THE PERFORMANCE OF SOUTH AFRICAN PILOTS ON COGNITIVE ABILITY ASSESSMENTS

by

Sinombongo Simpson

Submitted in partial fulfilment of the requirements for the degree of Masters of Commerce: Industrial and Organisational Psychology to be awarded at the Nelson Mandela Metropolitan University

17 May 2016

Supervisor: Dr Gert Louw (Posthumously)

Co-Supervisor: Prof R Van Niekerk

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the following people, who helped me to make this research project possible:

- To my Almighty God and Saviour, Jesus Christ, who gave me strength, hope and perseverance, when I had nothing left within me;
- My husband, KJ, for his support and encouragement and children for their understanding when I couldn't spend time with them;
- My family and friends, for their support and believing that I can do it;
- Dr Gerrit Louw (may his soul continue to rest in peace), my supervisor, for guiding the initial stages of this study;
- Prof Roelf Van Niekerk, my co-supervisor, for your honesty, encouragement and professional guidance during the study;
- Top Talent Solutions (TTS) – Prof Hennie Kriek, Dr Fred Guest, Ms. Zinhle Mthombeni and Mr. Renier Lewis. This study would not have been possible without the professional and personal assistance that you gave me;
- Dr Danie Venter for assisting with the statistical analysis;
- My employer and manager for allowing me to use the information for the research;
- My colleagues for their encouragement and support;
- A special thank you to Dr Jacqui Baumgardt who edited this research, for her zeal, encouragement and valuable guidance;
- Last but not least, to Esther Venter, my mentor and special friend who encouraged me to pursue this degree; your support means a lot.

DECLARATION

Student number: 213339110

I, Sinombongo Simpson declare that "**The performance of South African pilots on cognitive ability assessments**" is my own work. All sources used or quoted have been documented and recognized and according to my knowledge; this thesis has not been previously submitted in full or partial fulfilment of the requirements of an equivalent or higher qualification at any other recognized educational institution.

Sino Simpson Date

ABSTRACT

In South Africa, the issue of lack of black pilot skills is a most talked about topic in commercial airlines. Airlines need to find the right set of skills and attitude to operate safely and successfully. Airline safety remains a topic of discussion, therefore airlines need ensure robustness of their selection processes.

There are many requirements that a person has to meet before s/he can be offered a first officer job at an airline. The majority of the pilots in South Africa are white males. The study aimed to determine if there are statistically significant differences in cognitive ability test results between Black and White applicants for pilot positions with a South Africa airline. The test battery included a verbal reasoning test, numerical reasoning test, visual thinking, abstract reasoning test, spatial reasoning test, short term memory test, monitoring ability test, hand-eye coordination, sense of orientation test, reactivity test, and a multi-tasking test. The data were extracted with permission from a database maintained on behalf of the airline by an online test provider and subjected to a statistical analysis using measures of central tendency, and spread, in order to report on the significance of the differences between the groups. There is strong evidence that the White group performed better than the Black group on the majority of the tests.

It is recommended that further research is done in order to determine the causes of these differences. A comparative study of results on other airlines' recruitment selection tests and an exploratory study of the impact of socio-economic factors, education quality and language on cognitive ability tests are recommended.

Key Words

Intelligence, cognitive ability, selection, race, psychometric assessment; descriptive statistics, t-tests, ANOVA, airline industry

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1:

ORIENTATION

1.1 INTRODUCTION

The introduction of the *Employment Equity Act No. 55 of 1998* (EEA) (RSA), in South Africa was to redress the disparities caused by the apartheid system and other discriminatory laws and practices. This research seeks to determine how the performance of Blacks on cognitive ability assessments compares with performance of Whites on the same tests that are used for selection of trainee pilots.

1.2 BACKGROUND OF AND MOTIVATION FOR THE STUDY

South Africa's armed forces, known as the South African National Defence Force (SANDF) was created in 1994. Previously known simply as the South African Defence Force (SADF), the new force consists of the forces of the old SADF, as well as the forces of the African nationalist groups, namely Umkhonto we Sizwe (MK), Azanian People's Liberation Army (APLA), and the former homelands defence forces (Transkei, Bophuthatswana, Ciskei, Venda, Gazankulu, KaNgwane, KwaNdebele, KwaZulu, Lebowa, and QwaQwa). The SANDF is subdivided into four branches, the South African Army, the South African Air Force, the South African Navy, and the South African Military Health Services.

The South African Air Force (SAAF), formed on 1 February 1920, is the second oldest air force in the Commonwealth. The air arm played a major role in securing victory for the Allies during the Second World War, in the 1948/49 Berlin Airlift, and in Korea in the 1950s (Wingrin, 2013). The SAAF assisted Rhodesia in the 1960s and 1970s, made a major contribution to the ‘Border’ or ‘Bush’ war in South West Africa

and Angola, participated in the transition to a new democracy in South Africa and continuously supports South African peace missions in Africa. It has also assisted in countless relief and rescue missions in southern Africa throughout this entire period. However, despite democracy, transformation of the air force is still a central issue, specifically transformation of the mind-set, behaviour and the organisation.

In addition to the SAAF, there exists within South Africa a large private air industry with a number of airlines such as the national airline South African Airways and its subsidiaries, SA Express and Mango Air; Comair and its subsidiaries, British Airways and Kulula, Airlink and a host of smaller privately-owned airlines offering charter flights. In addition, there are cargo services. The common denominator among all these airlines, is their need for qualified pilots.

According to Superior Pilot Services (2010), citing the South African Civil Aviation Authority (SACAA), audited civil aviation statistics revealed that:

> of the 26 022 aviation personnel licence holders South Africa had in 2012, only 2 776 were African, 670 were Coloured, 553 were Indian whilst 17 346 were White. A further examination of those numbers revealed that a mere 1 668 of those licence holders were African women, 383 of them were Coloured women, 238 were Indian women and 2169 were White women (n.p).

The civil aviation industry therefore has some way to go to see transformation that reflects the demographics of the country which has blacks as majority and the whites as minority. According to Boeing's projections (Defenceweb, 2015) Africa will need 18 000 pilots over the next twenty years. However, successful trainees seem to be hard to find. While annual intake figures are not accessible, in the successful pilot

wing cohort of the SAAF in 2014, there was one Black student, one Indian student and four White students. In the navigator course of the same year, there were three Black students and one White female student (Wingrin, 2014). These are very small numbers comparative to the cited demand.

It seems that a consideration of how trainee pilots are recruited is pertinent. Here we consider both the South African Air Force as well as civil aviation. Since statistics from the Department of Defence are unobtainable, the following information has been taken from SAAF, the unofficial website of the South African Airforce[1] (SAAF, 2015). Of approximately 4 000 who apply every year, only about 30 are chosen. To be considered for direct entry to the flying training programme, candidates must have full Matric exemption and they must have passed English and a second language, Mathematics and Science, the latter two preferably with distinctions. The selection process involves checking other criteria as well such as height, weight and disabilities. Those identified for possible selection then undergo psychometric tests, interviews and medical examinations, until the final thirty candidates (depending on SAAF requirement) are selected. In the civil aviation industry, the selection requirements are very similar (The Joint Aviation Awareness Programme, 2010).

1.3 PROBLEM STATEMENT

Assessments are used in pilot selection to determine levels of cognitive ability, emotional intelligence and job fit. According to Mataboge (2010), South African Airways' (SAA) failure to increase the number of pilots from previously disadvantaged communities over the past 16 years shows that the aviation industry

1 The author of the website states "The views expressed on this site are mine and do not necessarily reflect those of the South African Air Force, the SANDF, the Department of Defence or any other organisation or individual. No inference should be made that any of the contents of this site is/was/will be official policy of the SAAF or that it is 100% accurate. All information was obtained from public sources".

is far from reaching its transformation deadline. The South African Airforce Forum has had to step in and assist South Africa in producing more Black African pilots.

According to Airline Pilots Association of South Africa (ALPASA), a trade union representing pilots working for SAA, SA Express and Comair (British Airways SA/Kulula), the biggest obstacle to training more Black African pilots was the high cost of the training, a factor also mentioned by De Kock and Schlechter (2009). According to Mataboge (2010), government intervention is needed to avoid a continual shortage of Black African pilots. However, Black trainee pilots tend to fail the course in their numbers, with few meeting the requirements to 'earn their wings' so to speak.

Subsidisation of training costs is only one issue, and Mataboge (2010) citing Ngema, the air force's chief director of policy and plans, proposed that the academic requirements for acceptance as a trainee pilot be lowered and the actual training be improved to accommodate students with lower academic marks. This is a contentious issue, with Gigaba (2013) disagreeing with the notion of lowering the minimum requirements. In fact, he stated that "there is a need to ensure that the high quality of standards required by South African Civil Aviation Authority (CAA) as well as International Civil Aviation Organisation are maintained" (n. p.). According to Mataboge (2010), again citing Ngema, the problem lies in historical disadvantage in the schooling system where students do not get the top class results and disinctions needed to meet the entry requirements for pilot training. In addition even when they do make the grade for selection, aspiring Black pilots do not not perform adequately in the cognitive ability assessments as part of the holistic psychometric assessment battery that is used to make the final selection. The questions that are raised are what is needed in order to become a successful pilot, why are there so few students

getting through the training, and even more pertinent, why are there so few Black students?

In the academic literature, two issues seem to dominate current pilot selection research. Firstly, Spearman's general cognitive ability (*g*) has been determined as a key factor in predicting pilot success (De Kock & Schlechter, 2009). The *g*-factor relates to innate cognitive ability that does not rely on the social or environmental circumstances in which a person is brought up or lives. The second is a concept called "fluid intelligence" denoted as *Gf* (Catell, 1963). Fluid intelligence refers to "a complex human ability that allows us to adapt our thinking to a new cognitive problem or situation" (Jaeggi, Buschkuehl, Jonides & Perrig, 2008, p. 6829). It is these concepts that are explored in depth in the literature review chapter of this thesis.

Therefore, the research problem is: Could cognitive ability be the reason why most Black Africans do not make it through the pilot selection process?

1.4 RESEARCH OBJECTIVES

The general objective of this research is to investigate the performance differences on cognitive ability assessments for the Black group in comparison to the White group of South Africans.

The objectives of this study are to:

- Define what constitutes cognitive ability;
- Describe the battery of cognitive ability tests used in the selection of candidate pilots;
- Understand the factors impacting achievement in cognitive ability tests;

- Compare the results of cognitive ability assessments between applicants from African groups (Blacks) and Caucasian (Whites) for jobs as pilots in the commercial airline;
- Determine the significance of the differences between Black and White pilots in results on cognitive ability assessments.

1.5 THE PARADIGMATIC PERSPECTIVE

This research will be conducted in the field of Industrial and Organisational Psychology also known as I-O Psychology. It is the scientific study of human behaviour in the workplace and organisations to improve productivity and quality of life of people at work (Schreuder & Coetzee, 2010).

Existing psychometric assessment results will be used to collect data in this research. Psychometrics is a sub-discipline of industrial psychology (Furr & Bacharach, 2013). Psychometric assessments are also referred to as psychological assessments and, according to Urbina (2014), include aptitude tests to predict potential future behaviour, achievement tests and tests of ability to evaluate academic or vocational skill, personality tests such as might be used in psychiatric diagnoses, and neuropsychological tests that evaluate the role of brain dysfunction in emotional, cognitive, and behavioural disorders. Psychological assessment is a process that entails choosing what to assess and which psychological assessment instruments will be applied (Urbina, 2014). Choosing what to assess and which tools to use are determined by the purpose and the context of a particular assessment. Psychological assessment is usually applied for a specific purpose in a specific context (Groth-Marnat, 2009). In this study, the results of psychological assessments used for the selection of pilots were used.

1.6 RESEARCH DESIGN

A brief overview of the research design is provided in this section. A detailed description of the research design is provided in Chapter 3.

Research design is a strategic framework for action that serves as a bridge between research questions and the execution or implementation of the research. Considering the research problem and the research aims, a non-experimental, quantitative, comparative enquiry, where the goal is to identify, analyse and explain similarities and differences between societies within a particular situation (Hantrais, 1995), was deemed appropriate for this study. It attempts to identify variables that exist in a given situation and tries to describe the relationship that exists between these variables (Terre Blanche, Durrheim, & Painter, 2006).

The researcher has used a research design that is both coherent and valid. A design is coherent when the techniques of sampling, data collection and interpretation, as well as the context of the study fit logically within the paradigm of the research and with the purpose of the research. A design is also seen as valid if internal and external validity is met. Internal validity entails that the findings of the study follow in a direct and unproblematic way from its methods and therefore the results arrived at are sustained by the design itself and cannot be explained in any other way. External validity refers to the fact that the findings or conclusions of the research can be generalised beyond the confines of the study and study setting (Terre Blanche, Durrheim, & Painter, 2006).

The design of the study is comparative in nature where performance on existing online psychometric assessment results of Black job applicants at a major South African Airline is evaluated and compared with the results of White applicants. This

study will be beneficial to the airline and its selection process of pilots and cadet pilots; it will also be beneficial to all the SA airlines that use psychometric assessments as part of their selection process.

1.6.1 Data collection

Data collection is very important as it is the base for reaching conclusions in the research. The data of existing online psychometric assessments is housed in a secure database by the product house (TTS Solutions). This airline uses Scales Assessments designed by “cut-e”, a German company represented in South Africa by a test distributor (Top Talent Solutions, TTS) based in Pretoria, which also maintains the database of the airline’s test results. Research data were extracted from this database. The data were collected by the airline during the recruitment process for the selection of pilots and cadet pilots. The applicants completed 11 cognitive ability assessments and a personality questionnaire. The tests consisted of the following: verbal reasoning test, numerical reasoning test, visual thinking, abstract reasoning test, spatial reasoning test, short term memory test, monitoring ability test, hand-eye coordination, sense of orientation test, reactivity test, and a multi-tasking test. Access to the data was requested from the supervising psychologist at the airline. No individual was exposed in this process as the researcher used only unique codes as identifiers of the participants.

1.6.2 Data Processing

Careful data analysis was done to ensure that the data were coherent and that they matched the research paradigm and answered the research questions. Statistical analysis was used to make sense of the data; it helped the researcher to describe the data more explicitly and to make inferences about the characteristics of the

population on the basis of data from samples. A software programme, Statistica, a programme devised by Statsoft, was used to analyse data.

1.6.2.1 Descriptive statistics and inferential statistics

Descriptive statistics were used for summarising data, looking for trends and patterns, means, frequencies, measures of variability. Descriptive data analysis aims to describe the data by investigating the distribution of scores on each variable, and by determining whether the scores on different variables are related to each other (Terre Blanche, Durrheim, & Painter, 2006). Descriptive statistics include the mean, and standard deviation.

Central measures of tendency were used to estimate the centre-most score in a distribution as they best represent the data collected for a variable. There are three measures of central tendency; the mode, median and the mean. These are the most important descriptive statistics and form the basis of most advanced inferential statistics procedures.

Pearson's product moment correlation coefficient was used to determine relationships between variables or factors. ANOVA or t-tests were used to compare one, two or more groups' means. The results were interpreted and reported in statistical tables.

1.7 ETHICAL CONSIDERATIONS

Ethics are very important in research. Lack of commitment to ethical considerations could adversely affect the credibility of the research, the autonomy of the researchers, the quality of the research or the rights of the participants.

In this research, the identity of the participants was not exposed. Existing assessment data were used with age, gender and race as moderator variables. Informed consent and permission to use assessment data of the employees were obtained from the airline with its permission. The researcher strove to maintain objectivity and integrity while conducting this research. The researcher did not fabricate or falsify the data, disclosed the methodology and techniques of analysis when required and adhered to ethical publishing practices (Mouton, 2012).

1.8 CHAPTER LAYOUT

The chapter layout of the study is as follows:

Chapter 1: Orientation of the study

This chapter presents the orientation of the study, objectives and the overview of the research design and methodology.

Chapter 2: Literature Review

The literature review provides an in-depth literature study on the concept of intelligence, the factors affecting cognitive ability with a view to determining possible causes of performance disparities on different pilots' job applicants in a major South African airline. A detailed description of the various cognitive ability tests used by the airline is included in this chapter.

