA Journal of Industrial Psychology*, 31 (3), 12-22.

Pérez-Escamilla, R., Cavallera, V., Tomlinson, M., & Dua, T. (2018). Scaling up Integrated Early Childhood Development programs: lessons from four countries. *Child: care, health and development*, *44*(1), 50-61.

Piaget, J. (1952). *The origins of intelligence in children* (Vol. 8, No. 5, pp. 18-1952). New York: International Universities Press.Princiotta, D and Goldstein, S. (2015). Intelligence as a Conceptual Construct: The Philosophy of Plato and Pascal. *Handbook of Intelligence: Evolutionary Theory, Historical Perspective and Current Concepts.* (pp. 85-92). Springer New York

Pollard, B., Johnston, M., & Dixon, D. (2007). Theoretical framework and methodological development of common subjective health outcome measures in osteoarthritis: a critical review. *Health and quality of life outcomes*, *5*(1), 14.

Porteus, S. D. (1950). The Porteus Maze Test and intelligence.

Portolese, L., Krause, J., & Bonner, J. (2016). Timed Online Tests: Do Students Perform Better With More Time?. American Journal of Distance Education, 30(4), 264-271.

Princiotta, D and Goldstein, S. (2015). Intelligence as a Conceptual Construct: The Philosophy of Plato and Pascal. Handbook of Intelligence: Evolutionary Theory, Historical Perspective and Current Concepts. (pp. 85-92). Springer New York

Raufelder, D., Regner, N., & Wood, M. A. (2018). Test anxiety and learned helplessness is moderated by student perceptions of teacher motivational support. *Educational Psychology*, *38*(1), 54-74.

Raven, J., Raven, J. C., & Court, J. H. (1998). Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 2: The Coloured Progressive Matrices. San Antonio, TX: Harcourt Assessment

Reams, J. (2014). A brief overview of developmental theory, or what I learned in the FOLA course. Integral Review, 10(1), 122-152.

Reynolds, C. R., & Kamphaus, R. W. (Eds.). (2003). *Handbook of psychological and educational assessment of children: Personality, behavior, and context* (Vol. 2). Guilford Press.

Saklofske, D. H., Van de Vijver, F. J., Oakland, T., Mpofu, E., & Suzuki, L. A. (2015). Intelligence and culture: History and assessment. *Handbook of Intelligence: Evolutionary Theory, Historical Perspective and Current Concepts.* (pp. 341-365). Springer New York.

Samuels, M. L., Taylor, S., Shepherd, D., van der Berg, S., Jacob, C., Deliwe, C. N., & Mabogoane, T. (2015). Reflecting on an impact evaluation of the Grade R programme: Method, results and policy responses. *African Evaluation Journal*, *3*(1), 10.

Santrock, J. (2005). *Psychology 7th edition*. McGraw-Hill

Sattler, J. M., & Hoge, R. D. (2006). Assessment of children: Behavioral, social, and clinical foundations. *Look inside this issue for the*, *123*.

Setati, M., Adler, J., Reed, Y., & Bapoo, A. (2002). Incomplete journeys: Code-switching and other language practices in mathematics, science and English language classrooms in South Africa. *Language and education*, *16*(2), 128-149.

Serpell, R. (2011). Social responsibility as a dimension of intelligence, and as an educational goal: Insights from programmatic research in an African society. Child Development Perspectives, 5(2), 126-133.

Shaffer, D. (2007) *Developmental Psychology: Childhood and Adolescence.* Belmount. Wadsworth*.*

Spelke, E. S. (2003). What makes us smart? Core knowledge and natural language. *Language in mind: Advances in the study of language and thought*, 277-311.

Shanks, D. R., Newell, B. R., Lee, E. H., Balakrishnan, D., Ekelund, L., Cenac, Z., ... & Moore, C. (2013). Priming intelligent behavior: An elusive phenomenon. *PloS one*, *8*(4), e56515.

Shuttleworth-Edwards, A. B., Van der Merwe, A. S., Van Tonder, P., & Radloff, S. E. (2013). WISC-IV test performance in the South African context: A collation of cross-cultural norms. *Psychological assessment in South Africa: Current research and practice*, 33-47.

Shye, S. (1988). Inductive and deductive reasoning: A structural reanalysis of ability tests. *Journal of Applied Psychology*, 73, 308–311.

Sternberg, R. J. (Ed.). (1982). *Handbook of human intelligence*. CUP Archive.

Sternberg, R.J. (2015). Multiple Intelligences in the New Age of Thinking. *Handbook of Intelligence: Evolutionary Theory, Historical Perspective and Current Concepts.* (pp. 105-120). Springer New York

Stein, N. (2017). Language in Schools. In *Basic Education Rights Handbook – Education Rights in South Africa* (pp. 205-217). Johannesburg, South Africa. Section27.

Stoop, C. (2017). Children's rights to mother-tongue education in a multilingual world: a comparative analysis between South Africa and Germany. *Potchefstroom Electronic Law Journal/Potchefstroomse Elektroniese Regsblad*, *20*(1).

Sweet, L.H. (2013). Symbol Search. In Bullard, S. E., Griss, M., Greene, S., & Gekker (Eds.), A Encyclopaedia Of Clinical Neuropsychology. Archives Of Clinical Neuropsychology, (pp. 2444-2445). New York: Springer

Taylor, S., & Yu, D. (2009). The importance of socio-economic status in determining educational achievement in South Africa. *Unpublished working paper (Economics). Stellenbosch: Stellenbosch University.*

The Professional Board for Psychology. (2009). *List of tests classified as being psychological tests.* (No. Form 207). Health Professions Council of South Africa.

Todor, J. I. (1979). Developmental differences in motor task integration: a test of Pascual-Leone's theory of constructive operators. *Journal of Experimental Child Psychology*, *28*(2), 314-322.

Truter, A. (2015). *Education in the best interests of the child: a case study of rural schools in KwaZulu-Natal* (Doctoral dissertation, University of Pretoria).

Van der Berg, S. (2007). Apartheid's enduring legacy: In equlities in education. *Journal of African Economics*, *16(5),* 849-880.

Van de Vijver, A.J.R. & Rothmann, S. (2004). Assessment in multicultural groups: The South African case. *SA Journal of Industrial Psychology*, 30 (4), 1 – 7.

Vrinioti, K., Einarsdottir, J., & Broström, S. (2010). Transitions from preschool to primary school. Early Years Transition Programme.

Wassenaar, D.R.(1998). A history of ethical codes in South African psychology: an insider's view. *South African Journal of Psychology, 28,* 135-145.

Weiβ, R & Osterland, J. (2013) Culture Fair Intelligence Test – Version One – Manual. Göttingen. Hogrefe.

Wechsler, D. (2003a). *Wechsler Preschool and Primary Scale of Intelligence—Thirds Edition*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2003b). Wechsler intelligence scale for children, Fourth Edition (WISC-IV). San Antonio, TX: Psychological Corporation.

Zuriff, G. E. (2000). Extra examination time for students with learning disabilities: An examination of the maximum potential thesis. Applied Measurement in Education, 13(1), 99–117

7. APPENDIX 1A: Example of consent letter sent to parents

XXX[7] Primary School / Laerskool

Dear Parent or Guardian

The Grade 2 learners from XXX Primary School have been chosen to take part in a research study conducted by myself Samantha Coppard from the University of Johannesburg. The purpose of the study is to assess whether we can improve the cultural fairness of an intelligence test by adapting the time limits. This study will contribute to the completion of my master's dissertation.

The study will be done using a translated version of the CFT1-R test. This test, which assess children's fluid intelligence levels, has been adapted, standardised and normed for use in Germany. The test is successfully used in Germany. It has been translated into isiZulu, Sesotho, English and Afrikaans for the purposes of a previous study. My study aims to establish whether or not the time limit given to the German learners is fair for use on South African learners. Learners from three other Johannesburg schools will also be assessed. The assessments will take place at the school, during school hours.

This letter serves to ask your consent for your child to be assessed. The assessments will be supervised by Prof Lara Ragpot from UJ. We will keep your child's identity confidential.

Your child's participation is entirely voluntary. He/she is free to choose not to participate. Should you and your child choose to participate, he/she can withdraw at any time without consequences of any kind.

The results of the assessment will be made available to the school principal.

Due to cost factors, only 30 children can be assessed per school, so if you are interested, please sign the below consent form and return to the school no later than the 11 March 2016.

If you have any questions, please feel free contact the school.

