Exploring the reliability of the Kilifi Development Inventory (KDI) for African children between the ages 6 – 35 months.

NOZIPHO THULADU MATHE

Student no. 202500235

MAY 2011

Submitted in partial fulfillment of the requirements of the degree of Masters in Social Sciences (Clinical Psychology), in the School of Psychology, University of KwaZulu-Natal, Pietermaritzburg.

Declaration

I, Nozipho T. Mathe, declare that

(i) The research reported in this dissertation, except where otherwise indicated, is my original work.

(ii) This dissertation/thesis has not been submitted for any degree or examination at any other university.

(iii) This dissertation/thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.

(iv) This dissertation/thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

(a) their words have been re-written but the general information attributed to them has been referenced;

(b) where their exact words have been used, their writing has been placed inside quotations marks, and referenced.

(v) Where I have reproduced a publication of which I am an author, co-author or editor, I have indicated in detail which part of the publication was actually written by myself alone and have fully referenced such publications.

(vi) This dissertation/thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the dissertation and in the References sections.

Signed: Date:

_____ _____

Recommended (APA) citation:

Mathe, N.T., (2011). *"Exploring the reliability of the Kilifi Development Inventory (KDI) for African children between the ages 6 – 35 months."* Unpublished Master's Thesis. University of KwaZulu-Natal, Pietermaritzburg.

Acknowledgements

This short-dissertation symbolizes many things; so many people need to be thanked:

I thank Dr. Bev Killian for her engaging, flexible and encouraging supervision style; and for providing a casual, relaxed, but intellectually invigorating space for me to locate my voice and writing style.

To the Valley Of 1000 Hills Research team, especially Dr. Jane Kvalsvig, thank you for allowing me to participate in this study and providing me access to your facilities and the research data.

To Mr. Oliver Bodhlyera, from the School of Actuarial Sciences, UKZN Pietermaritzburg, for assisting with the data analysis process.

To Stephen Olivier, a Research Masters Student from UKZN Pietermaritzburg, for also assisting with the analysis and interpretation of the data.

To my family for a supportive atmosphere conducive to studying and growing, and continuous encouragement in my career.

To my fellow classmates, friends, and interns, who shared great memories with me in the years.

To my lecturers, who made my studies in Pietermaritzburg an intellectually enriching
period of personal and professional development.

Lastly, to the Creator who has been my source of strength during life's journey.

Thank you.

Abstract

There is a dire need for psychological tests to be standardized for the African child and for appropriate norms to be developed. This study is part of two larger studies that began in 2008, namely the National Early Learning Standards Age Validation Study (NELDS) (Kvalsvig, Govender & Taylor; 2008) (Principal Investigator – Dr. S. Kauchali), and the INSTAPA Work Package 6 Study (Kvalsvig, Govender & Taylor; 2008*)* (Principle Investigator – Dr. J.D. Kvalsvig). This study considers aspects of the Kilifi Development Inventory (KDI) as a measurement tool used to provide a structured and standardized procedure for monitoring the cognitive development of children from the age of 6 to 35 months. It is situated in the positivist paradigm and has adopted the classical test theory in attempting to establish the reliability of the Kilifi Developmental Inventory. Stratified purposive sampling techniques were used to select the South African sample, which comprised 120 participants categorized by gender (males and females), age (6 – 18 months and 18 – 35 months) and area (urban and peri urban/rural). The internal reliability of the KDI was investigated. Results from the data analysis (using SPSS 18.0) suggest that the measurement tool as a whole had inadequate internal reliability and therefore, further investigations and modifications have to be undertaken in certain subtests of the measurement tool before the assessment tool can be standardized for the South African population.

Table of Contents

Acronyms used in this dissertation:

AIDS – Acquired Immune-Deficiency Syndrome **HIV**- Human Immune-Deficiency Virus

KDI – Kilifi Development Inventory (Kilifi Development Inventory, 2008)

NELDS - National Early Learning Standards Age Validation Study (Kvalsvig *et al*, 2009)

K-ABC – Kaufman Assessment Battery for Children (Kaufman, 1983)

CHAPTER 1

1. Introduction

The aim of this research is to explore the internal reliability of the Kilifi Development Inventory (KDI), which is a psychological measurement tool, used to assess the cognitive development of African children growing up in vulnerable and difficult circumstances. One hundred and twenty children between the ages of 6 – 35 months from three sites (KwaZulu-Natal, Western Cape Province, and North West Province) were selected using stratified purposive sampling to take into account different geographic regions. Trained examiners from the local areas administered various subtests of the adapted version of the KDI to the participants. The test scores were recorded and the data was entered into SPSS version 18.0.