Chapter 3: Methodology

This chapter comprises a research article which includes information regarding the actual research study – methodology, sample, data collection and data analysis.

Chapter 4: Results

In this chapter, statistical information is provided, the hypotheses are tested and results are discussed.

Chapter 5: Conclusions and Recommendations

In the final chapter, a concluding discussion of the findings is presented. Limitations are indicated and suggestions/recommendations for future research are noted.

1.9 CHAPTER SUMMARY

In this chapter, the background and motivation for the research were discussed. The problem statement, the aims, the research paradigm, the research design and method together with the layout of the chapters were provided. In Chapter 2, a literature review regarding pilot selection will be discussed.

CHAPTER 2:

LITERATURE REVIEW

2.1 INTRODUCTION

For many years, assessment of humans has been done by using a variety of tools. This chapter briefly discusses the theory of cognitive abilities as constructs of intelligence, followed by a description of selection and the psychological tests included in the assessment battery of the selection process. Due to the high costs of pilot training, the emphasis has been placed on the selection process to ensure that the "right" people are selected. A central point of discussion is, however, around the performance on cognitive ability assessments for African (Black, Indian, and Coloured) and Caucasian (a White person, a person of European origin) pilot job applicants. The chapter concludes with a short summary of the main issues.

2.2 THEORETICAL FRAMEWORK

The theoretical basis of this thesis is the theory of intelligence. This concept is broad and the definition of intelligence is not universally agreed as it is a very complex construct (Fry & Wigglesworth, 2010). Schlinger (2012, p. 15) called it a "myth", stating that intelligence is contextually-based, and that one cannot determine whether someone is intelligent without seeing a demonstration of it within a particular situation, for example, a pilot in training. Nevertheless, several key theories have been developed.

Spearman's (1904) seminal work provides the basis for general intelligence theory, using the *g*-factor (or general intelligence) as the singular dominant factor. This was further developed by Thurstone (1931) to include second-order factors such as

spatial reasoning and memory. Cattell (1963) distinguished between fluid intelligence (*gf*), the underlying capacity for reasoning which is biologically determined, and crystallised intelligence (*gc*), the more accidental environmental aspects of skills, knowledge and experience. Then Carroll‘s (1993) three-stratum model incorporated the earlier theories and reinforced the hierarchy common to all of them, with a singular *g* at the apex of a pyramid. The middle stratum contains broad memory, retrieval and learning abilities such as verbal information (stating what has been heard) and intellectual skills (discriminating between facts and following instructions). The lowest stratum has a large number of more narrowly defined abilities. A further development of intelligence theories was Gardner’s (1991) multiple intelligences framework which posited that:

> we are all able to know the world through language, logical-mathematical analysis, spatial representation, musical thinking, the use of the body to solve problems or to make things, an understanding of other individuals, and an understanding of ourselves. Where individuals differ is in the strength of these intelligences – the so-called profile of intelligences – and in the ways in which such intelligences are invoked and combined to carry out different tasks, solve diverse problems, and progress in various domains. (p. 12)

2.2.1 Intelligence and cognition

Intelligence is a “general mental capability that involves the ability to reason, plan, solve problems, think abstractly, comprehend ideas and language, and learn”

(Thurstone, 1938, p. 131). For the purposes of this study, intelligence is viewed as a general measure of overall ability or potential ability (Hendrickz, 2014; Kumar, 2014). Intelligence and cognition are interrelated, but they are not the same. Intelligence encompasses cognition which is the method by which people assimilate and integrate knowledge, while intelligence involves both the assimilation of knowledge as well as the ability to apply that knowledge; in other words, the proficiency one has in a given area (Khan, 2012). Cognitive ability is the ability to learn, or potential to learn and acquire new knowledge and skill (Atherton, 2013). Cognitive processes lead to intelligence. If one perceives, reacts to, evaluates and understands things properly, it means that the person is intelligent.

2.2.2 Cognitive ability as a construct of intelligence

Cognitive ability refers to an individual's understanding of concepts and causal relationships; it involves the ability to think rationally, and abstractly, and to demonstrate insight and foresight in problem solving. This has been called "fluid intelligence" (Jaeggi, Buschkuehl, Jonides, & Shah, 2011, p. 10081). Spector (2012) defined cognitive ability as a person's capacity to learn or carry out a particular task. This leads to proximal consequences, such as a better quality of work and more appropriate decisions about everyday life issues. Higher cognitive ability is seen to make a positive contribution to society because it allows people to devise creative solutions to problems (Hitt, Ireland, Sirmon, & Trahms, 2011), and has been seen to bring about positive impacts on wealth, health, democracy, political and economic liberty and tolerance as cultures become ever more complex (Rinderman, 2012). In certain jobs, such as aviation, psychomotor abilities are more important than for other jobs (Carretta & Ree, 2000).

2.2.3 Cognitive Ability in the Context of Race

Cognitive abilities and differences between groups have been a debated subject in research, with the differences between ‘‘Whites’’ (people of European origin) and ‘‘Blacks’’ (people of sub-Saharan African origin), in particular, being the subject of scientific and non-scientific conflict. The scientific debate goes back to the mid-19th century, starting with the widespread use of standardised mental tests (commonly known as IQ tests) in World War I, where differences were found in average results with respect to ethnic and racial groups (Rushton & Jensen, 2005). This is what is commonly known as the “nature-nurture debate” (Kan, Wicherts, Dolan, & van der Maas, 2013, p. 2420) which revolves around whether cognitive ability is attributable to social, economic, and cultural factors, or whether genetic factors are also involved. It could possibly be a combination of both groups of factors (Goldhaber, 2012). Jensen (as cited in Rushton & Jensen, 2005, p. 235) concluded that IQ tests measure general cognitive ability; individual differences in IQ have a high heritability, therefore they can be attributed to the genetic make-up of a person. Compensatory educational programmes have proved generally ineffective in raising the IQs or school achievement of individuals or groups; and that, controversially, it is likely that Black-White group differences can be attributed to heritability. For example, when a black African job applicant from a multi-racial school competes and performs better than another black African job applicant from a previously disadvantaged school and vice versa, it is difficult to determine whether this is linked to a schooling system, culture or genetic cognitive competence.

There has been a lack of resolution of this important and controversial issue because of the difficulty of the subject matter, the political issues associated with it and the

emotions they arouse (Rushton & Jensen, 2005). Many critics claim that IQ tests developed for use in Western countries are not appropriate for use with other groups as they are culturally biased (Valencia, 2012), and therefore not appropriate for sub-Saharan Africans. However, the main evidence to support a claim of external bias would be if the test failed to predict performance for Africans. As legislated by the *Employment Equity Act 1998* (RSA), psychological testing and other similar assessments of an employee are prohibited unless the test or assessment being used is scientifically shown to be valid and reliable; can be applied fairly to all employees; and is not biased against any employee or group.

Richter, Mabaso and Hsiao (2015) found that test scores for Blacks have equal predictive validity as those for White South Africans. They also found that many of the factors that influence scores in blacks are the same as those for Whites (e.g. coming from an urban vs. a rural environment; being a science rather than an arts student; or having had practice on the tests). This appears to indicate that nurture factors have a considerable influence on cognitive ability. Nevertheless, although predictive validity is used to ensure that there is minimal bias, critics also suggest that the items have different meanings for Africans than they do for Whites or other groups. The reasons are not clear but Rushton and Jensen (2005) concluded that African students were less interested, more anxious, worked less efficiently, or gave up sooner on items they found difficult, perhaps because the problems were not significant to them.

According to Hartmann, Kruuse and Nyborg (2007), Spearman's hypothesis, which states that "racial differences in IQ between Blacks and Whites are due primarily to

differences in the *g*[2] factor" (p. 1), has been confirmed in SA using test item analysis. For example, a study done on 309 university students (Rushton & Skuy, 2000) found that the more individual items from the Raven's Standard Progressive Matrices[3] measured *g* (estimated by its item-total correlation), the more it correlated with the standardised African vs White difference on that item. In South Africa, members of the White, Indian, and Coloured population groups have been found to obtain higher mean scores than does the indigenous Black African majority (Rushton & Jensen, 2005).

The traditional definitions of race and ethnicity relate to biological and sociological factors respectively. Race, as a biological construct, denotes a group of people who share similar and distinct physical characteristics such as bone structure and skin, hair or eye colour, while ethnicity refers to cultural factors including nationality, regional culture, ancestry and language (Heere et al., 2015).

Sternberg (2012) showed that individuals with larger brain volumes have higher IQ scores, a finding based on the assumption that larger brains contain more neurons and synapses than smaller brains, which apparently make them more efficient. Rushton and Rushton (2003) reported that

> African-descended people (Blacks) have average cranial capacities of 1267 cm^3, European-descended people (Whites) 1347 cm^3, and East Asian-descended people (East Asians) have a size of 1364 cm^3. These brain size differences, containing millions of brain cells and hundreds of millions of

[2] The *g* factor is a universal and reliably measured distinction among humans in their ability to learn reason, and solve problems (Gottfredson, 2004, p. 35).

> synapses, were hypothesized to underlie the race differences on IQ tests, in which Blacks average an IQ of 85, Whites 100, and East Asians 106. (p. 139)

They found that there were significant correlations between racial variables (including cranial traits) and brain size, saying that:

> if the races did not differ in brain size, these correlations could not have been found. It must be concluded that the race differences in average brain size are securely established. As such, brain size-related variables provide the most likely biological mediators of the race differences in intelligence (*ibid*.).

Sternberg (2012) disputed Rushton and Rushton's 2003 findings that there are racial-group differences in measured intelligence, and states that race is a "socially constructed rather than biological variable" (p. 235). This contradicts the 'culture-only' theory which has yet to offer an explanation as to why 'nurture variables' (the researcher's own derived expression) such as poverty, social class, religious beliefs, cultural practices, absence of fathers, and parenting styles account for so little variance within groups.

2.2.4 Intelligence and Aptitude

Intelligence and aptitude are closely linked. Aptitude refers to a person's capability and is the ability to learn or to develop proficiency in an area (if provided with appropriate education or training). Aptitude is often latent and needs to be developed. If a person has an aptitude for something, his current performance would be expected to improve over time with training.

[3] The SPM is a well-validated measure of basic cognitive functioning (Raven, 2000, p. 1)

2.3 RECRUITMENT AND SELECTION PROCESSES

Organisations need to use reliable methods to evaluate attributes, characteristics and processes and attach a value that will aid decision-making for fulfilling organisational needs. In this context, assessments are used to measure the potential performance for employees (Austin, 2013). A few concepts related to selection are explored and defined in the next section.

According to Moerdyk (2009), selection is a process of finding the most suitable people to meet the current and future manpower needs of the organisation. Mondy (2010) refers to selection as the process of choosing from a group of applicants those individuals best suited for a particular position in an organisation. The effectiveness of organisations depends on the effectiveness of its employees; therefore, organisations select candidates by means of evaluating them and this forms a critical human resources function.

The selection process involves matching a person's skills, knowledge, interests and capabilities with the specific requirements or competencies of a specific job (Ahsan, Ho, & Khan, 2013). This process begins with the definition of organisational goals. Job design is based on these goals, which involves defining the duties and responsibilities of the individual. The criteria for successful performance are defined, followed by the traits, skills and qualities required in an individual (called the job specification). Finally, selection instruments are chosen and used. These selection methods determine whether the applicant possesses the desired traits, characteristics and skills. Figure 2.1 below explicates this process.

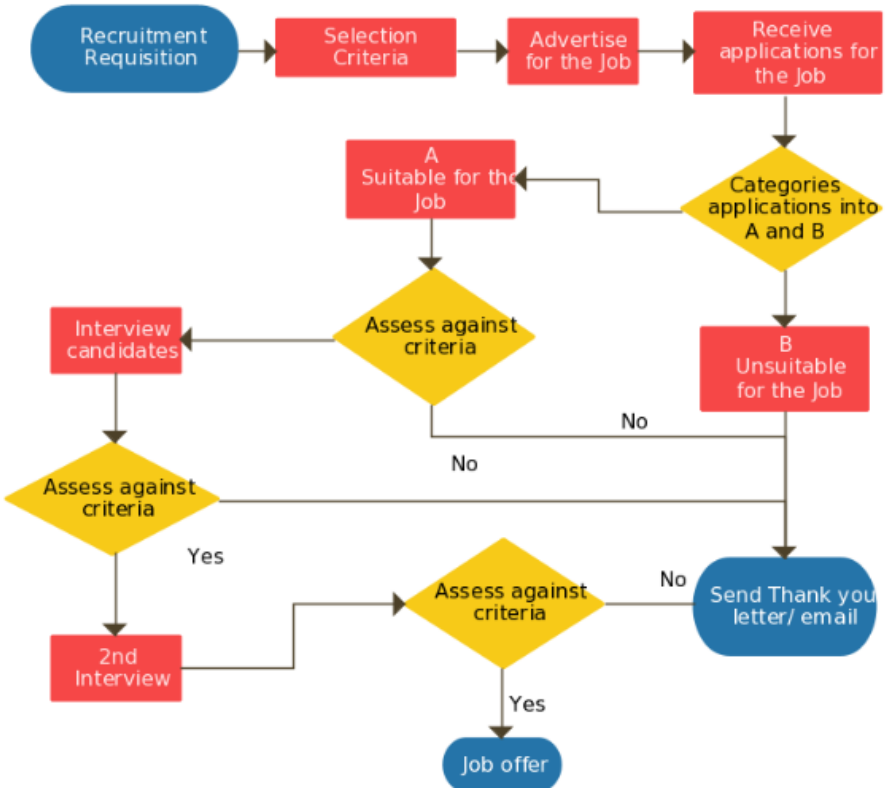

Figure 2.1: General recruitment and selection process

Source: Creately (2015)

Nell, Morris and Bohlander (2015) argued that there is no universal selection process and that most organisations use selection processes based on the “successive hurdles technique” (Flippo, 1947, p. 141). This means that to be selected for a position, applicants are required to be successful at each step of the process. In recent years, the recruitment process which precedes the selection process, has evolved from the traditional methods of direct mail postings and has since included recruitment agencies, newspaper and web advertising (Florea & Badea, 2013). The technology has become more advanced and includes social media, school alumni recruitment and e-recruitment. Schreuder and Coetzee (2010) echoed this view as their study found that that there would in future be an increase in the use of technology when it comes to recruitment and selection processes.

A key selection method that is used is that of psychological testing. Psychological assessment is a standardised procedure for sampling behaviour and describing it with categories and scores (Gregory, 2004). Psychological assessment and psychological testing are often used interchangeably. According to Flanagan and Harrison (2012), psychological assessments include psychometric testing as well as any other procedure used to assess human performance or potential. While psychological tests are only one step in the selection process, they can be extremely valuable as they have usually been validated by other test users and provide an objective method of measuring particular aspects of behaviour. Chakrabati and Fombonne (2014) suggested that by providing an understanding of a candidate's strengths and development areas, psychological tests add value to other assessment methods such as interviews in the broader selection process.

The psychometric approach considers intelligence as an aggregate of abilities (Srivastava & Misra, 2007). These ability tests include general and specific ability tests, personality questionnaires, motivation and interest tests, and emotional intelligence tests. Examples include traditional paper-and-pencil neuropsychological tests like the Raven's Colored Progressive Matrices; the Peabody Picture Vocabulary Test; the Category Test subtest of the Halstead Reitan Battery; SHL's personal, technical and critical reasoning test batteries and the Wisconsin Card Sorting Test (Parsons, 2011, p. 277) as well as a number of computerised tests on neurocognitive function such as "CogSport, ImPACT, ANAM, and HeadMinder" (ibid.). In South Africa, psychological assessment is regarded as a psychological act in terms of the *Health Professions Act, No 56 of 1974* (RSA). The industrial and organisational psychologist's knowledge should enhance the effectiveness of the

selection process by applying proper psychological assessments as a key selection method.