Kind Regards,

Samantha Coppard

[7] For confidentiality purposes the name of the participating schools was changed to xxx

8. APPENDIX 1B: Ethics clearance

NHREC Registration Number REC-110613-036

UNIVERSITY OF JOHANNESBURG

ETHICS CLEARANCE

Dear Samantha Coppard

Ethical Clearance Number: 2018-062

Assessing the culture fairness of an intelligence test by adjusting the test times and pictorial examples: a pilot study with Grade 2 learners in four Johannesburg schools

Ethical clearance for this study is granted subject to the following conditions:

- If there are major revisions to the research proposal based on recommendations from the Faculty Higher Degrees Committee, a new application for ethical clearance must be submitted.
- If the research question changes significantly so as to alter the nature of the study, it remains the duty of the student to submit a new application.
- It remains the student's responsibility to ensure that all ethical forms and documents related to the research are kept in a safe and secure facility and are available on demand.
- Please quote the reference number above in all future communications and documents.

The Faculty of Education Research Ethics Committee has decided to

☒ Grant ethical clearance for the proposed research.
☐ Provisionally grant ethical clearance for the proposed research
☐ Recommend revision and resubmission of the ethical clearance documents

Sincerely,

Dr David Robinson
Chair: FACULTY OF EDUCATION RESEARCH ETHICS COMMITTEE
21 August 2018

9. APPENDIX 2: Original SPSS data outputs

9.1 Description of the sample

Statistiken[a]

		Home_Language	Gender	Race	Age
N	Gültig	36	36	36	34
	Fehlend	0	0	0	2
Mittelwert		1,0000	1,5000	1,0000	7,0000
Std.-Abweichung		,00000	,50709	,00000	,85280
Varianz		,000	,257	,000	,727
Minimum		1,00	1,00	1,00	6,00
Maximum		1,00	2,00	1,00	9,00

a. Home_Language = Sesotho

Statistiken[a]

		Home_Language	Gender	Race	Age
N	Gültig	32	32	32	31
	Fehlend	0	0	0	1
Mittelwert		3,0000	1,5000	1,9063	7,3548
Std.-Abweichung		,00000	,50800	,39015	,55066
Varianz		,000	,258	,152	,303
Minimum		3,00	1,00	1,00	6,00
Maximum		3,00	2,00	3,00	8,00

a. Home_Language = Africaans

Statistiken[a]

		Home_Language	Gender	Race	Age
N	Gültig	37	37	37	33
	Fehlend	0	0	0	4
Mittelwert		2,0000	1,3243	1,0811	7,1212
Std.-Abweichung		,00000	,47458	,36350	,59987
Varianz		,000	,225	,132	,360
Minimum		2,00	1,00	1,00	6,00
Maximum		2,00	2,00	3,00	8,00

a. Home_Language = Isi-Zulu

Statistiken[a]

		Home_Language	Gender	Race	Age
N	Gültig	4	4	4	3
	Fehlend	0	0	0	1
Mittelwert		4,0000	1,2500	2,2500	7,0000
Std.-Abweichung		,00000	,50000	1,50000	,00000
Varianz		,000	,250	2,250	,000
Minimum		4,00	1,00	1,00	7,00
Maximum		4,00	2,00	4,00	7,00

a. Home_Language = English

9.2 Description of home language groups too small for statistical analysis

Statistiken[a]

		Home_Language	Gender	Race	Age
N	Gültig	2	2	2	2
	Fehlend	0	0	0	0
Mittelwert		5,0000	2,0000	1,0000	7,0000
Std.-Abweichung		,00000	,00000	,00000	,00000
Varianz		,000	,000	,000	,000
Minimum		5,00	2,00	1,00	7,00
Maximum		5,00	2,00	1,00	7,00

a. Home_Language = Venda

Statistiken[a]

		Home_Language	Gender	Race	Age
N	Gültig	2	2	2	2
	Fehlend	0	0	0	0
Mittelwert		7,0000	1,5000	1,0000	7,5000
Std.-Abweichung		,00000	,70711	,00000	,70711
Varianz		,000	,500	,000	,500
Minimum		7,00	1,00	1,00	7,00
Maximum		7,00	2,00	1,00	8,00

a. Home_Language = Sepedi

Statistiken[a]

		Home_Language	Gender	Race	Age
N	Gültig	6	6	6	5
	Fehlend	0	0	0	1
Mittelwert		6,0000	1,6667	1,0000	7,4000
Std.-Abweichung		,00000	,51640	,00000	,54772
Varianz		,000	,267	,000	,300
Minimum		6,00	1,00	1,00	7,00
Maximum		6,00	2,00	1,00	8,00

a. Home_Language = Tswana

Statistiken[a]

		Home_Language	Gender	Race	Age
N	Gültig	1	1	1	1
	Fehlend	0	0	0	0
Mittelwert		8,0000	2,0000	1,0000	8,0000
Minimum		8,00	2,00	1,00	8,00
Maximum		8,00	2,00	1,00	8,00

a. Home_Language = Xhosa

9.3 Description of performance per subtest of entire sample

Statistiken

		Subtest_1_CT	Subtest_2_sum score	Subtest_3_sum score	Subtest_4_sum score	Subtest_5_sum score	Subtest_6_sum score
N	Gültig	120	120	120	120	120	120
	Fehlend	0	0	0	0	0	0
Mittelwert		74,0583	13,5833	9,9917	5,6417	8,7333	7,3917
Std.-Abweichung		3,57229	2,29572	2,35324	3,11083	3,21167	3,74681
Varianz		12,761	5,270	5,538	9,677	10,315	14,039
Minimum		55,00	5,00	,00	1,00	1,00	,00
Maximum		75,00	15,00	14,00	13,00	14,00	14,00

9.4 Development of performances depending on test time: Subtest One

Statistik bei gepaarten Stichproben

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_1_T_70 correct answers per timeslot	33,0333	120	11,57071	1,05626
	S_1_T_100 correct answers per timeslot	45,5417	120	13,05959	1,19217
Paaren 2	S_1_T_100 correct answers per timeslot	42,5385	104	10,89038	1,06789
	S_1_T_130 correct answers per timeslot	53,2404	104	12,33001	1,20906
Paaren 3	S_1_T_130 correct answers per timeslot	49,0769	78	10,37951	1,17525
	S_1_T_160 correct answers per timeslot	59,1667	78	10,53802	1,19320
Paaren 4	S_1_T_160 correct answers per timeslot	53,5417	48	9,04588	1,30566
	S_1_T_190 correct answers per timeslot	63,2083	48	10,10995	1,45924
Paaren 5	S_1_T_190 correct answers per timeslot	55,1364	22	9,34164	1,99164
	S_1_T_220 correct answers per timeslot	63,6818	22	9,11507	1,94334
Paaren 6	S_1_T_220 correct answers per timeslot	59,2143	14	8,12573	2,17169
	S_1_T_250 correct answers per timeslot	65,8571	14	8,70771	2,32723
Paaren 7	S_1_T_250 correct answers per timeslot	60,0000	7	8,77496	3,31662
	S_1_T_280 correct answers per timeslot	67,0000	7	7,32575	2,76887
Paaren 8	S_1_T_280 correct answers per timeslot	63,2500	4	7,67572	3,83786
	S_1_T_290 correct answers per timeslot	68,0000	4	7,16473	3,58236

Test bei gepaarten Stichproben

		Gepaarte Differenzen							
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere	T	df	Sig. (2-seitig)
Paaren 1	S_1_T_70 correct answers per timeslot - S_1_T_100 correct answers per timeslot	-12,50833	4,55959	,41623	-13,33251	-11,68415	-30,051	119	,000
Paaren 2	S_1_T_100 correct answers per timeslot - S_1_T_130 correct answers per timeslot	-10,70192	4,85494	,47607	-11,64609	-9,75776	-22,480	103	,000
Paaren 3	S_1_T_130 correct answers per timeslot - S_1_T_160 correct answers per timeslot	-10,08974	3,11755	,35299	-10,79264	-9,38684	-28,583	77	,000
Paaren 4	S_1_T_160 correct answers per timeslot - S_1_T_190 correct answers per timeslot	-9,66667	4,14284	,59797	-10,86962	-8,46371	-16,166	47	,000
Paaren 5	S_1_T_190 correct answers per timeslot - S_1_T_220 correct answers per timeslot	-8,54545	2,85736	,60919	-9,81234	-7,27857	-14,028	21	,000
Paaren 6	S_1_T_220 correct answers per timeslot - S_1_T_250 correct answers per timeslot	-6,64286	3,07864	,82280	-8,42041	-4,86530	-8,073	13	,000
Paaren 7	S_1_T_250 correct answers per timeslot - S_1_T_280 correct answers per timeslot	-7,00000	4,43471	1,67616	-11,10142	-2,89858	-4,176	6	,006
Paaren 8	S_1_T_280 correct answers per timeslot - S_1_T_290 correct answers per timeslot	-4,75000	2,50000	1,25000	-8,72806	-,77194	-3,800	3	,032

9.5 Development of performances depending on test time: Subtest Two

Statistik bei gepaarten Stichproben

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_2_T_90 correct answers per timeslot	7,3361	119	2,13636	,19584
	S_2_T_120 correct answers per timeslot	8,9328	119	2,21601	,20314
Paaren 2	S_2_T_120 correct answers per timeslot	8,6916	107	2,04852	,19804
	S_2_T_150 correct answers per timeslot	10,1589	107	2,35183	,22736
Paaren 3	S_2_T_150 correct answers per timeslot	9,8557	97	2,24532	,22798
	S_2_T_180 correct answers per timeslot	11,1340	97	2,37447	,24109
Paaren 4	S_2_T_180 correct answers per timeslot	10,5658	76	2,16232	,24803
	S_2_T_210 correct answers per timeslot	11,6579	76	2,46875	,28319
Paaren 5	S_2_T_210 correct answers per timeslot	11,0500	60	2,33198	,30106
	S_2_T_240 correct answers per timeslot	11,7833	60	2,49808	,32250
Paaren 6	S_2_T_240 correct answers per timeslot	10,9762	42	2,56133	,39522
	S_2_T_270 correct answers per timeslot	11,7619	42	2,54529	,39275
Paaren 7	S_2_T_270 correct answers per timeslot	11,4063	32	2,39434	,42326
	S_2_T_300 correct answers per timeslot	12,3125	32	2,29217	,40520

Test bei gepaarten Stichproben

		Gepaarte Differenzen					T	df	Sig. (2-seitig)
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere			
Paaren 1	S_2_T_90 correct answers per timeslot - S_2_T_120 correct answers per timeslot	-1,59664	,93264	,08550	-1,76594	-1,42734	-18,675	118	,000
Paaren 2	S_2_T_120 correct answers per timeslot - S_2_T_150 correct answers per timeslot	-1,46729	,79287	,07665	-1,61925	-1,31533	-19,143	106	,000
Paaren 3	S_2_T_150 correct answers per timeslot - S_2_T_180 correct answers per timeslot	-1,27835	,70330	,07141	-1,42010	-1,13660	-17,902	96	,000
Paaren 4	S_2_T_180 correct answers per timeslot - S_2_T_210 correct answers per timeslot	-1,09211	,76903	,08821	-1,26784	-,91637	-12,380	75	,000
Paaren 5	S_2_T_210 correct answers per timeslot - S_2_T_240 correct answers per timeslot	-,73333	,68561	,08851	-,91044	-,55622	-8,285	59	,000
Paaren 6	S_2_T_240 correct answers per timeslot - S_2_T_270 correct answers per timeslot	-,78571	,68202	,10524	-,99825	-,57318	-7,466	41	,000
Paaren 7	S_2_T_270 correct answers per timeslot - S_2_T_300 correct answers per timeslot	-,90625	,81752	,14452	-1,20100	-,61150	-6,271	31	,000

9.6 Development of performances depending on test time: Subtest Three

Statistik bei gepaarten Stichproben

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_3_T_90 correct answers per timeslot	6,0442	113	1,63330	,15365
	S_3_T_120 correct answers per timeslot	7,2478	113	1,88761	,17757
Paaren 2	S_3_T_120 correct answers per timeslot	6,9432	88	1,83427	,19553
	S_3_T_150 correct answers per timeslot	8,1136	88	2,08680	,22245
Paaren 3	S_3_T_150 correct answers per timeslot	7,7692	65	2,06737	,25643
	S_3_T_180 correct answers per timeslot	8,6615	65	2,43847	,30246
Paaren 4	S_3_T_180 correct answers per timeslot	7,8611	36	2,26971	,37829
	S_3_T_210 correct answers per timeslot	8,8333	36	2,39643	,39940
Paaren 5	S_3_T_210 correct answers per timeslot	8,2105	19	2,83978	,65149
	S_3_T_240 correct answers per timeslot	9,0526	19	3,04546	,69868
Paaren 6	S_3_T_240 correct answers per timeslot	6,8571	7	3,28778	1,24267
	S_3_T_270 correct answers per timeslot	7,2857	7	3,63842	1,37519
Paaren 7	S_3_T_270 correct answers per timeslot	6,2000	5	3,76829	1,68523
	S_3_T_300 correct answers per timeslot	7,4000	5	4,03733	1,80555

Test bei gepaarten Stichproben

		Gepaarte Differenzen					T	df	Sig. (2-seitig)
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz Untere	Obere			
Paaren 1	S_3_T_90 correct answers per timeslot - S_3_T_120 correct answers per timeslot	-1,20354	,85731	,08065	-1,36334	-1,04374	-14,923	112	,000
Paaren 2	S_3_T_120 correct answers per timeslot - S_3_T_150 correct answers per timeslot	-1,17045	,83352	,08885	-1,34706	-,99385	-13,173	87	,000
Paaren 3	S_3_T_150 correct answers per timeslot - S_3_T_180 correct answers per timeslot	-,89231	,83147	,10313	-1,09834	-,68628	-8,652	64	,000
Paaren 4	S_3_T_180 correct answers per timeslot - S_3_T_210 correct answers per timeslot	-,97222	,81015	,13503	-1,24634	-,69811	-7,200	35	,000
Paaren 5	S_3_T_210 correct answers per timeslot - S_3_T_240 correct answers per timeslot	-,84211	,83421	,19138	-1,24418	-,44003	-4,400	18	,000
Paaren 6	S_3_T_240 correct answers per timeslot - S_3_T_270 correct answers per timeslot	-,42857	,53452	,20203	-,92292	,06578	-2,121	6	,078
Paaren 7	S_3_T_270 correct answers per timeslot - S_3_T_300 correct answers per timeslot	-1,20000	,83666	,37417	-2,23885	-,16115	-3,207	4	,033

9.7 Development of performances depending on test time: Subtest Four

Statistik bei gepaarten Stichproben

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_4_T_180 correct answers per timeslot	4,0256	78	2,50700	,28386
	S_4_T_210 correct answers per timeslot	4,6923	78	2,86178	,32403
Paaren 2	S_4_T_210 correct answers per timeslot	4,1525	59	2,70258	,35185
	S_4_T_240 correct answers per timeslot	4,5593	59	2,88439	,37552
Paaren 3	S_4_T_240 correct answers per timeslot	4,3000	40	2,46202	,38928
	S_4_T_270 correct answers per timeslot	4,6250	40	2,68603	,42470
Paaren 4	S_4_T_270 correct answers per timeslot	4,2963	27	2,44658	,47084
	S_4_T_300 correct answers per timeslot	4,6667	27	2,74563	,52840

Test bei gepaarten Stichproben

		Gepaarte Differenzen							
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere	T	df	Sig. (2-seitig)
Paaren 1	S_4_T_180 correct answers per timeslot - S_4_T_210 correct answers per timeslot	-,66667	,81650	,09245	-,85076	-,48258	-7,211	77	,000
Paaren 2	S_4_T_210 correct answers per timeslot - S_4_T_240 correct answers per timeslot	-,40678	,59069	,07690	-,56072	-,25284	-5,290	58	,000
Paaren 3	S_4_T_240 correct answers per timeslot - S_4_T_270 correct answers per timeslot	-,32500	,47434	,07500	-,47670	-,17330	-4,333	39	,000
Paaren 4	S_4_T_270 correct answers per timeslot - S_4_T_300 correct answers per timeslot	-,37037	,68770	,13235	-,64242	-,09832	-2,798	26	,010

9.8 Development of performances depending on test time: Subtest Five

Statistik bei gepaarten Stichproben

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_5_T_180 correct answers per timeslot	6,4878	41	3,20251	,50015
	S_5_T_210 correct answers per timeslot	7,0488	41	3,30115	,51555
Paaren 2	S_5_T_210 correct answers per timeslot	6,8636	22	3,35652	,71561
	S_5_T_240 correct answers per timeslot	7,0909	22	3,35104	,71444
Paaren 3	S_5_T_240 correct answers per timeslot	6,5833	12	3,23218	,93305
	S_5_T_270 correct answers per timeslot	7,0000	12	3,33030	,96138
Paaren 4	S_5_T_270 correct answers per timeslot	5,7500	8	2,81577	,99553
	S_5_T_300 correct answers per timeslot	6,0000	8	2,61861	,92582