2.4 ASSESSMENT OF PILOTS

It is important to identify the characteristics needed to be a good pilot and to determine the means to accurately measure those characteristics. It is further vital to understand the necessity and the importance of the rigorous pilot selection processes in relation to the above. Pilots operate in a demanding, dynamic environment that requires highly developed cognitive skills (Nicholson & O'Hare, 2014). In commercial aviation, the airlines emphasise public safety, optimal training and operating costs, customer satisfaction and making a profit. Pilot selection is a form of high-stakes selection due to the massive costs of training, high trainee ability requirements and costly repercussions of poor selection decisions (de Kock & Schlechter, 2009).

2.4.1 Historial perspective and recent trends in Pilot selection

In the literature on pilot selection, there are many references to the selection of military pilots; indeed, commercial aviation has often relied on the military for trained pilots. This means that there are fewer studies involving commercial pilots' selection processes.

Dockeray and Isaacs (1921) reported that prior to World War I, Italy was the first country with a pilot selection research programme. The Italians used measures of reaction time, emotional reaction, equilibrium, perception of muscular effort, and attention. At the same time, the French were investigating reaction time and

emotional stability. In the World War I era, Yerkes (1919) showed that measures of intelligence were valid predictors of pilot training success.

According to Carreta (2011), most of the World War I research reflected the use of Spearman's two-factor theory that demonstrated the existence of a general cognitive factor (*g*) and the test-unique factor (*gf*) of Thurstone (see Section 2.2 above). Thurstone (1938) changed the emphasis from Spearman's two-factor theory to the theory of multiple aptitudes which was later incorporated in multiple-aptitude batteries such as Differential Aptitude Tests (DAT), General Aptitude Test Battery, Armed Services Vocational Aptitude Battery (ASVAB) and the Air Force Officer Qualifying test (AFOQT) (Kay & Kennedy, 2013).

During World War II, there was a renewed interest in pilot selection which was influenced by Thurstone's multiple-aptitude theory (Carretta, 2000). Several ability measures for pilot selection which included intelligence, psychomotor skill, mechanical comprehension and spatial measures were used, mostly in a paper and pencil format (Carretta, 2011).

Modern selection methods differ completely from those used in World War I and II, according to Damos (2011) who stated that, in those early days, "No formal job analysis (as the term is understood today) was conducted to identify the knowledge, skills, and abilities needed for success as a pilot" (p. 24). Most modern pilot selection processes emphasise proficiency and fluency to ensure that pilots would be equipped to safely operate an aircraft in a complex, high-threat environment. According to IFATLA (2012), pilot training programmes should embrace the concept of competency (demonstrated ability) and fluency (ability to use skills instantly). Fluency is important because it can only be determined once a person has

demonstrated actions on an exercise without error. This concept becomes paramount in critical incidents.

Pilots are required to do recurrent training to keep their licences up to date and to adhere to the requirements of continuous professional development (Catton, 2009). This usually includes assessment on critical events which can lead to catastrophic loss of control if unrecognised. Prior to getting to the assessments stage, commercial pilots and applicants must demonstrate acceptable levels of academic, social psychomotor and social skills.

2.4.1.1 Academic requirements

Pilots should demonstrate basic mental abilities to learn and retain knowledge at an advanced level. In South Africa, the minimum academic entrance requirements are a Matriculation certificate with Mathematics or Statistics 101 and English coupled with one of the following: Physical Science or Computer Science (FlySAA, n. d.). These subjects form the foundation of the complex subjects they will have to grasp as pilots, such as physics, mechanics and aircraft systems. It is critical that these skills are evaluated before training. The nature of their job requires pilots to have mental acuity and be able to memorise and process vast amounts of information including numbers, procedures, rules and definitions, and make correct decisions that will result in the appropriate actions for any given situation while flying.

2.4.1.2 Social Skills

According to Martinussen (2014), a pilot should be able to perform his duties under stress and solve complex situations such as dealing with loss of an engine in flight,

bird strikes, or emergency landings. They have to be able to communicate with crew on board as well as ground staff such as air traffic controllers. Interacting with a variety of people from different backgrounds, cultures and educational levels requires a set of highly developed social skills.

2.4.1.3 Psychomotor skills

To safely operate a modern aircraft in a complex and technologically advanced cockpit, a pilot needs highly developed hand-eye coordination and the mental capacity to maintain spatial orientation under normal situations as well as in stressful emergency environment (McMahon & Newman, 2015). Most psychomotor skills are developed early in life. They can be refined by training but if the fundamentals are not in place, then it is unlikely that an adult could learn such skills from scratch (cut-e, 2013).

2.4.1.4 Physical fitness

Pilots need adequate levels of fitness to cope with the demands of the job, for example, long flying hours, which can cause mental fatigue. Nagai et al. (2014) found that pilots need strength in their backs and necks in order to cope with long hours of sitting in a cockpit. Bezerra, Shimano and Campos (2015) found that muscular fitness helps pilots to cope with the physical forces, such as G-load acceleration exerted on their bodies as well as manual manipulation of cockpit controls. Anaerobic fitness is an important physiological component of G-load tolerance and anaerobic training with weights and counter-resistance results in significant increases this tolerance. Bezerra et al. (2015) stated that “the stress loads that individuals from this unusual profession are subjected to are extreme, and thus

a lack of preparation or clear understanding of some physical symptoms can lead them to physical breakdown or even death" (p. 27). The lack of physical fitness is also regarded as a contributory factor in fatal aircraft accidents. Another aspect of physical fitness, is medical fitness. In order to operate an aircraft as a pilot, a Class 1 medical certificate (CAA, 2015) is required. Applicants for a Class 1 medical certificate:

> must be free from any risk factor, disease or disability which renders them either unable, or likely to become suddenly unable, to perform assigned duties safely. These may include effects and/or adverse effects from the treatment of any condition and drugs or substances of abuse. … [and] must be free from any of the following, if it results in a degree of functional incapacity likely to interfere with the safe operation of an aircraft of with the safe performance of their duties: (a) Congenital or acquired abnormality; (b) active, latent, acute or chronic disability, disease or illness; (c) wound, injury, or outcome of operation (n. p.).

2.4.2 Duties and Responsibilities of a Pilot

HCareers.com (2009) provides a detailed job description of pilots. Airline pilots are responsible for the safe and economic operation and management of aircraft carrying passengers and/or freight. They ensure that the controls of the aircraft are working properly, check weather conditions, and liaise with air traffic control. Their job is not routine and demands unconventional hours in a very complex workplace. Pilots are given formal written procedures that prescribe in detail how the aircraft must be operated in each phase of flight, and who is required to do what and in what sequence.

The job of a pilot comes with heavy responsibility and personal commitment. Stringent training courses have to be passed followed by recurrent training every six months in order to maintain the relevant licence required for the job. According to Gradireland.com (2014), a pilot's typical job function includes making sure all information on the route, weather, passengers and aircraft is received; using that information to create a flight plan, which details the altitude for the flight, route to be taken and amount of fuel required; ensuring the fuel levels balance safety with economy and supervising the loading and fuelling of the aircraft; and making sure all safety systems are working properly.

Pilots should also be able to display non-technical (cognitive and social) skills which are essential for their duties. These include being able to cooperate with others, resolve conflict, support others, help the crew in demanding situations and make decisions (Flin et al., 2003).

To emphasise the importance of these skills for pilots, in determining the causes of fatal aircraft accidents, pilot error has been predominant although this is often exacerbated by weather problems or mechanical failure (PlaneCrashInfo, 2015). This is corroborated by Stowell (2015) who stated that in general aviation, upward of 95% of aircraft crashes seem to be caused by pilot error, either tactical (based on a pilot's decision) or operational (based on omissions in training).

There is a clear need for the selection of suitable candidates for the pilot profession. While selection comes before training, it is important that candidates who have attributes in all the areas discussed above should undergo a rigorous selection process (AviationSafety, 2015). A thorough selection process will yield a robust

workforce and also save the company money by focusing on those whose chances for success is the greatest (IFALTA, 2012; UJUH, 2014).

2.5 ASSESSMENT METHODS

The assessments methods that are typically used in the selection of pilots are discussed next.

2.5.1 Cognitive Ability

There are many cognitive ability tests available in South Africa and around the world. As mentioned in section 2.3.2, these tests measure some form of knowledge which has been acquired up until the time that the candidate takes the test. Most of these tests are classified as crystallised ability (*gc*) measures, the results of which can be influenced by schooling or culture (McArdle & Woodcock, 2014).

While different positions require different abilities, a pilot's specialised profession relies heavily on a set of specific cognitive abilities in order to respond quickly and accurately in a wide variety of tasks that depend on high-level visual cognition (Carretta et al., 2014). Flying is a complex activity that takes place in a rapidly changing and uncertain environment. The pilot must not only know how to operate the aircraft, the procedures and rules for the flight, but also have accurate and up-to-date situation awareness (Carretta, 2013).

2.5.2 Personality assessments

Cognitive ability assessments alone are not sufficient in predicting future performance. Combining ability assessments and personality assessments adds more value in incremental validity to predicting performance. A personality trait is a

tendency to act in a specific way, across various situations (Spector, 2012). Personality is reflected in behaviours that are relatively stable over time and consistent across situations. While some studies report that personality adds little to the prediction of pilot success, others report that certain aspects of personality have incremental predictive validity in batteries (Carreta, 2013). The competence required is typically assessed using personality assessments which are trait and/or behaviourally inclined.

2.5.3 Situational Judgement Tests (SJTs)

Situational Judgement Tests (SJTs) or Situational Judgement Inventories (SJIs) are a type of psychological test which presents the test-taker with realistic, hypothetical scenarios and asks the individual to identify the most appropriate response or to rank the responses in the order they feel is most effective. These assessments are designed to assess how one would handle situations that they could encounter in the job they are applying for (SHL, 2015). According to Patterson, Ashworth, Zibarras, Coan, Kerring and O'Neill (2012), SJTs have proven to be a cost-efficient assessment method that can assess a wide range of non-academic constructs and can be adapted to the selection context, such as selection of pilot trainees. They concluded that "SJTs are found to demonstrate less adverse impact than IQ tests and are positively received by candidates" (p. 850) and that they have "good predictive validity" (p. 858), compared with tests of IQ and personality, which makes them useful for selection of staff.

2.5.4 Simulation assessments

Aviation software provides training for pilots (Mavin & Roth, 2015). This software is also used to evaluate flying skills and forms part of the selection processes for most

airlines. Candidates are assessed on their hand-flown, raw data instrument flying skills. As an assessment tool, simulation has “greater benefits than traditional paper-and-pencil tests because of the instant feedback, interactions, and visuals” that can be provided (Quellmalz, Timms, Silberglitt, & Buckley, 2012, p. 382). Animations, such as those provided by aircraft simulators, are considered particularly useful for providing “visualisations of dynamic phenomena that are not easily observable in real space and time scales” (*ibid*., p. 369), for example, emergency situations such as crash landings. They contribute to spatial reasoning ability, and tend to eliminate individuals’ initial differences, which makes them ideal for training and assessing all trainee pilots irrespective of their socio-economic status (SES).

2.5.5 Interviews

According to Shouksmith (2015), interviews are opportunities for personal contact between candidates and employers. They are used for a number of vital purposes including to determine the candidate’s experience, level of training and its applicability, and also to evaluate the candidate’s characteristics and cognitive functioning. Interviews may reveal the behaviour of candidate in a situation that can contribute to performance in the job being he is recruited for (Lewis, 2010). According to Bangeter, Rouilin and König (2012), structured interviews have predictive validity. Assessing predictive validity involves establishing that the scores from a measurement procedure (e.g. an interview) make accurate predictions about the construct they represent (e.g., constructs like intelligence or achievement) (Laerd, 2015) but this can only be proven if the measurement is taken some time after the test has been conducted.

2.5.6 Psychological profile of a pilot

According to IATA (2012), the following competencies are what is commonly evaluated in pilot aptitude screening. These may vary from one airline to the other and may not be equally applied to all four different groups of pilots (*ab initio*, first officers, senior first officers and captains).

- Basic abilities, including logical abilities, spatial reasoning and orientation, visual processing, concentration, perceptual speed, technical comprehension and memory capacity;
- Operational competencies, including psychomotor tasks (e.g., multitasking, information processing and reactivity) and strategic competencies (e.g., compensatory strategies, decision-making and problem-solving);
- Social competencies, including communication skills, cooperation, and assertiveness. Pilots must be able to communicate with their on-board crew, ground staff, controllers and passengers, and must be able to take charge in any situation. The word ‘captain’ is indicative of the status of an airline pilot in terms of leadership.
- Personality traits, including self-discipline, self-organisation and stress management. For example, pilots have to be good time managers, not drink alcohol before flying, and remain calm in emergency situations.
- Professional competencies, including abiding by regulations and procedures, as well as staying up to date through regular training.

The decision to use certain instruments over others should be based on job descriptions and job specifications. Job analysis is central to establishing job entry and performance requirements or competencies (Moerdyk, 2012). Each instrument

must be relevant and related to the position and should be used to assess whether the behaviour of the candidate is in line with the competencies required. Table 2.1 below illustrates the predictive power that comes with using a combination of different assessments in order to make a selection decision.

Table 2.1: Predictive power of assessments

1	**Perfect Prediction (does not exist)**
.63	Cognitive Ability Tests combined with structured interviews
.54	Work Sample Test
.51	Structured Interviews
.51	Cognitive Ability Tests
.40	Personality Tests
.37	Assessment Centres
.35	Biographical Data
.26	Reference Checks
.18	Years job experience
.10	Years of education
Printed Tests	
.51	General Information
.43	Mechanical Principles
.40	Spatial Orientation
.33	Biographical Inventory
.48	Instrument Comprehension
Apparatus Tests	
.42	Discrimination Reaction Time
.41	Complex Coordination
.40	Rudder Control
.36	Two-hand Coordination
.31	Rotary Pursuit
.18	Finger Dexterity
0	**Random Prediction**
-0.01	Age

Source: Schmidt and Hunter (1998); Flanagan (as cited in Damos, 2011, p. 22).

Based on the above table, personality questionnaires (0.40), structured interviews (0.51), and ability tests (0.63) have good predictive validity which means they can be used widely in selection tests. Further to this, it is essential in the rapidly changing, technology-driven, competitive world to ensure that the most competent employees of the highest quality are found and selected in order to optimise their productivity (Davis, 2013). Prior to making selection decisions, employers must get a good understanding of applicants' potential to contribute to organisation's goals by analysing abilities, interests and characteristics that best fit employers' needs (Armstrong & Taylor, 2014). The key to achieving organisational success is the early identification of candidates with high potential. Making the right selection decisions can reduce training costs, improve job performance and enhance organisational effectiveness (Jain, 2014). To achieve these organisational goals, the needs of the organisation and the job applicant must be matched (Boon, den Hartog, Boselie & Paauwe, 2011).

2.6 LEGISLATION, ETHICS, FAIRNESS AND CULTURE

The Professional Board for Psychology's rules of conduct (HPCSA, 2013) pertaining to assessment activities include guidelines on assessment within a psychological context which differ from other contexts such as academic contexts like schools, training institutions or universities where the focus is largely on academic assessment. The appropriate use of assessment methods, informed consent in assessments, test development, cultural diversity, communication of results, information for professional uses, interpreting and explaining assessment results, test scoring and interpretation services, release of test data, obsolete and outdated test results, and maintaining test security are important factors that are covered

(Society for Industrial and Organisational Psychology of South Africa (SIOPSA), 2006). Codes of conduct for assessment are usually designed based on ethical principles that protect the needs of those being assessed (Barnard, 2010).

When dealing with differences in scores, it is important to accept that people and groups of people differ in terms of their make-up and abilities as a result of different backgrounds and experiences. Taking into account the requirements of the *Employment Equity Act 1998* (RSA), it is important to move past the bias and unfairness of the past and to ensure that fair, ethical, reliable, valid measurement tools free from adverse impact and unfair discrimination are used for selection decisions. Multiple selection methods such as interviews, simulations and written tests can be used and integrated into the final selection decision.