Test bei gepaarten Stichproben

		Gepaarte Differenzen							
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere	T	df	Sig. (2-seitig)
Paaren 1	S_5_T_180 correct answers per timeslot - S_5_T_210 correct answers per timeslot	-,56098	,80774	,12615	-,81593	-,30602	-4,447	40	,000
Paaren 2	S_5_T_210 correct answers per timeslot - S_5_T_240 correct answers per timeslot	-,22727	,42893	,09145	-,41745	-,03709	-2,485	21	,021
Paaren 3	S_5_T_240 correct answers per timeslot - S_5_T_270 correct answers per timeslot	-,41667	,90034	,25990	-,98871	,15538	-1,603	11	,137
Paaren 4	S_5_T_270 correct answers per timeslot - S_5_T_300 correct answers per timeslot	-,25000	,46291	,16366	-,63700	,13700	-1,528	7	,170

9.9 Development of performances depending on test time: Subtest Six

Statistik bei gepaarten Stichproben

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_6_T_180 correct answers per timeslot	5,8889	18	2,98799	,70428
	S_6_T_210 correct answers per timeslot	7,1667	18	3,29438	,77649
Paaren 2	S_6_T_210 correct answers per timeslot	8,1250	8	3,27054	1,15631
	S_6_T_240 correct answers per timeslot	8,6250	8	3,42000	1,20915
Paaren 3	S_6_T_240 correct answers per timeslot	9,0000	5	2,44949	1,09545
	S_6_T_270 correct answers per timeslot	9,4000	5	2,40832	1,07703
Paaren 4	S_6_T_270 correct answers per timeslot	.	0[a]	.	.
	S_6_T_300 correct answers per timeslot	.	0[a]	.	.

a. Korrelation und T können nicht berechnet werden, da keine gültigen Paare vorhanden sind.

Test bei gepaarten Stichproben

		Gepaarte Differenzen							
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere	T	df	Sig. (2-seitig)
Paaren 1	S_6_T_180 correct answers per timeslot - S_6_T_210 correct answers per timeslot	-1,27778	1,07406	,25316	-1,81189	-,74366	-5,047	17	,000
Paaren 2	S_6_T_210 correct answers per timeslot - S_6_T_240 correct answers per timeslot	-,50000	,53452	,18898	-,94687	-,05313	-2,646	7	,033
Paaren 3	S_6_T_240 correct answers per timeslot - S_6_T_270 correct answers per timeslot	-,40000	,54772	,24495	-1,08009	,28009	-1,633	4	,178

9.10 Development of performances depending on test time and home language: Subtest One

Statistik bei gepaarten Stichproben[a]

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_1_T_70 correct answers per timeslot	35,6944	36	12,51472	2,08579
	S_1_T_100 correct answers per timeslot	47,7778	36	14,04709	2,34118
Paaren 2	S_1_T_100 correct answers per timeslot	43,6897	29	11,80830	2,19275
	S_1_T_130 correct answers per timeslot	53,8621	29	13,57709	2,52120
Paaren 3	S_1_T_130 correct answers per timeslot	50,6364	22	11,30758	2,41079
	S_1_T_160 correct answers per timeslot	60,8182	22	11,47103	2,44563
Paaren 4	S_1_T_160 correct answers per timeslot	52,0909	11	9,88387	2,98010
	S_1_T_190 correct answers per timeslot	61,3636	11	11,21849	3,38250
Paaren 5	S_1_T_190 correct answers per timeslot	50,4000	5	5,54977	2,48193
	S_1_T_220 correct answers per timeslot	60,8000	5	5,40370	2,41661
Paaren 6	S_1_T_220 correct answers per timeslot	58,7500	4	3,30404	1,65202
	S_1_T_250 correct answers per timeslot	67,5000	4	4,79583	2,39792
Paaren 7	S_1_T_250 correct answers per timeslot	65,5000	2	6,36396	4,50000
	S_1_T_280 correct answers per timeslot	73,5000	2	,70711	,50000
Paaren 8	S_1_T_280 correct answers per timeslot	73,0000	1[b]	.	.
	S_1_T_290 correct answers per timeslot	75,0000	1[b]	.	.

a. Home_Language = Sesotho

b. Korrelation und T können nicht berechnet werden, da die Summe der Fallgewichtungen kleiner oder gleich 1 ist.

Test bei gepaarten Stichproben[a]

		Gepaarte Differenzen							
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere	T	df	Sig. (2-seitig)
Paaren 1	S_1_T_70 correct answers per timeslot - S_1_T_100 correct answers per timeslot	-12,08333	3,84986	,64164	-13,38594	-10,78073	-18,832	35	,000
Paaren 2	S_1_T_100 correct answers per timeslot - S_1_T_130 correct answers per timeslot	-10,17241	7,27063	1,35012	-12,93801	-7,40681	-7,534	28	,000
Paaren 3	S_1_T_130 correct answers per timeslot - S_1_T_160 correct answers per timeslot	-10,18182	2,57527	,54905	-11,32363	-9,04001	-18,544	21	,000
Paaren 4	S_1_T_160 correct answers per timeslot - S_1_T_190 correct answers per timeslot	-9,27273	2,32770	,70183	-10,83650	-7,70896	-13,212	10	,000
Paaren 5	S_1_T_190 correct answers per timeslot - S_1_T_220 correct answers per timeslot	-10,40000	1,14018	,50990	-11,81571	-8,98429	-20,396	4	,000
Paaren 6	S_1_T_220 correct answers per timeslot - S_1_T_250 correct answers per timeslot	-8,75000	2,06155	1,03078	-12,03039	-5,46961	-8,489	3	,003
Paaren 7	S_1_T_250 correct answers per timeslot - S_1_T_280 correct answers per timeslot	-8,00000	7,07107	5,00000	-71,53102	55,53102	-1,600	1	,356

a. Home_Language = Sesotho

Statistik bei gepaarten Stichproben[a]

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_1_T_70 correct answers per timeslot	30,6216	37	10,87901	1,78850
	S_1_T_100 correct answers per timeslot	42,5946	37	13,03299	2,14261
Paaren 2	S_1_T_100 correct answers per timeslot	40,4545	33	11,69159	2,03524
	S_1_T_130 correct answers per timeslot	50,9697	33	12,96988	2,25777
Paaren 3	S_1_T_130 correct answers per timeslot	46,8462	26	11,24168	2,20467
	S_1_T_160 correct answers per timeslot	56,1154	26	10,89707	2,13709
Paaren 4	S_1_T_160 correct answers per timeslot	52,6842	19	9,82091	2,25307
	S_1_T_190 correct answers per timeslot	61,8947	19	11,41098	2,61786
Paaren 5	S_1_T_190 correct answers per timeslot	55,4000	10	12,03883	3,80701
	S_1_T_220 correct answers per timeslot	63,0000	10	12,22020	3,86437
Paaren 6	S_1_T_220 correct answers per timeslot	55,8333	6	10,66615	4,35444
	S_1_T_250 correct answers per timeslot	60,6667	6	10,80123	4,40959
Paaren 7	S_1_T_250 correct answers per timeslot	57,8000	5	9,17606	4,10366
	S_1_T_280 correct answers per timeslot	64,4000	5	7,12741	3,18748
Paaren 8	S_1_T_280 correct answers per timeslot	60,0000	3	5,00000	2,88675
	S_1_T_290 correct answers per timeslot	65,6667	3	6,65833	3,84419

a. Home_Language = Isi-Zulu

Test bei gepaarten Stichproben[a]

		Gepaarte Differenzen					T	df	Sig. (2-seitig)
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere			
Paaren 1	S_1_T_70 correct answers per timeslot - S_1_T_100 correct answers per timeslot	-11,97297	4,57946	,75286	-13,49984	-10,44610	-15,903	36	,000
Paaren 2	S_1_T_100 correct answers per timeslot - S_1_T_130 correct answers per timeslot	-10,51515	3,77592	,65730	-11,85403	-9,17627	-15,997	32	,000
Paaren 3	S_1_T_130 correct answers per timeslot - S_1_T_160 correct answers per timeslot	-9,26923	3,44741	,67609	-10,66167	-7,87679	-13,710	25	,000
Paaren 4	S_1_T_160 correct answers per timeslot - S_1_T_190 correct answers per timeslot	-9,21053	4,72086	1,08304	-11,48591	-6,93514	-8,504	18	,000
Paaren 5	S_1_T_190 correct answers per timeslot - S_1_T_220 correct answers per timeslot	-7,60000	3,30656	1,04563	-9,96537	-5,23463	-7,268	9	,000
Paaren 6	S_1_T_220 correct answers per timeslot - S_1_T_250 correct answers per timeslot	-4,83333	1,47196	,60093	-6,37806	-3,28861	-8,043	5	,000
Paaren 7	S_1_T_250 correct answers per timeslot - S_1_T_280 correct answers per timeslot	-6,60000	4,03733	1,80555	-11,61300	-1,58700	-3,655	4	,022
Paaren 8	S_1_T_280 correct answers per timeslot - S_1_T_290 correct answers per timeslot	-5,66667	2,08167	1,20185	-10,83781	-,49552	-4,715	2	,042

a. Home_Language = Isi-Zulu

Statistik bei gepaarten Stichproben[a]

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_1_T_70 correct answers per timeslot	35,7813	32	11,04459	1,95243
	S_1_T_100 correct answers per timeslot	49,4375	32	11,66450	2,06201
Paaren 2	S_1_T_100 correct answers per timeslot	45,7407	27	8,43798	1,62389
	S_1_T_130 correct answers per timeslot	56,9630	27	9,44409	1,81752
Paaren 3	S_1_T_130 correct answers per timeslot	52,6667	18	8,18895	1,93015
	S_1_T_160 correct answers per timeslot	63,5556	18	8,25493	1,94570
Paaren 4	S_1_T_160 correct answers per timeslot	56,4286	7	8,46280	3,19864
	S_1_T_190 correct answers per timeslot	67,5714	7	8,26352	3,12332
Paaren 5	S_1_T_190 correct answers per timeslot	50,0000	1[b]	.	.
	S_1_T_220 correct answers per timeslot	59,0000	1[b]	.	.
Paaren 6	S_1_T_220 correct answers per timeslot	59,0000	1[b]	.	.
	S_1_T_250 correct answers per timeslot	70,0000	1[b]	.	.
Paaren 7	S_1_T_250 correct answers per timeslot	.	0[c]	.	.
	S_1_T_280 correct answers per timeslot	.	0[c]	.	.
Paaren 8	S_1_T_280 correct answers per timeslot	.	0[c]	.	.
	S_1_T_290 correct answers per timeslot	.	0[c]	.	.

a. Home_Language = Africaans

b. Korrelation und T können nicht berechnet werden, da die Summe der Fallgewichtungen kleiner oder gleich 1 ist.

c. Korrelation und T können nicht berechnet werden, da keine gültigen Paare vorhanden sind.

Test bei gepaarten Stichproben[a]

		Gepaarte Differenzen							
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere	T	df	Sig. (2-seitig)
Paaren 1	S_1_T_70 correct answers per timeslot - S_1_T_100 correct answers per timeslot	-13,65625	5,25240	,92850	-15,54994	-11,76256	-14,708	31	,000
Paaren 2	S_1_T_100 correct answers per timeslot - S_1_T_130 correct answers per timeslot	-11,22222	3,28556	,63231	-12,52194	-9,92250	-17,748	26	,000
Paaren 3	S_1_T_130 correct answers per timeslot - S_1_T_160 correct answers per timeslot	-10,88889	3,49603	,82402	-12,62742	-9,15035	-13,214	17	,000
Paaren 4	S_1_T_160 correct answers per timeslot - S_1_T_190 correct answers per timeslot	-11,14286	5,49025	2,07512	-16,22049	-6,06522	-5,370	6	,002

a. Home_Language = Africaans

9.11 Development of performances depending on test time and home language: Subtest Two

Statistik bei gepaarten Stichproben[a]

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_2_T_90 correct answers per timeslot	6,4167	36	2,44219	,40703
	S_2_T_120 correct answers per timeslot	7,8889	36	2,53859	,42310
Paaren 2	S_2_T_120 correct answers per timeslot	7,8235	34	2,40543	,41253
	S_2_T_150 correct answers per timeslot	9,3824	34	2,64019	,45279
Paaren 3	S_2_T_150 correct answers per timeslot	9,2424	33	2,54988	,44388
	S_2_T_180 correct answers per timeslot	10,4545	33	2,85144	,49637
Paaren 4	S_2_T_180 correct answers per timeslot	9,6538	26	2,66747	,52313
	S_2_T_210 correct answers per timeslot	10,7308	26	3,02731	,59370
Paaren 5	S_2_T_210 correct answers per timeslot	10,0455	22	2,76848	,59024
	S_2_T_240 correct answers per timeslot	10,7727	22	3,05399	,65111
Paaren 6	S_2_T_240 correct answers per timeslot	10,2778	18	3,15866	,74450
	S_2_T_270 correct answers per timeslot	10,8333	18	2,97539	,70131
Paaren 7	S_2_T_270 correct answers per timeslot	10,6471	17	2,95680	,71713
	S_2_T_300 correct answers per timeslot	11,5882	17	2,73995	,66454

a. Home_Language = Sesotho

Test bei gepaarten Stichproben[a]

		Gepaarte Differenzen							
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere	T	df	Sig. (2-seitig)
Paaren 1	S_2_T_90 correct answers per timeslot - S_2_T_120 correct answers per timeslot	-1,47222	,97060	,16177	-1,80063	-1,14382	-9,101	35	,000
Paaren 2	S_2_T_120 correct answers per timeslot - S_2_T_150 correct answers per timeslot	-1,55882	,70458	,12083	-1,80466	-1,31298	-12,900	33	,000
Paaren 3	S_2_T_150 correct answers per timeslot - S_2_T_180 correct answers per timeslot	-1,21212	,69631	,12121	-1,45902	-,96522	-10,000	32	,000
Paaren 4	S_2_T_180 correct answers per timeslot - S_2_T_210 correct answers per timeslot	-1,07692	,89098	,17474	-1,43680	-,71705	-6,163	25	,000
Paaren 5	S_2_T_210 correct answers per timeslot - S_2_T_240 correct answers per timeslot	-,72727	,70250	,14977	-1,03874	-,41580	-4,856	21	,000
Paaren 6	S_2_T_240 correct answers per timeslot - S_2_T_270 correct answers per timeslot	-,55556	,51131	,12052	-,80982	-,30129	-4,610	17	,000
Paaren 7	S_2_T_270 correct answers per timeslot - S_2_T_300 correct answers per timeslot	-,94118	,96635	,23437	-1,43803	-,44433	-4,016	16	,001

a. Home_Language = Sesotho

Statistik bei gepaarten Stichproben[a]

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_2_T_90 correct answers per timeslot	7,1351	37	1,65264	,27169
	S_2_T_120 correct answers per timeslot	8,9189	37	1,55239	,25521
Paaren 2	S_2_T_120 correct answers per timeslot	8,8333	36	1,48324	,24721
	S_2_T_150 correct answers per timeslot	10,0556	36	1,75572	,29262
Paaren 3	S_2_T_150 correct answers per timeslot	9,9714	35	1,70614	,28839
	S_2_T_180 correct answers per timeslot	11,2571	35	1,46213	,24714
Paaren 4	S_2_T_180 correct answers per timeslot	11,1765	34	1,40282	,24058
	S_2_T_210 correct answers per timeslot	12,2647	34	1,72870	,29647
Paaren 5	S_2_T_210 correct answers per timeslot	11,7407	27	1,53404	,29523
	S_2_T_240 correct answers per timeslot	12,5185	27	1,69548	,32629
Paaren 6	S_2_T_240 correct answers per timeslot	11,8947	19	1,62941	,37381
	S_2_T_270 correct answers per timeslot	12,9474	19	1,39338	,31966
Paaren 7	S_2_T_270 correct answers per timeslot	12,2500	12	1,21543	,35086
	S_2_T_300 correct answers per timeslot	13,0833	12	1,31137	,37856

a. Home_Language = Isi-Zulu

Test bei gepaarten Stichproben[a]

		Gepaarte Differenzen							
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere	T	df	Sig. (2-seitig)
Paaren 1	S_2_T_90 correct answers per timeslot - S_2_T_120 correct answers per timeslot	-1,78378	,82108	,13498	-2,05755	-1,51002	-13,215	36	,000
Paaren 2	S_2_T_120 correct answers per timeslot - S_2_T_150 correct answers per timeslot	-1,22222	,76012	,12669	-1,47941	-,96504	-9,648	35	,000
Paaren 3	S_2_T_150 correct answers per timeslot - S_2_T_180 correct answers per timeslot	-1,28571	,71007	,12002	-1,52963	-1,04180	-10,712	34	,000
Paaren 4	S_2_T_180 correct answers per timeslot - S_2_T_210 correct answers per timeslot	-1,08824	,75348	,12922	-1,35114	-,82533	-8,421	33	,000
Paaren 5	S_2_T_210 correct answers per timeslot - S_2_T_240 correct answers per timeslot	-,77778	,69798	,13433	-1,05389	-,50167	-5,790	26	,000
Paaren 6	S_2_T_240 correct answers per timeslot - S_2_T_270 correct answers per timeslot	-1,05263	,77986	,17891	-1,42851	-,67675	-5,883	18	,000
Paaren 7	S_2_T_270 correct answers per timeslot - S_2_T_300 correct answers per timeslot	-,83333	,57735	,16667	-1,20016	-,46650	-5,000	11	,000

a. Home_Language = Isi-Zulu

Statistik bei gepaarten Stichproben[a]