2.7 FACTORS INFLUENCING ASSESSMENT RESULTS

There are a few factors that influence assessments results. These factors are socio-economic status (SES), environment and education; they are briefly discussed next.

2.7.1 Socioeconomic Status

Socio-economic status (SES) is often measured as a combination of education, income and occupation. It is commonly conceptualised as the social standing or class of an individual or group. Low SES and its correlates, such as lower education, poverty and poor health, ultimately affect our society as a whole.

Research indicates that children from low-SES households and communities develop academic skills more slowly compared to children from higher SES groups

(Morgan, Farkas, Hillemeier, & Maczuga, 2009). Initial academic skills are linked to the home environment, where low literacy environments and chronic stress negatively affect a child's pre-academic skills. According to Beach (2013), the school systems in "low-SES communities are often under-resourced, negatively affecting students' academic progress" (p. 21) and achievement.

2.7.2 Environment

Communities are often segregated by socio-economic status, race, and ethnicity. These communities commonly share characteristics of developing nations: low economic development, poor health conditions, and low levels of educational attainment (Hysa, 2013). The quality of human settlement (housing type and construction, infrastructure and lifelines) is important for understanding social vulnerability and influence in these communities. In addition to the environment that people grow up in, language is generally accepted as the most important moderator of performance in psychological testing (Foxcroft, 2008) and is closely tied to culture. Entry requirements for pilot training include the ability to communicate in English (Department of Public Enterprises, 2012).

2.7.3 Education

In South Africa, quality of education varies based on the different schooling systems that are offered (public vs private/independent schools; universities vs FET colleges and technikons). This gives those who have had a good education a much greater advantage than those who have had a poorer education in terms of quality and curriculum offering. The following factors have been found to improve the quality of schools in SES neighbourhoods: a focus on improving teaching and learning,

creation of an information-rich environment, building of a learning community, continuous professional development, involvement of parents and increased funding and resources (Mujis, Harris, Chapman, Stoll, & Russ, 2009).

While some researchers feel that a battery of tests should be developed for those with a lower level of education and that they should be explained in many languages, McArdle and Woodcock (2014) explain that many tests available in South Africa measure crystallised abilities (*gc*) which are strongly influenced by culture and one's schooling where specific cultural groups have had a better standard of schooling and development opportunities. They also say that because of this, a need has arisen in South Africa not just to assess skill, but rather potential for development, despite the skills gap that may exist due to previous disadvantage.

2.8 CHAPTER SUMMARY

In this chapter a review of literature was provided focusing on selection as a process in general, and pilot selection and psychological assessments as part of pilot selection in the South African context. The importance of selection in the context of pilots with specific reference to psychometric assessments and cognitive assessments in particular was discussed. Legislation governing the use of these assessments in the multi-cultural SA was briefly reviewed in relation to issues of fairness, bias and culture in relation to selection. Social economic status factors that influence assessments were discussed. The literature review formed the basis for empirically testing the relationship between certain variables. The next chapter outlines the methodology of the study.

CHAPTER 3:

METHODOLOGY

3.1 INTRODUCTION

This chapter presents the methods that have been used in this study to investigate, describe and compare the performance on cognitive ability assessments for the Black and the White groups of pilots in a South African commercial airline. This chapter describes the research design, target population, sample, data collection, data analysis, validity and reliability of the study and ethical considerations.

3.2 RESEARCH DESIGN

Research design as described by Terre Blanche, Durrheim, and Painter (2006) is a strategic framework for action that serves as a bridge between research questions and the execution or implementation of the research. A quantitative, descriptive research design was chosen for this study in order to give a detailed description of the pilots' assessments results and the performance differences. Ezzy (2013) defines quantitative research as a formal, objective, systematic process to describe and test relationships and examine cause and effect interactions among variables. This method was selected because it is a flexible research design that provides an opportunity to examine all aspects of the problem being studied, it strives to develop new knowledge and the data may lead to suggestions of hypotheses for future studies (Babbie, 2013; Creswell, 2003; Terre Blanche et al., 2006). Quantitative researchers collect data in the form of numbers and use statistical data analysis.

3.3 RESEARCH SETTING

The study was conducted at a South African commercial airline which has head offices in Johannesburg. This airline has been operating successfully within Southern Africa for more than six decades with an internationally recognised safety record. The airline operates local and regional flights and employs an excess of 200 pilots (first officers and captains).

3.4 POPULATION AND SAMPLING

A population is the total of all the individuals who have certain characteristics that are of interest to a researcher (Babbie, 2013; Burns & Grove, 2003; Dillon, Madden, & Firtle, 2000). A common term for this is a population of interest or target population. From the target population, a researcher must generally identify an accessible population to which the conclusions of the study can be applied (Explorable.com, 2009). This population is a subset of the target population and is also known as the study population. The population for this study was the applicants for a commercial airline's First Officer position as they met the requirements for inclusion which were:

- Grade 12 or equivalent (Essential);
- English as a first language minimum symbol D or equivalent OR English second language minimum symbol C or equivalent (Essential);
- Must be a South African citizen (Essential);
- A valid South African Class 1 Medical Flying Certificate (Essential);
- A valid South African Commercial Pilots Licence (Essential);
- Instrument Rating (Essential);
- Flying hours – total of 1500 hours (Essential);
- One or both of the following: 200 hours on multi-engine aeroplanes (Essential) OR

multi-engine rating plus 200 hours on a technically sophisticated single engine pressurized turbine (Essential);

- A valid South African Airline Transport Pilots Licence (Desirable);
- Multi crew experience (Desirable).

3.4.1 Sample

Sampling is the process of selecting a group of subjects for a study in such a way that the individuals represent the larger group from which they were selected (Denscombe, 2014). Terre Blanche et al. (2006) describe sampling as a selection of research participants from an entire population, which involves decisions about which people, settings, events, behaviours, and/or social processes to observe.

Although a sample should be representative of the population of interest (Blair, Czaja, & Blair, 2013), it was not logistically possible to reach all the members of the target population in this study. Convenience sampling was therefore applied which means that the most readily available or most convenient group of subjects was used for the sample (Cohen, Manion, & Morrison, 2000). The parameters of generalisability in this sample are negligible and the study did not seek to generalise the findings to the wider population.

The sample selected, therefore, consisted of a convenience sample of the First Officer job applicants at one airline. This study included all shortlisted applicants, as per inclusion criteria indicated on section 3.4, who were selected to proceed to the next phase of the selection process between March and December 2013. The sample consisted of 598 pilots who met these inclusion criteria. The study excluded

all the First Officer job applicants that applied pre-March 2013 because of the difference in the selection battery that was used. This readily available sample was chosen in order to highlight potential disparities between Black and White applicants. Problems that are anticipated with the sample are that Blacks that are perceived to be from the previously disadvantaged backgrounds may not necessarily have been subjected to the disadvantaged schooling system. Nevertheless, based on the cost involved to get a Private Pilot Licence (PPL), it is assumed that the parents may have taken their children to less disadvantaged (multi-racial) schools because they could afford to do so; therefore, not every black African pilot job applicant has been subjected to previously disadvantaged schooling system. Because the sample utilised is a convenience one, it may not be a precise replica of the universe and caution should be exercised in making generalisations to the entire population of candidate airline pilots in South Africa.

3.5 DATA COLLECTION

3.5.1 Data Collection Instruments

There are many cognitive ability tests available in South Africa and elsewhere which are continually being developed and redeveloped (Frederiksen, Mislevy, & Bejar, 2012). For example, the CAT3 test developed by GL Assessments, the UK's most widely used test of reasoning abilities, which now also has test batteries for non-verbal and spatial reasoning (GL Assessments, 2015). Gowda (2010) mentioned the redevelopment of Raven's SPM tests since they were first introduced in 1938. Cognitive ability tests measure a candidate's aptitude or potential to solve job-related problems (Society for Industrial and Organizational Psychology, 2015). Most of these tests are classified as "crystallised intelligence" (*gc*) measures (McArdle &

Woodcock, 2014, p. 31), the results of which can be strongly affected by schooling or culture.

For purposes of this study, existing psychometric assessment data of First Officer applicants were used. The data were obtained from the test provider and a South African commercial airline, with the permission of the organisational development department of the airline. The pilot selection battery used by the airline consists of 11 cognitive ability scales (CAS) tests and a Shapes Management personality test which is used to determine a talent match to the airline's specific pilot competency profile.

The instruments that measure cognitive ability are verbal reasoning test, numerical reasoning test, visual thinking, abstract reasoning test, spatial reasoning test, short term memory test, monitoring ability test, hand-eye coordination, sense of orientation test, reactivity test, and a multi-tasking test (with 3 subscales) (cut-e, 2010). All these tests were administered. They are adaptive, online assessments which were delivered using an unproctored method. Each of these scales tests is explained below.

3.5.1.1 Verbal and numerical reasoning test

3.5.1.1.1 Aim and description

The verbal reasoning test measures the ability to draw logical conclusions from complex verbal information. Additionally, this test measures the ability to retrieve relevant information when confronted with different type of questions. The verbal reasoning test is the most commonly used psychometric test (or aptitude test) to measure the suitability and potential of job applicants (The Institute of Psychometric

Coaching, n.d.). In the current study, a complex test was used which included reading passages to measure the applicant's ability to understand verbal information, think logically about written information, draw logical, well-founded conclusions, produce written reports, and convey information to others simply and clearly.

The numerical test measures the ability to draw logical conclusions from complex numerical information that is presented in tables and charts. It is not a test of mathematical ability. Additionally, this test measures the ability to retrieve relevant information when confronted with different type of questions. There is also correlation between verbal, numerical and other scales tests including abstract reasoning and concentration test (The Institute of Psychometric Coaching, n.d.)

3.5.1.1.2 Development of verbal and numerical reasoning tests

According to Sternberg (2000), the earliest tests were devised by the Heim team, although these have their foundation in the work of Galton in the late 1800s and Binet in 1910. The Alice Heim (AH) series of tests is a well-known and widely-used set of tests designed in the 1960s and 1970s to assess reasoning ability in both the graduate (AH5 and AH6) and general (AH2, AH3 and AH4) population (Psychometrics Ltd., n. d.). These are also labelled as Graduate General Reasoning Test 1 (GRT1) and General Reasoning Test 2 (GRT2). GRT1 was designed to assess applicants for graduate, managerial and professional roles, and other jobs that require an above average level of reasoning ability, and tests fluid intelligence (Psychometrics Ltd., n. d.). This is applicable in the current study.

Verbal and numerical tests differ in terms of content and context but not in terms of construction, concept and structure. During development of test items, importance is attached to the view that no classical distractors should exist but that the items have

to be answered on the basis of the given information. The response alternatives, which are identical for all items, are therefore “true”, “false” and “cannot say”. The response alternative “undefined” means that the statement cannot be answered without further information. Each administration of the numerical tests consists of 37 items, each administration of the verbal tests consists of 49 items, with each item containing 6 charts presenting different information.

3.5.1.2 Spatial reasoning test (spr)

3.5.1.2.1 Aim and description

There are two commonly used spatial reasoning tests: the Mental Rotation Test which was originally developed by Shepard and Metzler in 1972 (Ashrafi, 2012), and the Embedded Figures Test developed by Witkin (1969). The Mental Rotation Test consists of geometric shapes with labelled areas. The task is to create a pattern which, if folded, forms a shown geometric shape. A shape can be rotated interactively. The Embedded Figures Test requires “locating a simple shape embedded within a background of overlapping target-irrelevant scene elements” (Almeida, Dickinson, Maybery, Badcock, & Badcock, 2010, p. 9).

These tests are important for pilots because having spatial awareness is the ability to understand and remember the spatial relations among objects and the ability to mentally re-arrange objects without physically touching them. It is essential for solving everyday problems such as how to use a map to guide you, playing 3D games, or predicting where the ball is going to land during ball games. Sports such as basketball, football soccer, tennis involve spatial reasoning.

3.5.1.2.2 Development and administration

The task is to create a pattern which, if folded, forms a shown geometric shape. Different markers are given that have to be used for completion of the template. The candidate has to choose one marker for each free field. In a *spr* administration, each participant has 12 tasks from 8 different and increasingly difficult categories and has 14 different unwind patterns which represent patterns for the relevant geometric field. Furthermore, *spr* has 10 different graphic symbols with each symbol available in 4 orientations (rotated 90°). These symbols are the response alternatives and markers for the free fields of the template.

3.5.1.3 Short term memory test (stm)

3.5.1.3.1. Aim and description

This test measures the ability to make exact observations and the capacity of the short term memory. A picture or pictures are initially presented, and in several subsequent sections the participant sees a number of different pictures. The task is to memorise these pictures and to mark those pictures which are identical to the ones seen before. It is suitable for potential analysis and screening for all jobs, which require people to make exact observations, as is expected of a pilot.

3.5.1.3.2. Development and administration

Each *stm* administration consists of 10 items. Each item consists of a presentation phase and a response phase. After each item the participant has a break of exactly 8 seconds before the next presentation phase starts. During the presentation phase different objects are shown for exactly one second at randomly chosen screen positions. Each object of the presentation phase is shown exactly three times; e.g. If

five different objects are to be presented then these 5 objects are shown in a random order. This order is repeated three times with randomly chosen screen positions. During the response phase 8 objects are always shown. The participant has to decide whether the objects were part of the presentation phase or not. In each response phase, 4 of the 8 objects had been presented before.

3.5.1.4 Complex control test (*xw*)

3.5.1.4.1. Aim and description

The test measures the hand-eye coordination in a complex control task. The task is to avoid collisions in a tube whilst flying through it and encountering rotating obstacles with openings in the tube. It is suitable for all professions which require good hand-eye coordination such as pilots and machine operators.

3.5.1.4.2. Development and administration

The task of a participant who is completing the *scales xw* test is to manoeuvre with the mouse device through a tube, passing objects successfully without colliding with them. The operation is assisted by a small-sized indication on the top left side of the screen in which a little black dot shows the position within the tube. The test thereafter generates the obstacles at run-time, so that every ‘flight’ in the test is unique. The obstacles are positioned at regular intervals within the tube and can stagnate as well as rotate in different speeds and directions. During the administration, the plane accelerates continually. The participant can intervene insofar as he can accelerate or decelerate the flight with the aid of two keys. Acceleration therefore describes a linear function, whereas deceleration describes

an exponential function. There are 6 different types of obstacles, which appear repeatedly during the administration of the test.

3.5.1.5 Visual thinking (spc)

3.5.1.5.1 Aim and description

The test measures visual thinking and spatial ability. The task is to find mistakes in the copy of a pattern, which can be rotated, mirrored or rotated and mirrored. It is suitable for professions which require visual thinking, spatial ability or visual comparison of pattern.

3.5.1.5.2 Development and administration

Since the task is to find mistakes in the copy of a pattern, 10 geometrical figures are available which can be presented in 4 different directions each, so that 40 (10*4) different figures are provided. Every *spc* administration includes 6 tasks as well as four different difficulty levels.

3.5.1.6 Sense of orientation test (ndb)

3.5.1.6.1 Aim and description

This test measures the sense of orientation. The task is to specify the position and course of a plane relative to a non-directional beacon with the aid of a gyrocompass and a radio compass. It is suitable for all activities which require a good sense of orientation such as those carried out by pilots and air traffic controllers.

3.5.1.6.2 Development and administration

The participant's task is to specify the position and course of a plane relative to a non-directional beacon with the aid of a gyrocompass and a radio compass. For the gyrocompass, 4 different indications of course are available: north, east, south and west. The radio compass provides 8 possible indications to specify the position of the non-directional beacon: ahead, right ahead, right, right behind, behind, left behind, left, left ahead. The participant has to specify the direction in which the plane is flying with the aid of the indications and place it on a matrix relative to the non-directional beacon. This test can generate in each case 32 (4*8; 4 indications of the gyrocompass, 8 indications of the radio compass) different items within the same difficulty level at run-time.

3.5.1.7 Monitoring ability (cmo)

3.5.1.7.1 Aim and description

The task in these tests is to quickly identify a number of moving objects. It is suitable for professions which require high levels of monitoring ability as well as high cognition and perception speed such as controlling, safety inspectors, pilots and air traffic controllers.