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_2_T_90 correct answers per timeslot	8,4063	32	1,81142	,32022
	S_2_T_120 correct answers per timeslot	10,1250	32	1,99596	,35284
Paaren 2	S_2_T_120 correct answers per timeslot	9,6800	25	1,79629	,35926
	S_2_T_150 correct answers per timeslot	11,2800	25	2,11187	,42237
Paaren 3	S_2_T_150 correct answers per timeslot	10,8000	20	2,06729	,46226
	S_2_T_180 correct answers per timeslot	12,2500	20	2,26820	,50719
Paaren 4	S_2_T_180 correct answers per timeslot	11,0833	12	2,19331	,63315
	S_2_T_210 correct answers per timeslot	12,2500	12	2,37888	,68672
Paaren 5	S_2_T_210 correct answers per timeslot	11,7778	9	2,58736	,86245
	S_2_T_240 correct answers per timeslot	12,3333	9	2,59808	,86603
Paaren 6	S_2_T_240 correct answers per timeslot	9,6667	3	3,21455	1,85592
	S_2_T_270 correct answers per timeslot	10,3333	3	3,78594	2,18581
Paaren 7	S_2_T_270 correct answers per timeslot	12,5000	2	,70711	,50000
	S_2_T_300 correct answers per timeslot	13,5000	2	2,12132	1,50000

a. Home_Language = Africaans

Test bei gepaarten Stichproben[a]

		Gepaarte Differenzen							
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere	T	df	Sig. (2-seitig)
Paaren 1	S_2_T_90 correct answers per timeslot - S_2_T_120 correct answers per timeslot	-1,71875	1,05446	,18640	-2,09892	-1,33858	-9,221	31	,000
Paaren 2	S_2_T_120 correct answers per timeslot - S_2_T_150 correct answers per timeslot	-1,60000	,76376	,15275	-1,91527	-1,28473	-10,474	24	,000
Paaren 3	S_2_T_150 correct answers per timeslot - S_2_T_180 correct answers per timeslot	-1,45000	,60481	,13524	-1,73306	-1,16694	-10,722	19	,000
Paaren 4	S_2_T_180 correct answers per timeslot - S_2_T_210 correct answers per timeslot	-1,16667	,57735	,16667	-1,53350	-,79984	-7,000	11	,000
Paaren 5	S_2_T_210 correct answers per timeslot - S_2_T_240 correct answers per timeslot	-,55556	,72648	,24216	-1,11398	,00287	-2,294	8	,051
Paaren 6	S_2_T_240 correct answers per timeslot - S_2_T_270 correct answers per timeslot	-,66667	,57735	,33333	-2,10088	,76755	-2,000	2	,184
Paaren 7	S_2_T_270 correct answers per timeslot - S_2_T_300 correct answers per timeslot	-1,00000	1,41421	1,00000	-13,70620	11,70620	-1,000	1	,500

a. Home_Language = Africaans

9.12 Development of performances depending on test time and home language: Subtest Three

Statistik bei gepaarten Stichproben[a]

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_3_T_90 correct answers per timeslot	5,9118	34	1,28788	,22087
	S_3_T_120 correct answers per timeslot	6,8824	34	1,53277	,26287
Paaren 2	S_3_T_120 correct answers per timeslot	6,8621	29	1,52887	,28390
	S_3_T_150 correct answers per timeslot	8,0690	29	1,83091	,33999
Paaren 3	S_3_T_150 correct answers per timeslot	7,6667	21	1,49443	,32611
	S_3_T_180 correct answers per timeslot	8,6667	21	2,28765	,49921
Paaren 4	S_3_T_180 correct answers per timeslot	8,0000	13	1,95789	,54302
	S_3_T_210 correct answers per timeslot	8,8462	13	1,72463	,47833
Paaren 5	S_3_T_210 correct answers per timeslot	8,8889	9	1,90029	,63343
	S_3_T_240 correct answers per timeslot	9,7778	9	2,53859	,84620
Paaren 6	S_3_T_240 correct answers per timeslot	8,3333	3	2,08167	1,20185
	S_3_T_270 correct answers per timeslot	9,0000	3	2,00000	1,15470
Paaren 7	S_3_T_270 correct answers per timeslot	7,0000	1[b]	.	.
	S_3_T_300 correct answers per timeslot	7,0000	1[b]	.	.

a. Home_Language = Sesotho

b. Korrelation und T können nicht berechnet werden, da die Summe der Fallgewichtungen kleiner oder gleich 1 ist.

Test bei gepaarten Stichproben[a]

		Gepaarte Differenzen							
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere	T	df	Sig. (2-seitig)
Paaren 1	S_3_T_90 correct answers per timeslot - S_3_T_120 correct answers per timeslot	-,97059	,93696	,16069	-1,29751	-,64367	-6,040	33	,000
Paaren 2	S_3_T_120 correct answers per timeslot - S_3_T_150 correct answers per timeslot	-1,20690	,77364	,14366	-1,50117	-,91262	-8,401	28	,000
Paaren 3	S_3_T_150 correct answers per timeslot - S_3_T_180 correct answers per timeslot	-1,00000	1,14018	,24881	-1,51900	-,48100	-4,019	20	,001
Paaren 4	S_3_T_180 correct answers per timeslot - S_3_T_210 correct answers per timeslot	-,84615	,68874	,19102	-1,26235	-,42995	-4,430	12	,001
Paaren 5	S_3_T_210 correct answers per timeslot - S_3_T_240 correct answers per timeslot	-,88889	,92796	,30932	-1,60218	-,17560	-2,874	8	,021
Paaren 6	S_3_T_240 correct answers per timeslot - S_3_T_270 correct answers per timeslot	-,66667	,57735	,33333	-2,10088	,76755	-2,000	2	,184

a. Home_Language = Sesotho

Statistik bei gepaarten Stichproben[a]

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_3_T_90 correct answers per timeslot	5,9459	37	1,43267	,23553
	S_3_T_120 correct answers per timeslot	7,3514	37	1,53145	,25177
Paaren 2	S_3_T_120 correct answers per timeslot	7,2581	31	1,61178	,28949
	S_3_T_150 correct answers per timeslot	8,4516	31	1,89453	,34027
Paaren 3	S_3_T_150 correct answers per timeslot	8,0800	25	1,82392	,36478
	S_3_T_180 correct answers per timeslot	8,8000	25	1,80278	,36056
Paaren 4	S_3_T_180 correct answers per timeslot	7,9167	12	1,31137	,37856
	S_3_T_210 correct answers per timeslot	8,8333	12	1,80067	,51981
Paaren 5	S_3_T_210 correct answers per timeslot	8,4000	5	2,19089	,97980
	S_3_T_240 correct answers per timeslot	9,6000	5	2,07364	,92736
Paaren 6	S_3_T_240 correct answers per timeslot	6,0000	1[b]	.	.
	S_3_T_270 correct answers per timeslot	6,0000	1[b]	.	.
Paaren 7	S_3_T_270 correct answers per timeslot	6,0000	1[b]	.	.
	S_3_T_300 correct answers per timeslot	7,0000	1[b]	.	.

a. Home_Language = Isi-Zulu

b. Korrelation und T können nicht berechnet werden, da die Summe der Fallgewichtungen kleiner oder gleich 1 ist.

Test bei gepaarten Stichproben[a]

		Gepaarte Differenzen					T	df	Sig. (2-seitig)
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere			
Paaren 1	S_3_T_90 correct answers per timeslot - S_3_T_120 correct answers per timeslot	-1,40541	,76229	,12532	-1,65956	-1,15125	-11,215	36	,000
Paaren 2	S_3_T_120 correct answers per timeslot - S_3_T_150 correct answers per timeslot	-1,19355	,94585	,16988	-1,54049	-,84661	-7,026	30	,000
Paaren 3	S_3_T_150 correct answers per timeslot - S_3_T_180 correct answers per timeslot	-,72000	,54160	,10832	-,94356	-,49644	-6,647	24	,000
Paaren 4	S_3_T_180 correct answers per timeslot - S_3_T_210 correct answers per timeslot	-,91667	1,08362	,31282	-1,60517	-,22816	-2,930	11	,014
Paaren 5	S_3_T_210 correct answers per timeslot - S_3_T_240 correct answers per timeslot	-1,20000	,83666	,37417	-2,23885	-,16115	-3,207	4	,033

a. Home_Language = Isi-Zulu

Statistik bei gepaarten Stichproben[a]

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_3_T_90 correct answers per timeslot	6,6552	29	1,47057	,27308
	S_3_T_120 correct answers per timeslot	8,0000	29	1,62569	,30188
Paaren 2	S_3_T_120 correct answers per timeslot	7,5000	20	1,39548	,31204
	S_3_T_150 correct answers per timeslot	8,7500	20	1,44641	,32343
Paaren 3	S_3_T_150 correct answers per timeslot	8,6154	13	1,55662	,43173
	S_3_T_180 correct answers per timeslot	9,8462	13	2,03495	,56439
Paaren 4	S_3_T_180 correct answers per timeslot	8,8750	8	1,72689	,61055
	S_3_T_210 correct answers per timeslot	10,1250	8	2,10017	,74252
Paaren 5	S_3_T_210 correct answers per timeslot	9,6667	3	3,21455	1,85592
	S_3_T_240 correct answers per timeslot	10,0000	3	2,64575	1,52753
Paaren 6	S_3_T_240 correct answers per timeslot	11,0000	1[b]	.	.
	S_3_T_270 correct answers per timeslot	12,0000	1[b]	.	.
Paaren 7	S_3_T_270 correct answers per timeslot	12,0000	1[b]	.	.
	S_3_T_300 correct answers per timeslot	14,0000	1[b]	.	.

a. Home_Language = Africaans

b. Korrelation und T können nicht berechnet werden, da die Summe der Fallgewichtungen kleiner oder gleich 1 ist.

Test bei gepaarten Stichproben[a]

		Gepaarte Differenzen							
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere	T	df	Sig. (2-seitig)
Paaren 1	S_3_T_90 correct answers per timeslot - S_3_T_120 correct answers per timeslot	-1,34483	,85673	,15909	-1,67071	-1,01894	-8,453	28	,000
Paaren 2	S_3_T_120 correct answers per timeslot - S_3_T_150 correct answers per timeslot	-1,25000	,78640	,17584	-1,61805	-,88195	-7,109	19	,000
Paaren 3	S_3_T_150 correct answers per timeslot - S_3_T_180 correct answers per timeslot	-1,23077	,72501	,20108	-1,66889	-,79265	-6,121	12	,000
Paaren 4	S_3_T_180 correct answers per timeslot - S_3_T_210 correct answers per timeslot	-1,25000	,70711	,25000	-1,84116	-,65884	-5,000	7	,002
Paaren 5	S_3_T_210 correct answers per timeslot - S_3_T_240 correct answers per timeslot	-,33333	,57735	,33333	-1,76755	1,10088	-1,000	2	,423

a. Home_Language = Africaans

9.13 Development of performances depending on test time and home language: Subtest Four

Statistik bei gepaarten Stichproben[a]

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_4_T_180 correct answers per timeslot	3,8800	25	2,48864	,49773
	S_4_T_210 correct answers per timeslot	4,7200	25	3,20832	,64166
Paaren 2	S_4_T_210 correct answers per timeslot	4,0476	21	2,81915	,61519
	S_4_T_240 correct answers per timeslot	4,2857	21	3,01899	,65880
Paaren 3	S_4_T_240 correct answers per timeslot	4,2353	17	2,75067	,66714
	S_4_T_270 correct answers per timeslot	4,4706	17	2,87484	,69725
Paaren 4	S_4_T_270 correct answers per timeslot	3,6364	11	2,37793	,71697
	S_4_T_300 correct answers per timeslot	3,8182	11	2,27236	,68514

a. Home_Language = Sesotho

Test bei gepaarten Stichproben[a]

		Gepaarte Differenzen							
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere	T	df	Sig. (2-seitig)
Paaren 1	S_4_T_180 correct answers per timeslot - S_4_T_210 correct answers per timeslot	-,84000	1,06771	,21354	-1,28073	-,39927	-3,934	24	,001
Paaren 2	S_4_T_210 correct answers per timeslot - S_4_T_240 correct answers per timeslot	-,23810	,43644	,09524	-,43676	-,03943	-2,500	20	,021
Paaren 3	S_4_T_240 correct answers per timeslot - S_4_T_270 correct answers per timeslot	-,23529	,43724	,10605	-,46010	-,01049	-2,219	16	,041
Paaren 4	S_4_T_270 correct answers per timeslot - S_4_T_300 correct answers per timeslot	-,18182	,40452	,12197	-,45358	,08994	-1,491	10	,167

a. Home_Language = Sesotho

Statistik bei gepaarten Stichproben[a]

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_4_T_180 correct answers per timeslot	4,0370	27	1,91113	,36780
	S_4_T_210 correct answers per timeslot	4,7037	27	2,16288	,41625
Paaren 2	S_4_T_210 correct answers per timeslot	4,4211	19	2,38783	,54781
	S_4_T_240 correct answers per timeslot	5,1579	19	2,36322	,54216
Paaren 3	S_4_T_240 correct answers per timeslot	4,2500	12	1,65831	,47871
	S_4_T_270 correct answers per timeslot	4,6667	12	1,92275	,55505
Paaren 4	S_4_T_270 correct answers per timeslot	4,8000	10	2,04396	,64636
	S_4_T_300 correct answers per timeslot	5,2000	10	2,44040	,77172

a. Home_Language = Isi-Zulu

Test bei gepaarten Stichproben[a]

		Gepaarte Differenzen							
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere	T	df	Sig. (2-seitig)
Paaren 1	S_4_T_180 correct answers per timeslot - S_4_T_210 correct answers per timeslot	-,66667	,73380	,14122	-,95695	-,37639	-4,721	26	,000
Paaren 2	S_4_T_210 correct answers per timeslot - S_4_T_240 correct answers per timeslot	-,73684	,65338	,14989	-1,05176	-,42192	-4,916	18	,000
Paaren 3	S_4_T_240 correct answers per timeslot - S_4_T_270 correct answers per timeslot	-,41667	,51493	,14865	-,74384	-,08950	-2,803	11	,017
Paaren 4	S_4_T_270 correct answers per timeslot - S_4_T_300 correct answers per timeslot	-,40000	,51640	,16330	-,76941	-,03059	-2,449	9	,037

a. Home_Language = Isi-Zulu

Statistik bei gepaarten Stichproben[a]

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_4_T_180 correct answers per timeslot	5,1111	18	2,94836	,69493
	S_4_T_210 correct answers per timeslot	5,7222	18	3,00599	,70852
Paaren 2	S_4_T_210 correct answers per timeslot	5,3846	13	2,56705	,71197
	S_4_T_240 correct answers per timeslot	5,7692	13	2,80339	,77752
Paaren 3	S_4_T_240 correct answers per timeslot	5,2222	9	2,63523	,87841
	S_4_T_270 correct answers per timeslot	5,6667	9	3,00000	1,00000
Paaren 4	S_4_T_270 correct answers per timeslot	5,4000	5	3,04959	1,36382
	S_4_T_300 correct answers per timeslot	6,2000	5	3,70135	1,65529

a. Home_Language = Africaans

Test bei gepaarten Stichproben[a]

		Gepaarte Differenzen							
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere	T	df	Sig. (2-seitig)
Paaren 1	S_4_T_180 correct answers per timeslot - S_4_T_210 correct answers per timeslot	-,61111	,60768	,14323	-,91331	-,30892	-4,267	17	,001
Paaren 2	S_4_T_210 correct answers per timeslot - S_4_T_240 correct answers per timeslot	-,38462	,65044	,18040	-,77767	,00844	-2,132	12	,054
Paaren 3	S_4_T_240 correct answers per timeslot - S_4_T_270 correct answers per timeslot	-,44444	,52705	,17568	-,84957	-,03932	-2,530	8	,035
Paaren 4	S_4_T_270 correct answers per timeslot - S_4_T_300 correct answers per timeslot	-,80000	1,30384	,58310	-2,41893	,81893	-1,372	4	,242

a. Home_Language = Africaans

9.14 Development of performances depending on test time and home language: Subtest Five

Statistik bei gepaarten Stichproben[a]

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_5_T_180 correct answers per timeslot	6,7692	13	2,55453	,70850
	S_5_T_210 correct answers per timeslot	7,6923	13	2,78043	,77115
Paaren 2	S_5_T_210 correct answers per timeslot	7,8571	7	2,73426	1,03345
	S_5_T_240 correct answers per timeslot	8,1429	7	2,54484	,96186
Paaren 3	S_5_T_240 correct answers per timeslot	7,6667	6	2,42212	,98883
	S_5_T_270 correct answers per timeslot	8,3333	6	1,86190	,76012
Paaren 4	S_5_T_270 correct answers per timeslot	7,2500[b]	4	,95743	,47871
	S_5_T_300 correct answers per timeslot	7,2500[b]	4	,95743	,47871

a. Home_Language = Sesotho

b. Korrelation und T können nicht berechnet werden, da der Standardfehler der Differenz gleich 0 ist.