3.5.1.7.2. Development and administration

The participant's task is to identify quickly the number of moving objects on the screen and to click on the respective number below. Therefore, 10 different categories with respective difficulty levels are available. The instrument starts with the first category and adapts its difficulty to the candidate during the course of the test. This means that the test becomes more difficult if the participant answers

correctly and easier if the participant answers incorrectly. This means that every item within a category is generated randomly at run-time.

3.5.1.8 Reactivity test (rt)

3.5.1.8.1 Aim and description

This test measures the participant´s reactivity. The task is to react as fast as possible when two equal objects appear on the screen. It requires fast reactions, high concentration and attentiveness. It is also suitable to determine general processing capacity and general cognitive performance.

3.5.1.8.2 Development and administration

In this test, the presented objects are generated randomly at run-time. The pool includes five different objects (circle, cross, star, square and triangle). For every sequence *scales rt* provides 5*5 symbols. It can be assumed that at an average of 120 pairs of objects are shown per administration.

3.5.1.9 Discovering rules/abstract reasoning (ix)

3.5.1.9.1 Aim and description

This test measures the ability of logical thinking. The task is to find out the object that does not match the rule. It is also suitable for general diagnostics of inductive reasoning.

3.5.1.9.2 Development and administration

The task is to detect the object corporate rule and find out the object that does not match this rule. Every series includes 9 objects. The test generates for every item a series of objects and randomly substitutes one of them with an incorrect object. The

participant has to identify the particular incorrect object and mark it; the test has different difficulty levels.

3.5.1.10 Multi-tasking test (mt)

3.5.1.10.1 Aim and description

This test measures the ability to perform several tasks simultaneously. The measurement takes place in subsequent segments, so that the gradient of the multi-tasking performance can be measured too. The test has three subscales called catching, calculating and checking.

3.5.1.10.2 Development and administration

This test incorporates the gradient of multi-tasking performance and each item is generated by an item generator at run-time. The administration requires that a participant to complete three tasks at the same time. The test has a task called “catching” which requires a participant to catch as many balls as possible with a basket. For catching, the basket can be moved with the mouse horizontally from the left to the right side of the task frame. The other task is called “calculating” and it requires deciding quickly if a given equation is correct or not. After providing the answer by clicking on the respective button, the equation disappears and after 5 seconds a new equation is shown. The progression of the 5 seconds is indicated by a progression bar. The third and the last task is called “checking” and it requires checking within 10 seconds if any character appears twice within a string of 7 characters. If a character appears twice, the okay-button be clicked and after exactly 10 seconds a new string is shown. This test is important for pilots because it

assesses the ability to perform more than one task at a given time (Cut-e, 2010; McDonald, 2011).

3.5.2 Psychometric Properties of the Tests

There are three psychometric properties that need to be considered in cognitive ability tests. These are standardisation, reliability and validity and they are briefly discussed next.

3.5.2.1 Standardisation

Standardisation means that tests are standardised against populations similar to those for which they were designed (e.g. people in technical, managerial, professional and scientific roles). Standardisation ensures that the scores obtained on these tests can be interpreted by relating them to a relevant distribution of scores. According to cut-e (2015a), the tests they administer are "scientifically sound – ensured through continuous standardisation and validation studies in cooperation with companies and universities" (n. p.). Standardisation is measured using the split-half test which is measured by using the Spearman-Brown prophecy coefficient. A common rule of thumb for adequate reliability is a coefficient of .80 and for good reliability.90 or higher (Garson, 2009). Struwig and Stead (2004) maintain the score shoud be higher than 0.85.

3.5.2.2 Reliability

Reliability is a measure of the consistency with which the measuring instrument measures. It deals with errors of measurement, and the error component is random (Moerdyk, 2009). It is assessed against the principles of temporal stability and

internal consistency. Temporal stability relies on a statistical principle called test-retest reliability which basically says that the results of tests with a group of testees should be stable (remain the same) when the tests are repeated even though they may be taken at different times. Should the results differ, then the test can be said to be unreliable. Internal consistency implies that if testees score well on one part of a test, they should also score well on other parts of the test (Gregory, 2004).

3.5.2.3 Validity

While reliability is important, it may affect validity which defines the meanings of the test scores. If the instrument in unreliable, it is not possible for it to be valid. Validity is the most fundamental testing and selection issue. The concept of validity was formulated by Kelley (1927, p. 14) who stated that 'a test is valid if it measures what it claims to measure'. It should be noted that a reliable test is not necessarily a valid test, which is why reliability is normally assessed before validity is evaluated. A test's validity is the extent to which scientifically valuable or practically useful inferences can be drawn from the scores (Carreta, 2000). Much effort is made to develop selection methods based on theories of the relations between personal characteristics and performance but they are worthless without theory proving validity. Therefore, validity is concerned with what is being measured by the test and how well it is being measured (Barnard, 2010; Foxcroft & Roodt, 2001). Validity is the extent to which the test is free of irrelevant or contaminating influences (Moerdyk, 2009).

The validity of a test is assessed against the principles of construct validity and criterion validity. Construct validity can be viewed as "an overarching term to assess

the validity of the measurement procedure" (e.g., a reasoning test) that is used to measure a given construct (e.g., cognitive ability) (Lund Research Limited, 2012, n. p.). Constructs are the building blocks of theories, helping to explain how and why certain phenomena behave the way that they do. In this thesis, the constructs are intelligence, aptitude and cognitive ability. Construct validity is measured by using correlation tests, such as Pearson's correlation coefficient. Criterion validity, according to Pallant (2005, p. 6) "concerns the relationship between scale scores and some specified measurable criterion". For example, a valid test would be expected to predict cognitive ability in a pilot applicant (Table 2.1 in Chapter 2 confirms that cognitive ability tests have good predictive power).

A further aspect of validity is external validity which refers to the extent to which study findings can be generalised beyond the sample used (Burns & Grove, 2003). Terre Blanche et al. (2006) further describe external validity as the extent to which it is possible to generalise from the data and context of the research study to the broader populations and settings. The researcher used readily available data for analysis and is aware of the threats associated with generalisability of the findings to participants. In order to counter this, the findings obtained need to be replicated with a different sample (Terre Blanche et al., 2006).

3.5.2.3 Evaluation of tests in terms of psychometric properties

Table 3.1 and 3.2 below shows that psychometric properties of the tests listed above have been researched and meet the required standards as stipulated above. The information was requested from the researchers at cut-e as a wide search of the internet did not reveal any studies. Only six out of 11 tests had statistical evidence

that was provided to substantiate the claims of the test developers. Cut-e confirmed that there are no validation studies for these tests: ix, ndb, spc and xw and that the organisation is in the process of producing them (R. Justenhoven, personal communication, January 12, 2016). The reliability studies of the six tests are illustrated in Table 3.1 below.

Table 3.1: Split-half reliability

Test Name	**N**	**Mean**	**SD**	**Reliability coefficient**
STM	317	58.9	6.4	0.93
SPR	211	13.0	5.6	0.81
Numerical Industry	314	15.24	4.31	0.81
Verbal Industry	314	14.31	3.95	0.83
Multitasking	1163	105.2	18.3	0.91
• Catching	1163	54.8	1.8	0.93
• Calculating	1163	18.6	3.0	0.82
• Checking	1163	21.1	2.4	0.84

Going by the parameters for reliability mentioned above, all the above cut-e test, the reliability co-efficient is acceptable at 0.80 and higher. Table 3.2 below shows the test-retest reliability evidence.

Table 3.2: Re-test Reliability

Test Name	**N**	**Mean**	**SD**	**Reliability coefficient**
STM	116	59.6	6.3	0.82
SPR	100	13.7	3.9	0.83
RT	100	496.83	55.23	0.74

Test Name	N	Mean	SD	Reliability coefficient
Numerical Industry	84	17.14	5.01	0.83
Verbal Industry	84	13.86	4.89	0.81
Multitasking	186	108.4	17.9	0.81
• Catching	186	56.5	1.8	0.83
• Calculating	186	19.2	3.0	0.74
• Checking	186	21.7	2.3	0.76

The following tests: STM, SPR, RT, numerical industry, verbal industry, multi-tasking overall performance and multi-tasking catching had acceptable reliability coefficient at 0.80. However, reactivity (RT), multi-tasking calculating and multi-tasking checking had reliability coefficient of less than the acceptable parameter of 0.80.

Since these tests recorded an acceptable level of reliability on split-half, it is not too concerning that the test re-test reliability was lower than 0.8. According to Shuttleworth (2009), even if a test-retest reliability process is applied with no sign of intervening factors, there will always be some degree of error, and therefore some methods such as split-half are better to use (n.p.). He further argued that perfection is impossible and some researchers accept a lower level of reliability at 0.7.

3.5.3 Scoring the tests

The scoring for all the tests follows the same process. The raw score is computed from the number of items completed correctly compared to the number of items attempted within a specified time. Each test is generated by an item generator at run-time which prevents faking (cut-e, 2010). This prevention is good when running unsupervised online administrations. Each test had its own instructions and provides

an opportunity for practice before starting with the completion of the timed test, with different time limits.

3.6 DATA ANALYSIS

The purpose of the data analysis is to test the hypothesis and the general aims of the research, as set out in Chapter 1. Data analysis is the systematic organisation and synthesis of the research data and the testing of research hypotheses using those data. It also entails categorising, ordering, manipulating and summarising the data and describing them in meaningful terms (Brink, 1996). For analysis, a statistical computer programme called Statistica was used.

3.6.1 Descriptive Statistics

The purpose of descriptive statistical analysis is to provide an overview of the data that were collected for the research. It helps to summarise data in a meaningful way such that patterns might emerge from the data (Mouton, 1996). It enables presentation of data in a more meaningful way, which allows simpler interpretation of data.

Since descriptive statistics are not developed on the basis of probability theory, different measures are used to describe descriptive statistics. The data is described through statistics and graphs using measures of central tendency (mean, median and mode), measures of variability or dispersion and spread (range, quartiles, absolute deviation, variance and standard deviation). Frequency tables were drawn from these and the data were presented in tables.

3.6.1.1 Measures of central tendency

These are ways of describing the central position of a frequency distribution for a group of data. This central position is described using a number of statistics, including the mode, median, and mean. Mean and median are used in this study.

- Mean: The goal of central tendency is to find the best representative value of distribution. The mean has a good advantage of being closely related to variance and standard deviation. The mean is a valuable measure for the purposes of inferential statistics, which will be discussed in the next section (Yount, 2006).
- Median: When a distribution has a few extreme scores, i.e. those that are very different in value from most of the others, the mean may not provide a good representation of distribution (Gravetter & Wallnau, 2013). Extreme values can cause the mean to be displaced. In these cases, it becomes suitable to use the median.

3.6.1.2 Measures of spread

These are ways of summarising a group of data by describing how spread out the scores are (Buglear, 2012). To describe this spread, a number of statistics are used including the range, quartiles, absolute deviation, variance and standard deviation. Range and standard deviation are used in this study.

- Range: The range is the distance between the largest score and the smallest score in the distribution. It is the difference between upper real limit and lower real limit.

- Standard Deviation: Standard deviation (σ) is the most commonly used and the most important measure of variability. Standard deviation uses the mean of the distribution as a reference point and measures the distance of a score from the mean (Laerd, 2013).

3.6.2 Inferential Statistics

Inferential statistics are used to determine the extent to which the results can be generalised to the population from which the sample was drawn. Inferential statistics arise out of the fact that sampling naturally incurs sampling error and thus a sample is not expected to perfectly represent the population. Since certain inferences were made from the data, a number of statistical methods were used.

- To test the hypotheses; *p* values were used.
- To test the statistical significance of the difference between the means of more than two groups, t-test or ANOVA (Analysis of Variance) was used (Creswell, 2003).

3.6.3 Research Hypotheses

The following research hypotheses were tested. H_1 and H_2 are the null hypotheses and H_{1A}, H_{2A} are the alternative hypotheses.

H_1 – There is no significant difference in test scores (performance) of black and white group of applicants.

H_{1A} – There is a significant difference in test scores (performance) of black and white group applicants.

H_2 – There is no significant difference in test scores (performance) of English home language (speaking) and English second language group of applicants.

H_{2A} – There is a significant difference in test scores (performance) of English home language (speaking) and English second language group of applicants.

3.7 ETHICAL CONSIDERATIONS

Conducting research not only requires expertise and diligence, but also honesty and integrity in order to protect the rights of humans. According to Terre Blanche et al. (2006), some studies are exempt from ethical review if they do not involve human participants and/or are based on the information that is already in the public domain. In this research, the identity of the participants is not exposed. Existing assessment data were used and were identified by using unique participant identifier, age, gender and race as moderators. It is, however, still crucial that all researchers are aware of research ethics. The measures discussed next were taken to ensure that the study was conducted without any potential risks.

3.7.1 Confidentiality

Confidentiality is very important when it comes to the use of human data. Anonymity is one way in which confidentiality is maintained. A research project guarantees anonymity when the researcher can identify the responses given by an individual but essentially promises not to do so publicly (Terre Blanche et al., 2006).

3.7.2 Honesty

Honesty, as described by McNamara (1994), means accurately reporting both the methods and the results of the surveys to professional colleagues in the academic community. Because advancements in academic fields come through honesty and openness, the researcher assumed the responsibility to report problems and weaknesses experienced as well as the positive results of the study. Manipulation of the data was avoided as the services of an independent statistician were used in analysing the data. The researcher did not falsify or misrepresent any data or information presented in this study.

3.8 CHAPTER SUMMARY

This researcher used a quantitative, comparative design. Existing assessments results were used from a convenient sample of subjects. The sample characteristics included job applicants for a first officer position at an airline. The chapter described the research methodology, including population, sample, data collection, data analysis (using both descriptive and inferential statistics), reliability and validity of the study and ethical considerations. The analysis of the data and the findings are discussed in the next chapter.

CHAPTER 4:

RESEARCH FINDINGS

4.1 INTRODUCTION

The researcher conducted quantitative, descriptive research to investigate various factors related to the performance of pilots on cognitive abilities at a South African Airline. This chapter describes the analysis of data followed by a discussion of the research findings. The first section describes the information derived from analysis of each variable through descriptive statistics and the second section is the presentation of the results using inferential statistics. The findings relate to the research objectives that guided the study.

4.2 DATA ANALYSIS

4.2.1 Data Analysis Strategies

Prior to undertaking the quantitative analysis, the following was done to prepare the data for analysis:

- Coding – data were coded and prepared for analysis using the statistical programme;
- Cleaning and screening data – data were cleaned to ensure intelligibility. Data that could not be used were deleted.

4.2.2 Presentation of Data

The quantitative data of the sample were made available from the database of the airline's results and were captured into an Excel spreadsheet before being analysed

by the Unit for Statistical Consultations at NMMU and TTS. The analysis of the data is presented as follows:

- Sample characteristics of the participants;
- Descriptive statistical analysis including the results of central tendency t-tests and degrees of freedom and the frequency to which the applicants have completed each of the ability tests;
- T-tests of the sample on each test by race, gender and language; and
- Data interpretation.

4.3 SAMPLE CHARACTERISTICS

The sample was made up of 598 individuals that applied to be First Officers at a South African commercial airline. Each participant was given 11 online assessments to complete; however, not all participants completed all the tests but a maximum sample was used for each test. This means that the sample size differed for each criterion. The demographic details of the sample according to ethnic origin, gender and ages are shown in the next table:

Table 4.1: Biographical details of the sample (n = 598)

Variable	N	Percentage (%)
Gender		
Male	547	91.5%
Female	51	8.5%
Total	598	100%
Age		
20-29	293	49.2%
30-39	239	40.1%

Variable	N	Percentage (%)
40-49	57	9.6%
50-59	6	1.0%
60-69	1	0.2%
Total	596	100%
Race		
Black	29	4.8%
White	564	94.3%
Other/Not provided	5	0.8%
Total	598	100%
Ethnic group within race		
Other/Not Provided	5	1%
African	20	3%
Coloured	4	1%
Indian	5	1%
White	564	94%
Total	**598**	**100%**
English Language		
Not Provided	1	0.2%
Home Language	328	54.8%
Second language	269	45.0%
Total	**598**	**100%**
Other languages		
Other/Not Provided	9	1.5%
Afrikaans	250	41.8%
English	327	54.7%
South Sotho	2	0.3%
Tswana	3	0.5%
Venda	2	0.3%
Xhosa	4	0.7%
Zulu	1	0.2%
Total	**598**	**100%**

Variable	N	Percentage (%)
Education		
Not Provided	8	1.3%
Grade 10 or below*	3	0.5%
Grade 12	442	73.9%
Grades 10 and 11	6	1.0%
B Degree or Diploma	103	17.2%
Post Graduate Qualification	36	6.0%
Total	**598**	**100%**

*Although Grade 12 is required for entry, these 3 testees indicated that they had a qualification at Grade 10 or below

4.3.1 Gender

The sample consisted of 51 females (8.5%) and 547 males (91.5%). It is apparent that the majority of the sample was male. This is as a result of the historical and, to a large extent current, demographics of the world-wide commercial airline pilots' profession which was previously seen as a male job. It currently remains male-dominated.

4.3.2 Age

A majority of sample were in the age group 20 – 29 with 49.2% followed by the age group of 30 – 39 with 40.1%. The retirement age for a commercial airline pilot as stipulated by the Federal Aviation Authority (2012) and adopted internationally is 65. It is therefore not surprising that there are only 7 (1.2%) participants who fall in the 50 – 59 and 60 – 69 age categories. The dispersion of these ages is illustrated below.

Table 4.2. Central tendency and dispersion: Age (n = 596)

Mean	S.D.	Minimum	Quartile 1	Median	Quartile 3	Maximum
31.18	6.20	20.00	27.00	30.00	34.00	61.00

The youngest participant was 20 years old and the oldest was 61 years old with the mean for these ages being 31.18.

4.3.3 Race

The white group had a majority sample representation of 94.3%, while the black group had a representation of 5%. The “other” category constitutes 0.5% and it has been excluded from the analysis together with the ‘not provided’ group which is 0.3%. In the black group category, there were 20 Africans, 4 Coloureds and 5 Indians. For the purposes of this study, Coloureds, Africans and Indians were included in the black group. The White group’s majority could possibly be attributed to being advantaged and that the participants could afford to pay for the costs associated with training to become a pilot. Conversely, the Black group’s minority could be attributed to being previously disadvantaged and unable to afford the costs associated with training to become a pilot.

4.3.4 Language

The sample had 328 (54.8%) participants who had English as a home language and 268 (45%) participants who had English as a second Language. Of the 268 participants that indicated that English as their second language 250 were Afrikaans, 2 were South Sotho, 3 were Tswana, 2 were Venda, 4 were Xhosa, 1 was Zulu. Seven 7 participants did not provide their home languages.

4.3.5 Education

A total of 442 participants had Grade 12 education, 102 participants had a B Degree or Diploma qualification, 36 participants had a post-graduate qualification, 6 participants had post-grade 10 (i.e. grade 11 and 12), 3 participants had Grade 10 or below and 8 participants did not provide their level of education. The majority of the participants (74%) pursued flying as a career after grade 12.

4.4 DATA PROCESSING

4.4.1 Statistical data processing

The assessment data and the biographical information were analysed. The tests that were utilised were t-tests, degree of freedom and Cohen's d.

4.4.2 Descriptive statistics

The descriptive statistics of the sample were obtained in order to understand their nature better. These statistics seek to describe the data in terms of minimum and maximum values, mean, standard deviation for each test and are described in Table 4.3.

Table 4.3: Descriptive statistics per test – central tendency and dispersion

Test Code & Name	n	Mean	S.D.	Minimum	Quartile 1	Median	Quartile 3	Maximum
306 STM	594	63.38	10.76	-64.00	58.00	64.00	70.00	80.00
310 SPR	587	9.37	4.41	0.00	6.00	8.00	12.00	24.00
321 Numerical reasoning industry	586	8.34	4.56	-3.00	5.00	8.00	11.00	27.50
322 Verbal Reasoning industry	577	8.92	4.20	-3.00	6.00	9.00	11.50	22.50
325 SPC visual thinking	583	16.60	7.60	-6.00	11.00	16.00	21.00	39.00
326 CMO perceptual speed (Monitoring ability)	582	128.53	53.16	36.00	93.00	120.00	150.00	369.00
327 IX discover rules	577	114.72	28.67	20.00	96.00	116.00	132.00	191.00
329 XW complex	580	61.38	42.79	-63.00	35.00	64.00	92.00	182.00
330 NDB spatial orientation	579	58.29	45.17	-22.00	23.50	49.00	84.00	223.00
333 MT Calculating	577	19.22	5.25	-2.00	16.00	20.00	23.00	31.00
333 MT Catching	577	57.13	21.00	-96.00	55.00	55.00	64.00	102.00
333 MT Checking	577	19.19	5.45	-18.00	16.00	20.00	23.00	32.00
365 TR2 Reactivity	574	424.86	81.12	-579.38	398.63	437.80	469.02	574.65

From the above, the test with the highest completion rate is short term memory followed by spatial reasoning and numerical reasoning. The test with the lowest completion rate is reactivity followed by multi-tasking and verbal reasoning test.

4.5 DATA ANALYSIS

4.5.1 T-test

A t-test is utilised to analyse the data obtained from two different groups in order to scrutinise whether the group mean difference score is so large that it cannot reasonably be the result of chance (Christensen, 1997). It is most commonly applied when the test statistic would follow a normal distribution if the value of a scaling term in the test statistic were known. The t-test was used to investigate the mean difference scores of the White and Black groups regarding performance on the basis of race and home language.

4.5.1.1 Effect sizes

The effect sizes of the t-test were interpreted. According to Cohen (1992), the accepted benchmark for effect size is 0.20 for minimal solutions which are significant in social science research, 0.50 is for medium effect and 0.80 and above is regarded to be a large effect size. Cohen (1988) defined the effect size as the extent to which the phenomenon is found within the population or, in the context of statistical significance testing, the degree to which H_0 is false. Snyder and Lawson (1993) argued that the effect size indicates the extent to which the dependent variable can be controlled, predicted and explained by the independent variable (s). On the use of the benchmarks as guidelines, Cohen (1988) stated that “there is a certain risk inherent in offering conventional operational definitions for those terms for use in

power analysis in as diverse a field of inquiry as behavioural science" (p. 25). Table 4.4 provides the results indicating the significance of the differences.

4.5.1.2 Level of statistical significance

Two types of errors can be made when testing a null hypothesis (Good, 2013). The first one is rejecting the null hypothesis when it should be accepted (type-1 error) and the second one is accepting the null hypothesis when it should be rejected (type-2 error). Statistical significance is calculated by determining if the probability differences between sets of data occurred by chance or if they were a result of the experimental treatment. In testing the hypotheses, *p*-values were used.

As explained by Bruton, Conway and Holgate (2000), *p*-value means the probability of obtaining a result equal to or more extreme than what was actually observed. The *p*-value was first introduced by Karl Pearson in his Pearson's chi-squared test. The smaller the *p*-value, the larger the significance because it tells the investigator that the hypothesis under consideration may not adequately explain the observation. Most researchers use levels of significance of 0.05 and 0.01 (Fenton & Neil, 2012). A 0.05 level of significance was chosen for this research project. If the obtained *p*-value is smaller than 0.05, the null hypothesis is rejected.

Statistical significance does not necessarily mean real significance; if sample size is large, even very small differences can have a low *p*-value. Conversely, lack of significance does not necessarily mean that the null hypothesis is true because if sample size is small, there could be a real difference, but we are not able to detect it (Schneider, 2013).

Table 4.4: T-tests: Scores by ethnic group

Variable	Ethnic Group	n	Mean	S.D	Difference	t	d.f.	*p*	Cohen's d
306 STM	Black White	29 561	58.21 63.61	12.45 10.63	-5.40	-2.65	588	**.008**	**0.50 Medium**
310 SPR	Black White	25 557	6.28 9.49	3.25 4.39	-3.21	-3.62	580	**<.0005**	**0.74 Medium**
321 Numerical reasoning industry	Black White	28 554	6.36 8.48	4.64 4.52	-2.12	-2.42	580	**.016**	**0.47 Small**
322 Verbal Reasoning industry	Black White	25 548	8.66 8.97	4.45 4.17	-0.31	-0.36	571	.721	n/a
325 SPC visual thinking	Black White	28 551	10.89 16.90	6.20 7.56	-6.01	-4.14	577	**<.0005**	**0.80 Large**
326 CMO perceptual speed (Monitoring ability)	Black White	28 550	116.61 129.26	46.63 53.47	-12.66	-1.23	576	.220	n/a
327 IX discover rules	Black White	27 546	101.81 115.41	40.45 27.95	-13.59	-2.41	571	**.016**	**0.47 Small**
329 XW complex	Black White	27 549	21.52 63.29	35.30 42.33	-41.77	-5.04	574	**<.0005**	**0.99 Large**
330 NDB spatial orientation	Black White	26 549	28.31 59.68	30.16 45.19	-31.38	-3.50	573	**<.0005**	**0.70 Medium**
333 MT Calculating	Black White	26 547	17.23 19.32	5.50 5.23	-2.09	-1.98	571	**.048**	**0.40 Small**
333 MT Catching	Black White	26 547	44.04 57.79	36.04 19.91	-13.75	-3.28	571	**.001**	**0.66 Medium**
333 MT Checking	Black White	26 547	17.08 19.30	8.01 5.28	-2.22	-2.03	571	**.042**	**0.41 Small**
365 TR2 Reactivity	Black White	27 543	361.26 427.91	124.64 77.34	-66.64	-4.22	568	**<.0005**	**0.83 Large**

Statistically significant differences are seen in a number of tests and the difference is statistically highly significant at $p<0.05$ and lower.

4.5.3 Analysis of Variance (ANOVA)

ANOVA is a general technique that can be used to test the hypothesis that the means among two or more groups are equal, under the assumption that the sampled populations are normally distributed (Pallant, 2013). ANOVA was used to describe possible the difference between the two groups in terms of the English language, assuming that Black candidates generally use English as a second language. Two-way ANOVA was the preferred method because there was one independent variable independent (assessment results) and three dependent variables (gender/race and language). The results of the performance by English language in Table 4.5.

Table 4.5: T-tests: Scores by English language

Variable	English Level	n	Mean	S.D	Difference	t	d.f.	p	Cohen's d
306 STM	Home Language 2nd language	325 268	63.19 63.60	11.76 9.44	-0.41	-0.46	591	.644	n/a
310 SPR	Home Language 2nd language	320 266	9.34 9.38	4.49 4.32	-0.04	-0.10	584	.923	n/a
321 Numerical reasoning industry	Home Language 2nd language	318 267	8.90 7.66	4.66 4.37	1.24	3.31	583	**.001**	**0.27 Small**
322 Verbal Reasoning industry	Home Language 2nd language	315 261	9.39 8.37	4.12 4.23	1.01	2.90	574	**.004**	**0.24 Small**
325 SPC visual thinking	Home Language 2nd language	317 265	16.32 16.92	7.33 7.92	-0.61	-0.96	580	.337	n/a
326 CMO perceptual speed (Monitoring ability)	Home Language 2nd language	316 265	128.71 128.17	54.50 51.65	0.55	0.12	579	.902	n/a
327 IX discover rules	Home Language 2nd language	312 264	114.04 115.56	28.34 29.14	-1.52	-0.63	574	.526	n/a
329 XW complex	Home Language 2nd language	316 263	62.66 59.86	41.97 43.85	2.80	0.78	577	.434	n/a
330 NDB spatial orientation	Home Language 2nd language	317 261	60.29 55.48	45.34 44.55	4.80	1.28	576	.202	n/a
333 MT Calculating	Home Language 2nd language	314 262	19.32 19.10	5.33 5.17	0.22	0.51	574	.612	n/a
333 MT Catching	Home Language 2nd language	314 262	57.38 56.85	20.26 21.92	0.52	0.30	574	.766	n/a
333 MT Checking	Home Language 2nd language	314 262	19.26 19.09	5.74 5.09	0.17	0.36	574	.716	n/a
365 TR2 Reactivity	Home Language 2nd language	311 262	424.03 425.81	87.39 73.30	-1.78	-0.26	571	.794	n/a

4.6 INTERPRETATION OF THE RESULTS

4.6.1 Hypothesis testing regarding differences in group performance based on race

The means of the two race groups were explored to establish whether race has an influence on the test scores. For statistical purposes, the Black group consists of Africans, Indians and Coloureds and the second group is Whites.

In all the tests, the mean for the White group was higher than that of the Black group. This suggests that there is sufficient evidence to infer that the alternative hypothesis is true and that there are differences in group performance on test results between the Black and the White group. There is no test where the Black group performed better than the White group. As mentioned above, there are only 2 tests where there were no statistically significant differences. Testing of the group differences on race was done in terms of the following hypothesis:

H_1: $p = 0$ indicating that there is no significant difference in test scores (performance) of Black and White groups of applicants.
H_{1A}: $p \neq 0$ indicating that there is a significant difference in test scores (performance) of Black and White groups of applicants.

There is, therefore, strong evidence to infer that there is a statistically significant difference between the mean scores of 11 out of 13 (multi-tasking with 3 subscales) assessments (as discussed above) of the Black and the White (race) groups. This constitutes 85% of the assessments written. The null hypothesis is therefore rejected.

4.6.2 Hypothesis testing regarding differences in group performance based on language

The group was further divided according to those who have English as a home language and those who have English as a second language. The English home language numerical reasoning industry tests consisted of 318 applicants and the second language group consisted of 267 applicants. The English home language for the verbal reasoning industry tests consisted of 315 applicants and the second language group consisted of 261 applicants. On these two tests, English home language speakers performed better than the English second language speakers as indicated in Table 4.5 above.

On the numerical reasoning industry test, English home language speakers had a higher mean of 8.90 than the second language group at 7.66. The statistical significance of this difference is small at 0.27. On the verbal reasoning industry test, English home language speakers also had a relatively higher mean of 9.39 and females 8.37. The statistical significance of this difference is small at 0.24. No statistically significant differences were reported on the rest of the tests on the basis of English language level. Testing of the group difference on English level was done by means of the following hypothesis:

H2: p = 0 indicating that there is no significant difference in test scores (performance) of English speaking (home language) and English speaking (second language) group of applicants.

$H2_{A:}$ $p \neq 0$ indicating that the is a significant difference in test scores (performance) of English speaking (home language) and English speaking (second language) group of applicants

The English second language speakers' performance was lower than that of the home language speakers on both verbal and numerical reasoning industry tests. These are the only tests that require reading, understanding and interpretation of the English language in order to complete these tests. The development process of these tests is exactly the same as mentioned in section 3.5.1.1.2 of chapter 3. This result could be due to the fact that the English first language speakers understood and interpreted the tests better than the English second language speakers, hence they performed better. There is therefore strong evidence to conclude that there is a statistically significant difference in test scores (performance) of English speaking (home language) and English speaking (second language) group of applicants. The null hypothesis is confidently rejected.

4.7 CHAPTER SUMMARY

This chapter provided an analysis of the quantitative data obtained from evaluating the test results. The data has been arranged and categorised in order to be able to answer the objectives of the study. The results of the analysis indicate that statistically significant differences exist on the performance of Black and White group on cognitive ability scores of 11 out of 13 tests. The results further indicate that there is a statistically significant difference between the performance of the English first language speaking group and the English second language speaking group. The null hypotheses 1 and 2 were rejected. These results were interpreted to inform

conclusions and recommendations of the study which are discussed in the next, final chapter of the study.

CHAPTER 5:

CONCLUSIONS AND RECOMMENDATIONS

5.1 INTRODUCTION

The purpose of the study was to compare the differences that exist in the psychometric assessment (cognitive ability test) scores performance of Black and White pilots in South Africa. Chapter 1 dealt with the motivation and the background of the study. Chapter 2 dealt with literature review with emphasis on cognitive ability assessments as a greater part of selection process and a construct of intelligence. Chapter 3 discussed the research design and methodology. The empirical results that were obtained for the first officer applicants at a commercial airline in South Africa were presented in chapter 4. Finally, in this chapter, conclusions are drawn on the basis of the findings of the literature as well as the empirical studies. The limitations of this project are presented before making recommendations for future research.