Test bei gepaarten Stichproben[a]

		Gepaarte Differenzen							
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere	T	df	Sig. (2-seitig)
Paaren 1	S_5_T_180 correct answers per timeslot - S_5_T_210 correct answers per timeslot	-,92308	,95407	,26461	-1,49962	-,34654	-3,488	12	,004
Paaren 2	S_5_T_210 correct answers per timeslot - S_5_T_240 correct answers per timeslot	-,28571	,48795	,18443	-,73699	,16556	-1,549	6	,172
Paaren 3	S_5_T_240 correct answers per timeslot - S_5_T_270 correct answers per timeslot	-,66667	1,21106	,49441	-1,93760	,60426	-1,348	5	,235

a. Home_Language = Sesotho

Statistik bei gepaarten Stichproben[a]

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_5_T_180 correct answers per timeslot	6,9231	13	3,22649	,89487
	S_5_T_210 correct answers per timeslot	7,5385	13	3,33205	,92414
Paaren 2	S_5_T_210 correct answers per timeslot	6,8750	8	2,79987	,98990
	S_5_T_240 correct answers per timeslot	7,1250	8	2,69590	,95314
Paaren 3	S_5_T_240 correct answers per timeslot	5,5000[b]	4	1,91485	,95743
	S_5_T_270 correct answers per timeslot	5,5000[b]	4	1,91485	,95743
Paaren 4	S_5_T_270 correct answers per timeslot	5,6667	3	2,30940	1,33333
	S_5_T_300 correct answers per timeslot	6,0000	3	2,64575	1,52753

a. Home_Language = Isi-Zulu

b. Korrelation und T können nicht berechnet werden, da der Standardfehler der Differenz gleich 0 ist.

Test bei gepaarten Stichproben[a]

		Gepaarte Differenzen							
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere	T	df	Sig. (2-seitig)
Paaren 1	S_5_T_180 correct answers per timeslot - S_5_T_210 correct answers per timeslot	-,61538	,86972	,24122	-1,14095	-,08982	-2,551	12	,025
Paaren 2	S_5_T_210 correct answers per timeslot - S_5_T_240 correct answers per timeslot	-,25000	,46291	,16366	-,63700	,13700	-1,528	7	,170
Paaren 4	S_5_T_270 correct answers per timeslot - S_5_T_300 correct answers per timeslot	-,33333	,57735	,33333	-1,76755	1,10088	-1,000	2	,423

a. Home_Language = Isi-Zulu

Statistik bei gepaarten Stichproben[a]

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_5_T_180 correct answers per timeslot	7,8889[b]	9	3,10018	1,03339
	S_5_T_210 correct answers per timeslot	7,8889[b]	9	3,10018	1,03339
Paaren 2	S_5_T_210 correct answers per timeslot	8,7500	4	2,62996	1,31498
	S_5_T_240 correct answers per timeslot	9,0000	4	2,70801	1,35401
Paaren 3	S_5_T_240 correct answers per timeslot	11,0000	1[c]	.	.
	S_5_T_270 correct answers per timeslot	12,0000	1[c]	.	.
Paaren 4	S_5_T_270 correct answers per timeslot	.	0[d]	.	.
	S_5_T_300 correct answers per timeslot	.	0[d]	.	.

a. Home_Language = Africaans

b. Korrelation und T können nicht berechnet werden, da der Standardfehler der Differenz gleich 0 ist.

c. Korrelation und T können nicht berechnet werden, da die Summe der Fallgewichtungen kleiner oder gleich 1 ist.

d. Korrelation und T können nicht berechnet werden, da keine gültigen Paare vorhanden sind.

Test bei gepaarten Stichproben[a]

		Gepaarte Differenzen							
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere	T	df	Sig. (2-seitig)
Paaren 2	S_5_T_210 correct answers per timeslot - S_5_T_240 correct answers per timeslot	-,25000	,50000	,25000	-1,04561	,54561	-1,000	3	,391

a. Home_Language = Africaans

9.15 Development of performances depending on test time and home language: Subtest Six

Statistik bei gepaarten Stichproben[a]

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_6_T_180 correct answers per timeslot	4,5714	7	3,20713	1,21218
	S_6_T_210 correct answers per timeslot	5,8571	7	3,80476	1,43806
Paaren 2	S_6_T_210 correct answers per timeslot	4,0000	2	2,82843	2,00000
	S_6_T_240 correct answers per timeslot	4,5000	2	3,53553	2,50000
Paaren 3	S_6_T_240 correct answers per timeslot	7,0000	1[b]	.	.
	S_6_T_270 correct answers per timeslot	8,0000	1[b]	.	.
Paaren 4	S_6_T_270 correct answers per timeslot	.	0[c]	.	.
	S_6_T_300 correct answers per timeslot	.	0[c]	.	.

a. Home_Language = Sesotho

b. Korrelation und T können nicht berechnet werden, da die Summe der Fallgewichtungen kleiner oder gleich 1 ist.

c. Korrelation und T können nicht berechnet werden, da keine gültigen Paare vorhanden sind.

Test bei gepaarten Stichproben[a]

		Gepaarte Differenzen					T	df	Sig. (2-seitig)
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere			
Paaren 1	S_6_T_180 correct answers per timeslot - S_6_T_210 correct answers per timeslot	-1,28571	,95119	,35952	-2,16542	-,40601	-3,576	6	,012
Paaren 2	S_6_T_210 correct answers per timeslot - S_6_T_240 correct answers per timeslot	-,50000	,70711	,50000	-6,85310	5,85310	-1,000	1	,500

a. Home_Language = Sesotho

Statistik bei gepaarten Stichproben[a]

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_6_T_180 correct answers per timeslot	8,2857	7	1,49603	,56544
	S_6_T_210 correct answers per timeslot	9,8571	7	1,21499	,45922
Paaren 2	S_6_T_210 correct answers per timeslot	10,2000	5	1,30384	,58310
	S_6_T_240 correct answers per timeslot	10,8000	5	1,09545	,48990
Paaren 3	S_6_T_240 correct answers per timeslot	10,6667	3	1,15470	,66667
	S_6_T_270 correct answers per timeslot	11,0000	3	1,00000	,57735
Paaren 4	S_6_T_270 correct answers per timeslot	.	0[b]	.	.
	S_6_T_300 correct answers per timeslot	.	0[b]	.	.

a. Home_Language = Isi-Zulu

b. Korrelation und T können nicht berechnet werden, da keine gültigen Paare vorhanden sind.

Test bei gepaarten Stichproben[a]

		Gepaarte Differenzen					T	df	Sig. (2-seitig)
					95% Konfidenzintervall der Differenz				
		Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes	Untere	Obere			
Paaren 1	S_6_T_180 correct answers per timeslot - S_6_T_210 correct answers per timeslot	-1,57143	1,27242	,48093	-2,74822	-,39464	-3,267	6	,017
Paaren 2	S_6_T_210 correct answers per timeslot - S_6_T_240 correct answers per timeslot	-,60000	,54772	,24495	-1,28009	,08009	-2,449	4	,070
Paaren 3	S_6_T_240 correct answers per timeslot - S_6_T_270 correct answers per timeslot	-,33333	,57735	,33333	-1,76755	1,10088	-1,000	2	,423

a. Home_Language = Isi-Zulu

Statistik bei gepaarten Stichproben[a]

		Mittelwert	N	Std.-Abweichung	Standardfehler des Mittelwertes
Paaren 1	S_6_T_180 correct answers per timeslot	5,0000	1[b]	.	.
	S_6_T_210 correct answers per timeslot	5,0000	1[b]	.	.
Paaren 2	S_6_T_210 correct answers per timeslot	.	0[c]	.	.
	S_6_T_240 correct answers per timeslot	.	0[c]	.	.
Paaren 3	S_6_T_240 correct answers per timeslot	.	0[c]	.	.
	S_6_T_270 correct answers per timeslot	.	0[c]	.	.
Paaren 4	S_6_T_270 correct answers per timeslot	.	0[c]	.	.
	S_6_T_300 correct answers per timeslot	.	0[c]	.	.

a. Home_Language = Africaans

b. Korrelation und T können nicht berechnet werden, da die Summe der Fallgewichtungen kleiner oder gleich 1 ist.

c. Korrelation und T können nicht berechnet werden, da keine gültigen Paare vorhanden sind.