The study set out to determine whether there were differences in performance on cognitive ability assessments for the Black and White South African groups of pilots. The objectives of the study set out in Chapter 1 were as follows:

- Define what constitutes cognitive ability;
- Describe the battery of cognitive ability tests used in the selection of candidate pilots;
- Understand the factors impacting achievement in cognitive ability tests;

- Compare the results of cognitive ability assessments between applicants from African groups (Blacks) and Caucasian (Whites) for jobs as pilots in the commercial airline;
- Determine the significance of the differences between Black and White pilots in results on cognitive ability assessments.

This chapter evaluates whether the study has met these objectives.

5.2 CONCLUSIONS FROM THE LITERATURE

Objective 1: Define what constitutes cognitive ability

As far as the literature review is concerned, the researcher experienced difficulty in pinpointing the exact differences between aptitude, ability and intelligence as there are huge overlaps between these concepts. The researcher found competing definitions. Sternberg (1987) stated that “viewed narrowly, there seems to be almost as many definitions of intelligence as there were experts asked to define it”.

Objective 2: Describe the battery of cognitive ability tests used in the selection of candidate pilots

In chapter 2, section 2.3, the recruitment and selection of pilots was discussed and it was found that there are four basic skills areas that are evaluated in the recruitment and selection process: academic, social, psychomotor and physical. The need for rigorous selection was illustrated by referring to the statistics on fatal aircraft accidents where it was determined that up to 95% of such accidents can be attributed to pilot error. Section 2.5 showed that cognitive ability tests have good predictive value in determining whether job applicants meet the requirements for

selection. This then led to a description of the range of cognitive ability tests used by the airline selected for study; these tests are provided by the “cut-e” organisation. It was not possible, despite an extensive search of the literature, to determine the psychometric properties of these tests with any real certainty, although the organisation makes claims that the tests are standardised, valid and reliable. The organisation has provided assessments for over 100 companies worldwide, including some of the world’s largest aviation companies, such as Lufthansa, Boeing, Airbus, Swissair and Aer Lingus, among others (cut-e, 2015b), and it is assumed that these companies would have rigorous selection requirements and would have evaluated the efficacy of the cut-e tests before using them.

Objective 3: Understand the factors impacting achievement in cognitive ability tests

It was determined that, amongst other things, socio-economic status, environment and quality of education have a decided impact on achievement in cognitive ability tests (Chapter 2, sections 2.7.1; 2.7.2 and 2.7.3). People with low SES tend to perform poorly on cognitive ability tests; in terms of environment, knowledge of the English language was seen to be the main moderating factor in achievement while quality of education, with culture as a moderating influence, also has an impact on achievement in cognitive ability tests.

5.3 CONCLUSIONS FROM THE EMPIRICAL INVESTIGATION

Objective 4: Compare the results of cognitive ability assessments between applicants from African groups (Blacks) and Caucasian (Whites) for jobs as pilots in the commercial airline.

A range of statistical tests was conducted using descriptive statistics of median and mean, Pearson's *p*-value and t-tests (ANOVA). It was comprehensively proven that there are differences in all but two of the tests, namely verbal reasoning and monitoring ability. The size of the differences on each test is summarised in Table 5.1 below.

Table 5.1: Summary of size differences

Size criterion	Test	Difference
No significant differences exist (d: <020)	Verbal reasoning industry test	-
	CMO perceptual speed (monitoring ability)	-
Statistically small differences exist (d: 0.20 – 0.49)	Multi-tasking calculating	d = 0.40
	Multi-tasking checking	d = 0.41
	Discover rules	d = 0.47
	Numerical reasoning industry test	d = 0.47
Statistically medium differences exist (d: 0.50 – 0.79)	Short term memory (STM)	d = 0.50
	Multi-tasking catching	d = 0.66
	Spatial orientation (NDB)	d = 0.70
	Spatial reasoning (SPR)	d = 0.74
Statistically large differences exist (d: 0.80+)	Visual thinking (SPC)	d = 0.80
	Reactivity (TR2)	d = 0.83
	Complex control (XW)	d = 0.99

Objective 5: Determine the significance of the differences between Black and White pilots in results on cognitive ability assessments

Table 5.1 (extracted from Table 4.4) shows that there are differences in performance on these tests between the Black and White groups, with the White group performing better across the whole range of tests. The question that arises is whether the socio-economic, environmental and educational background of the candidates (briefly

discussed in Chapter 2, section 6) influenced these results. The scope of the study did not include an evaluation of these factors and it is therefore not possible to draw conclusions with any degree of certainty in this regard.

5.4 LIMITATIONS OF THE STUDY

Since the study was a comparison of the performance of the Black and White groups of the first officer applicants at the South African commercial airline, existing assessment results were used. The results do not indicate whether these applicants succeeded in the selection process and were actually taken on as trainee pilots. This may have resulted in a skewed sample. The researcher wishes to highlight that this could be the reason for a lack of Black pilots employed by the commercial airlines in South Africa. Even thou the Black sample size was small, relative to the White sample size, statistically significant differences are still applicable.

The sample was one of convenience as it was readily available. While the sample is big enough to make generalisations, there is no equal representation of the samples as the Black group was much smaller than the White group.

Despite these identified limitations, the researcher believes that the research field and the South African commercial airline fraternity will benefit from the project because it provides evidence of the existence of a difference in performance between the Black and White groups on a battery of cognitive tests used for selection of trainee pilots.

5.5 RECOMMENDATIONS

It is recommended that the airline should spend time and resources thinking about the next generation of the pilots, like it does with the next generation of aircrafts. The airline, together with other South African airlines should offer bursaries or training bonds for school leavers that wish to pursue flying, the same way it is done for other professions (e.g., Engineering or Actuarial Science). In this way, those who have a passion for flying would be able to pursue the career without worrying about the costs that are associated with it. There is no doubt that there is socio-political pressure on companies to implement affirmative action as diversity remains an important aspect of corporate South Africa. If the best pilots happen to be Black, they should get the jobs, but the standards should not be lowered for anyone.

5.6 RECOMMENDATIONS FOR FURTHER RESEARCH

- Recruitment and selection strategies could be explored in order to provide a sustainable solution to the shortage of Black pilot skill in South Africa.
- An examination of the group differences in terms of interviews, simulator assessments and training results could be carried out.
- A comparative study could be undertaken to determine whether the results for cognitive ability tests conducted by other airlines reflect the same or different results.
- An explanatory study could be conducted to determine the extent of the impact of socio-economic, environmental and educational factors on the candidacy of Black pilots.

- The quality of the schooling systems and the matric results of the people that pursue flying as a career.

REFERENCES

Ahsan, K., Ho, M., & Khan, S. (2013). Recruiting project managers: A comparative analysis of competencies and recruitment signals from job advertisements. *Project Management Journal, 44*(5), 36-54.

Aikens, N. L., & Barbarin, O. (2008). Socioeconomic differences in reading trajectories: The contribution of family, neighborhood, and school contexts. *Journal of Educational Psychology, 100*, 235-251.

Almeida, R. A., Dickinson, J. E., Maybery, M. T., Badcock, J. C., & Badcock, D. R. (2010). A new step towards understanding Embedded Figures Test performance in the autism spectrum: The radial frequency search task. *Neuropsychologia*, *48*(2), 374-381.

Armstrong, M., & Taylor, S. (2014). *Armstrong's handbook of human resource management practice*. London, England: Kogan Page Publishers.

Ashrafi, L. (2012). *Design of a user interface to investigate the role of 3D spatial reasoning in surgical training*. (Masters dissertation). Available from Electronic Thesis and Dissertation Repository. (http://ir.lib.uwo.ca/etd/552).

Atherton, J.S. (2013). *Learning and teaching; cognitive theories of learning.* Retrieved from http://www.learningandteaching.info/learning/cognitive.htm.

Austin, R.D. (2013). Measuring and managing performance in organizations. Boston, MA: Addison-Wesley.

AviationSafety. (2015). *Industry's role in aviation safety.* Retrieved from http://web.archive.org/web/20110629183647/http://www.boeing.com/:

Babbie, E. (2010). *The practice of social research.* Ontario, Canada: Wadsworth Cengage Learning.

Babbie, E. (2013). *The practice of social research.* (13th ed.). Belmont, CA: Wadsworth Publishing Company.

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

Baguley, T. (2009). Standardized or simple effect: what should be reported? *British Journal of Psychology, 100(3)*, 603-617.

Bangerter, A., Roulin, N., & König, C. J. (2012). Personnel selection as a signaling game. *Journal of Applied Psychology*, *97*(4), 719-738. doi: 10.1037/a0026078

Barnard, M. (2010). Psychological assessment: Predictors of human behaviour. In M. Coetzee, & D. Schreuder, *Personnel Psychology: An Applied Perspective*, (pp. 134-169). Cape Town, South Africa: Oxford.

Beach, C. (2013). *At-risk students: transforming student behaviour*. Plymouth, England: Rowman and Littlefield Education.

Bergh, Z. (2009). Assessment of personality and individual differences. In Z. C. Bergh, & A. L. Theron (Eds.) *Psychology in the Work Context* (4th ed.) (pp. 406-432). Cape Town, South Africa: Oxford University Press.

Bezerra, T. A. R., Shimano, A. C., & Campos, F. A. D. (2015). Analysis of the forces exerted in flight by aviator cadets of the Brazilian Air Force. *Aviation in Focus – Journal of Aeronautical Sciences, 5*(2), 61-67.

Blair, J., Czaja, R. F., & Blair, E. A. (2013). *Designing surveys: A guide to decisions and procedures*. Thousand Oaks, NJ: Sage Publications.

Boon, C., Den Hartog, D. N., Boselie, P., & Paauwe, J. (2011). The relationship between perceptions of HR practices and employee outcomes: examining the role of person-organisation and person-job fit. *The International Journal of Human Resource Management, 22*(01), 138-162.

Bowler, J. (2013). The practice of research. *EZ514 Masters course class handout*, 1-15.

Brink, H. I. (1996). *Fundamentals of research methodology for health care professionals.* Cape Town, South Africa: Juta.

Bruton, A., Conway, J. H., & Holgate, S. T. (2000). Reliability: what is it, and how is it measured? *Physiotherapy*, *86*(2), 94-99.

Buglear, J. (2012). *Quantitative methods for business and management*. London, England: Pearson Higher Ed.

Burns, S. N., & Grove, S. K. (2003). *Understanding nursing research.* (3rd ed.). Philadelphia, PA: Saunders.

Carretta, T. R. (2000). *US Air Force pilot selection and training methods* (No. AFRL-HE-WP-TR-2000-0122). *71,* 950-956.

Carretta, T. R. (2011). Pilot candidate selection method. Still an effective predictor of us air force pilot training performance. *Aviation Psychology and Applied Human Factors, 1*(1), 3-8.

Carretta, T. R. (2013). Predictive validity of pilot selection instruments for remotely piloted aircraft training outcome. *Aviation, Space, and Environmental Medicine*, *84*(1), 47-53.

Carretta, T. R., Teachout, M. S., Ree, M. J., Barto, E. L., King, R. E., & Michaels, C. F. (2014). Consistency of the relations of cognitive ability and personality traits to pilot training performance. *The International Journal of Aviation Psychology, 24*(4), 247-264.

Carretta, T., & Ree, M.J. (2000). General and specific cognitive and psychomotor abilities in personnel selection: The prediction of training and job performance. *International Journal of Selection and Assessment*, *8,* 227-236.

Carroll, J.B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. London, England: Cambridge University Press.

Cattell, R.B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*(1), 1-22.

Catton, C. (2009). *Attitudes toward pilot recurrent training*. (Honours thesis), OpenSIUC, Southern Illinois University, United States of America.

Chakrabarti, S., & Fombonne, E. (2014). Pervasive developmental disorders in preschool children: confirmation of high prevalence. *American Journal of Psychiatry*, *162*(6), 1133-1141.

Civil Aviation Authority. (2015). *Civil aviation regulations, 1977*. Retrieved from: http://www.caa.co.za/Legislation/SACATS-%20ARM%2014-08-07.pdf

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillside, NJ: Erlbaum.

Cohen, J. (1992). Qualitative methods in psychology: A power primer. *Psychological Bulletin, 112(1)*, 155-159.

Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education.* (5th ed.). London, England: RoutledgeFalmer.

Creately. (2015). *Recruitment process*. Retrieved from: http://creately.com/diagram/example/gsy8pdq4f/Recruitment%20Process

Creswell, J. H. (2003). *Research design: qualitative, quantitative, and mixed methods approaches.* (2nd ed.). New Delhi, India: SAGE Publications.

cut-e. (2010). *Scales. Quality characteristics*. Retrieved from http://www.cute.com/

cut-e. (2013). *Scales - ability testing online*. Retrieved from: www.cute.com.

cut-e. (2015a). *Cut-e product finder*. Retrieved from http://www.cut-e.com/online-assessment/cut-e-product-finder/cut-e-product-finder-by-topic/cuteproduct/Product/list/

cut-e. (2015b). *References*. Retrieved http://www.cut-e.com/references/

Damos, D. E. L. (2011). *A summary of the technical pilot selection literature (No. DAS-2011-05)*. Gurnee, IL: Damos Aviation Services Inc.

Davis, R. J. (2013). *The validation of a psychological assessment battery for the selection of customer service agents in a south african commercial airline company.* (Unpublished masters thesis). University of South Africa, Pretoria, South Africa.

De Kock, F., & Schlechter, A. (2009). Fluid intelligence and spatial reasoning as predictors of pilot training performance in the South African Air Force (SAAF). *SA Journal of Industrial Psychology*, *35*(1), 31-38.

Denscombe, M. (2014). *The good research guide: for small-scale social research projects*. London, England: McGraw-Hill Education

Department of Public Enterprises. (2012). *Youth development programmes 2012 – 2013*. Retrieved from: http://www.dpe.gov.za/resourcecentre/publications/Youth%20Booklet/Youth%20Booklet%20DPE%20Youth%20Development%20Programmes%202012_2013.pdf

Dillon, W. R., Madden, T. J., & Firtle, N. H. (2000). *The essentials of marketing research.* (3rd ed.). London, England: Richard and Irwin.

Dror, I. E., Kosslyn, S. M., & Waag, W. L. (1993). Visual-spatial abilities of pilots. *Journal of Applied Psychology*, 763-773.

Employment Equity Act 1998. (RSA). No 55 of 1988.

Endsley, M. R., & Jones, W. (2013). Situation awareness. *The Oxford Handbook of Cognitive Engineering*, 88-108.

Explorable.com (2009). *Research population*. Retrieved from https://explorable.com/research-population

Ezzy, D. (2013). *Qualitative analysis*. London, England: Routledge.

Federal Aviation Authority. (2012). The age 65 law - information, questions and answers. *Fair Treatment of Experienced Pilots Act*, 1-5.

Fenton, M., & Neil, M. (2012). *Risk assessment and decision analysis with Bayesian networks.* Boca Raton, FL: CRC Press.

Flanagan, D. P., & Harrison, P. L. (Eds.). (2012). *Contemporary intellectual assessment: Theories, tests, and issues.* New York, NY: Guilford Press.

Flippo, E.B. (1976). *Principles of personnel management*. (4th ed.). Tokyo, Japan: McGraw-Hill Kogakusha.

Florea, V. N., & Badea, M. (2013). Acceptance of new technologies in HR: E-Recruitment in organizations. In *Proceedings of the European Conference on Information Management & Evaluation*, 344-352.

FlySAA. (No Date). *Pilot programme*. Retrieved from http://www.flysaa.com/za/en/SAA_Careers/RelatedLinks/cadet.html

Foxcroft, C. D. (2008). Planning a psychological test in the multicultural South African context*. SA Journal of Industrial Psychology, 38*(4), 8-15.

Foxcroft, C., & Roodt, R. (2001). *An introduction to psychological assessment in the South African context.* Cape Town, South Africa: Oxford.

Frederiksen, N., Mislevy, R. J., & Bejar, I. I. (Eds.). (2012). *Test theory for a new generation of tests*. London, England: Routledge.

Fry, L. W. J., & Wigglesworth, C. G. (2013). Toward a theory of spiritual intelligence and spiritual leader development. *International Journal on Spirituality and Organization Leadership, 47*.

Furr, R. M., & Bacharach, V. R. (2013). *Psychometrics: an introduction*. Thousand Oaks, CA: Sage.

Gardner, H. (1983). *Frames of mind*. New York, NY: Basic Books.

Gardner, H. (1991). *The unschooled mind. How children think and how schools should teach*. New York, NY: Basic Books

Garson, G. D. (2009) *Reliability analysis*. Retrieved from http://tx.liberal.ntu.edu.tw/~purplewoo/Literature/!DataAnalysis/Reliability%20Analysis.htm

Gay, L. R. (1987). *Educational research: competencies for analysis and application.* (3rd ed.). Columbus, OH: Merrill Publishing Company,

Gigaba, M. (2013). *On SAA's pilot cadet programme*. Retrieved from www.politicsweb.co.za

GL Assessment. (2015). *Introduction to the cognitive abilities test: fourth edition (CAT4)*. Retrieved from: www.gl-assessment.co.uk/cat4

Goldhaber, D. (2012). *The nature-nurture debates: Bridging the gap*. Cambridge, England: Cambridge University Press.

Good, P. (2013). *Permutation tests: a practical guide to resampling methods for testing hypotheses*. Berlin, Germany: Springer Science & Business Media.

Gowda, N. S. (2010) *Learning and the learner*: *Insights into the processes of learning and teaching.* New Delhi, India: Phi Learning Private Limited.

Gradireland.com. (2014). *Job description: airline pilot*. Retrieved from http://www.lit.ie/Careers/Job%20Profiles/Airline%20pilot.pdf.

Gravetter, F., & Wallnau, L. (2013). *Essentials of statistics for the behavioral sciences*. Boston, MA: Cengage Learning

Gregory, R. J. (2004). *Psychological testing: History, principles, and applications*. Boston, MA: Pearson Education Group.

Gregory, R. J. (2004). *Psychological testing: history, principles, and applications.* Boston, MA: Pearson Education Group.

Groth-Marnat, G. (2009). *Handbook of psychological assessment*. London, England: John Wiley & Sons.

Hantrais, L. (1995). Comparative research methods. *Social Research Update, 13*, 11-18.

Hartmann, P., Kruuse, N., & Nyborg, H. (2007). Testing the cross-racial generality of Spearman's hypothesis in two samples. *Intelligence, 35*, 47−57.

HCareers.com. (2009). *What is the job description of an airline pilot?* Retrieved from: http://www.hcareers.com/us/resourcecenter/tabid/306/articleid/617/default.aspx

Health Professions Act (RSA). No 56 of 1974.

Health Professions Council of South Africa. (2010). *Form 208. Policy on the classification of psychometric measuring devices, instruments, methods and techniques*. Retrieved from: http://www.hpcsa.co.za.

Health Professions Council of South Africa. (2013). *Form 233. Professional conduct & ethics*. *Rules of conduct pertaining specifically to psychology*. Retrieved from: http://www.hpcsa.co.za/Conduct

Heere, B., Walker, M., Gibson, H., Thapa, B., Geldenhuys, S., & Coetzee, W. (2015). Questioning the validity of race as a social construct: examining race and ethnicity in the 'Rainbow Nation'. *African Social Science Review, 7*(1), 2.

Hitt, M. A., Ireland, R. D., Sirmon, D. G., & Trahms, C. A. (2011). Strategic entrepreneurship: creating value for individuals, organizations, and society. *The Academy of Management Perspectives, 25*(2), 57-75.

Hysa, E. (2013). Defining a 21st century education: Case study of development and growth course. *Journal of Educational and Social Research, 3*(7), 704.

IATA. (2012). *Guidance material and best practices for pilot aptitude testing*. Geneva, Switzerland: IATA.

International Federation of Airline Pilots' Association. (2012). *Pilot training standards: Guide for best practices*. (pp. 1-32). London, England: IFALTA

International Test Commission. (2013). *ITC Guidelines on Test Use* (pp. 1-30). Surrey, England: International Test Commission.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences*, *105*(19), 6829-6833.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short-and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences, 108*(25), 10081-10086.

Jensen, A. R. (1992). Spearman's hypothesis: Methodology and evidence. *Multivariate Behavioral Research, 27,* 225-234.

Jensen, A. R. (2000). Testing: The dilemma of group differences. *Psychology, Public Policy, and Law, 6*, 121-127.

Kan, K. J., Wicherts, J. M., Dolan, C. V., & van der Maas, H. L. (2013). On the nature and nurture of intelligence and specific cognitive abilities the more heritable, the more culture dependent. *Psychological Science*, 0956797613493292.

Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology Monograph*, 74,657-690.

Kay, G. G., & Kennedy, C. H. (Eds.). (2013). *Aeromedical psychology*. Burlington, VT: Ashgate Publishing, Ltd.

Kelley, T. L. (1927). *Interpretation of educational measurements.* New York, NY*:* Macmillan.

Kendall, I. M., Verster, M. A., & von Mollenhorf, J. W. (1988). Test performance of Blacks in Southern Africa. In S. H. Irvine, & J. W. Berry, *Human Abilities in Cultural Context* (pp. 299-339). Cambridge, England: Cambridge University Press.

Labour Relations Act (1995) (RSA). No 66 of 1995.

Laerd Statistics. (2013). *Descriptive and inferential statistics.* Retrieved from https://statistics.laerd.com/statistical-guides/descriptive-inferential-statistics.php

Laerd. (2015). *What is predictive validity*? Retrieved from http://dissertation.laerd.com/criterion-validity-concurrent-and-predictive-validity-p2.php

Lewis, A. (2010). *Competency based interview questions made easy*. Retrieved from: http://www.blueskyinterviews.co.uk/int_art10.htm

Martinussen, M. (2014). Commentary on the article by Hoermann and Goerke: Do pilots need social competence? *The International Journal of Aviation Psychology, 24*(1), 32-35.

Mataboge, M. (2010). *Black pilot training fails to take off.* Retrieved from: http://mg.co.za/article/2010-09-28-black-pilot-training-fails-to-take-off

Mavin, T. J., & Roth, W. M. (2015). Optimizing a workplace learning pattern: a case study from aviation. *Journal of Workplace Learning, 27*(2), 112-127.

McArdle, J. J., & Woodcock, R. W. (Eds.). (2014). *Human cognitive abilities in theory and practice*. Hove, England: Psychology Press.

McDonald, R. P. (2011). Test theory: A unified treatment. Upper Saddle River, NJ: Routledge.

McIntire, S. A., & Miller, L. A. (2007). *Foundations of psychological testing*. Thousand Oaks, CA: Sage.

McMahon, T. W., & Newman, D. G. (2015). A methodology to determine the psychomotor performance of helicopter pilots during flight maneuvers. *Aerospace Medicine and Human Performance*, *86*(7), 641-646.

Mnguni, V. Q. (2011). *The predictive validity of a psychological test battery for the selection of cadet pilots in a commercial airline.* Unpublished masters dissertation, University of South Africa, Pretoria, South Africa.

Moerdyk, A. (2009). *The principles and practice of psychological assessment*. Pretoria, South Africa: Van Schaik Publishers.

Mondy, R. W. (2010). *Human resources management*. Upper Saddle River, NJ: Prentice Hall.

Montello, D. R., & Raubal, M. (2013). Functions and applications of spatial cognition. In D Waller & L Nadel (Eds.). *The APA Handbook of Spatial Cognition*. Washington D.C., WA: American Psychological Society.

Morgan, P. L., Farkas, G., Hillemeier, M. M., & Maczuga, S. (2009). Risk factors of learning-related behaviour problems at 24 months of age: Population based estimates. *Journal of Abnormal Child Psychology, 37*, 401-413.

Mouton, J. (1996). *Understanding social research.* Pretoria, South Africa: Van Schaik.

Mouton, J. (2012). *How to succeed in your masters and doctoral studies: A South African guide and resource book.* Pretoria, South Africa: Van Schaik Publishers.

Mouton, J., & Marais, H. C. (1991). *Basic concepts in the methodology of the social sciences.* Pretoria, South Africa: HSRC.

Mujis, D., Harris, A., Chapman, C., Stoll, L., & Russ, J. (2009). Improving schools in socioeconomically disadvantaged areas – A review of research evidence. *Social Effectiveness and School Improvement, 15*, 149-175.

Nagai, T., Abt, J. P., Sell, T. C., Clark, N. C., Smalley, B. W., & Lephart, S. M. (2014). Neck proprioception, strength, flexibility, and posture in pilots with and without neck pain history. *Aviation, Space, and Environmental Medicine*, *85*(5), 529-535.

Nell, V. (2000). *Cross-cultural neuropsychological assessment: Theory and practice.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Nicholson, B., & O'Hare, D. (2014). The effects of individual differences, prior experience and cognitive load on the transfer of dynamic decision-making performance. *Ergonomics*, *57*(9), 1353-1365.

Pallant, J. (2005). *SPSS survival manual. A step by step guide to data analysis using SPSS for windows (version 12).* Crows Nest, Australia: Allen & Unwin.

Parsons, T. D. (2011). Neuropsychological assessment using virtual environments: enhanced assessment technology for improved ecological validity. In *Advanced Computational Intelligence Paradigms in Healthcare 6. Virtual Reality in*

Psychotherapy, Rehabilitation, and Assessment (pp. 271-289). Heidelberg, Germany: Springer.

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.

Patterson, F., Ashworth, V., Zibarras, L., Coan, P., Kerrin, M., & O'Neill, P. (2012). Evaluations of situational judgement tests to assess non-academic attributes in selection. *Medical Education*, *46*(9), 850-868.

PlaneCrashInfo. (2015). *Accidents statistics.* Retrieved from: www.planecrashinfo.com.

Psychometrics Ltd. (No Date). *General & graduate reasoning tests*. Bedfordshire, England: Psytech International Limited.

Quellmalz, E. S., Timms, M. J., Silberglitt, M. D., & Buckley, B. C. (2012). Science assessments for all: Integrating science simulations into balanced state science assessment systems. *Journal of Research in Science Teaching*, *49*(3), 363-393.

Richter, L., Mabaso, M., & Hsiao, C. (2015). Predictive power of psychometric assessments to identify young learners in need of early intervention: data from the birth to twenty plus cohort, South Africa. *South African Journal of Psychology*, 0081246315599476.

Rinderman, H. (2012). African cognitive ability: Research, results, divergencies and recommendations. *Personality and Individual Differences.* Retrieved from http://dx.doi.org/10.1016/j.paid.2012.06.022

Rosenthal, R., Rosnow, R. L., & Rubin, D. B. (2000). *Contrasts and effect sizes in behavioral research.* Cambridge, England: Cambridge University Press.

Ruscio, J., & Mullen, T. (2012). Confidence intervals for probability of superiority effect size measure and the area under a receiver operating characteristic curve. *Multivariate Behavioural Research, 47(2)*, 201-223.

Rushton, J. P., & Jensen, A. R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law*, *11*(2), 235-294

Rushton, J. P., & Skuy, M. (2000). Performance on Raven's Matrices by African and White university students in South Africa. *Intelligence, 28,* 251–265.

SAAF. 2015. *Recruitment – aircrew – pilot or navigator*. Retrieved from http://www.saairforce.co.za/recruitment/pilot-or-navigator

Schlinger, H.D. (2012). The myth of intelligence. *The Psychological Record, 53*(1), 15-32.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*(2), 262 - 274.

Schneider, J. W. (2013). Caveats for using statistical significance tests in research assessments. *Journal of Informetrics*, *7*(1), 50-62.

Schreuder, D., & Coetzee, M. (2010). An overview of industrial and organisational psychology research in South Africa: A preliminary study. *SA Journal of Industrial Psychology*, *36*(1)

SHL. (2015). *Situational judgement.* Retrieved from www.shldirect.com.

Shouksmith, G. (2014). *Assessment through interviewing*. London, England: Elsevier.

Shuttleworth, M. (2009). *Test retest reliability.* Retrieved from https://explorable.com/test-retest-reliability

Society for Industrial and Organisational Psychology of South Africa (SIOPSA). (2006). Code of practice for psychological and other similar assessment in the workplace. *SIOPSA*, 1-16.

Snell, S., Morris, S., & Bohlander, G. (2015). *Managing human resources*. Boston, MA: Cengage Learning.

Society for Industrial and Organizational Psychology. (2015). *Types of employment tests*. Retrieved from: http://www.siop.org/workplace/employment%20testing/testtypes.aspx

Spearman, C. (1904). General intelligence objectively determined and measured. *American Journal of Psychology, 15*(2), 201-203.

Spector, P. E. (2012). *Industrial and organisational psychology: Research and practice*. Singapore, Malaysia: Wiley.

Sternberg, R. (2000). *The handbook of intelligence*. London, England: Cambridge University Press.

Sternberg, R. J. (2012). Intelligence. Wiley interdisciplinary reviews: *Cognitive Science, 3*(5), 501-511.

Stowell, R. (2015). *How modern flight instruction is training pilots to make fatal mistakes*. Retrieved from: http://www.aviationsafetymagazine.com/airplane/Flight-Training-Mistakes.html

Struwig, F. W., & Stead, G. B. (2004). *Planning, designing and reporting research.* Cape Town, South Africa: Pearson Education.

Terre Blanche, M., Durrheim, K., & Painter, D. (2006). *Research in practice: applied methods for the social sciences.* (2nd ed.). Cape Town, South Africa: University of Cape Town Press.

The Institute of Psychometric Coaching. No Date. *Verbal reasoning test guide.* Retrieved from http://www.psychometricinstitute.com.au/Psychometric-Test-Guide/Verbal-Reasoning-Test-Guide.html.

The Joint Aviation Awareness Programme, 2010. *The A-Z of careers in aviation*. Retrieved from http://www.caa.co.za/Documents/A-Z%20Booklet%202010.pdf.

Thurstone, L. L. (1931). Multiple factor analysis. *Psychological Review, 38*, 406-427.

Thurstone, L. L. (1938). Primary mental abilities. *Psychometrics Monographs, 1*.

UJUH. (2014). *SAA & SA Express team up to train pilots*. Retrieved from: http://www.ujuh.co.za/saa-sa-express-team-up-to-train-pilots/

Urbina, S. (2014). *Essentials of psychological testing*. London, England: John Wiley & Sons.

Valencia, R. R. (Ed.). (2012). *The evolution of deficit thinking: Educational thought and practice*. London, England: Routledge.

Van Der Merwe, R. P. (2002). Psychometric testing and human resource management. *SA Journal of Industrial Psychology, 28*(2), 77-86.

Van Rensburg, H., Basson, J. S., & Carrim, N. M. (2011). The establishment and early history of the South African Board for People Practices (SABPP). *Journal of Human Resources Management, 9*(1), 1-12.

Vroom, V. H. (1964). *Work and motivation.* New York, NY: Wiley.

Wickens, C. D. (2002). Situation awareness and workload in aviation. *American Psychological Society, 11*(4), 128-132.

Wingrin, D. (2013). *Tumult in the clouds: Stories from the South African Air Force 1920-2010.* Solihull: Helion.

Wingrin, D. (2014). *SAAF gains nine new pilots. Retrieved from* http://www.saairforce.co.za/news-and-events/1326/saaf-gains-nine-new-pilots

Witkin, H. A. (1969). *Embedded figures test*. Sunnyvale, South Africa: Consulting Psychologists Press.

Yerkes, R. M. (1919). Report of the psychology committee of the National Research Council. *Psychological Review*, 83-149.

Yount, R. (2006). *Research design and statistical analysis for Christian ministry.* (4th ed.). Fort Worth, TX: W.R. Yount.