This study formed part of a broader study, namely the NELDS, which assessed cognitive, language and emotional development of children between the ages of 6 – 48 months. For an explanation of how this study is situated within the NELDS study, see Diagram A on page 30. The broader study aimed to "scientifically undertake an age validation of the National Early Learning Standards Age Validation Study (NELDS) in children from birth to four years in selected Early Childhood Development (ECD) sites in the country" (Kvalsvig *et al.*, 2009, p. 4). Its objective was to determine whether the development indicators provided by the NELDS are appropriate and valid for African children between the ages of 6 – 48 months. Thereafter, recommendations to the relevant government departments will be made on any significant adjustments on the NELDS, in terms of age validity and socio-demographic factors (Kvalsvig *et al.*, 2009).

Infant assessment is defined as "the process of investigating and evaluating the developmental status of children who are between the newborn period and about 30 months of age" (Kopp, 1994, p.268). Assessing children

plays an important role in evaluating the infant's progress and determining which interventions or remedial services would best benefit the infant and family (Kopp, 1994). The KDI is a neuro-cognitive psychological assessment battery that provides "a structured and standardized procedure for monitoring the development of children from 6 to 35 months" (Kilifi Development Inventory, 2008; p.2). In this study, the KDI is used to establish whether the child is following a normative pattern of development specific to the identified age range, with the focus being on physical and cognitive development.

Child development is an exciting field of study that focuses on the following domains: physical, cognitive, emotional and social development. The KDI was designed to focus particularly on cognitive and physical development of children between ages 6 and 35 months, therefore the normal pattern of development is briefly discussed in these two areas, before considering some of the contextual factors that may affect child development.

Physical development can be briefly described as "changes in body size, proportions, appearance, and the functioning of various body systems; brain development; perceptual and motor capacities; and physical health" (Berk, 2003, p.5). Thelen (2000) suggests that simple motor skills in infants are learned through a process where they continuously modify their current abilities to specific tasks. This process requires that a child acquires the ability to coordinate the movements of their limbs when crawling or walking. "Opportunities for action depend on the presence of desired objects, suitable support surfaces, and helping social support" (Thelen, 2000, p.102). Cognitive development can be regarded as the "development of a wide variety of thought processes and intellectual abilities, including attention, memory, academic and everyday knowledge, problem

solving, imagination, creativity, and the uniquely human capacity to represent the world through language" (Berk, 2003, p.5).

In sub-Saharan Africa, malnutrition, cerebral malaria, HIV/AIDS, and numerous parasitic infections are among the common diseases which can have adverse effects on cognitive functioning (Holding, Taylor, Kazungu, Mkala, Gona, Mwamuye, & Stevenson, 2004). The KDI will therefore help tap into a broad range of these cognitive functions to enable the researcher to assess the neuro-cognitive sequelae of these diseases in very young children.

Chapter 2 will discuss the statistical processes involved when looking at the psychometric properties of a measurement scale. It will include the different methods used in establishing reliability; the difference between nominal and interval measurement; item selection when constructing an index; the process of analysis data with missing values; the role of factor analysis in establishing reliability; and issues related to the assessment of young children. In addition, previous and current literature in the field of early cognitive development in children aged 6 – 35 months will be outlined in order to understand what the normal path of cognitive development is in children of this age. Socio-cultural factors that may impact on cognitive development will also be mentioned. Chapter 3 presents the research methodology, sampling procedures, the instrument, and the data analytic techniques used in this study. In Chapter 4, the results from this study are presented and discussed in detail, and a comparison of the results with previous and current studies is made. Chapter 5 summarizes the overall findings of this study, specifically focusing on the internal reliability of the KDI. The conclusions are presented in Chapter 6.

CHAPTER 2

2. Literature Review

The aim of this research was to analyze and interpret the data collected from the three geographical areas in South Africa (KwaZulu-Natal, Western Cape and North West Province), and to explore the internal reliability of the Kilifi Development Inventory (KDI). In this chapter, issues related to the reliability of an assessment tool will be outlined, with a focus on different methods of establishing reliability; the process of analyzing missing data, item selection, different types of measurements, the role of factor analysis in establishing reliability, as well as issues related to assessing young children. Empirical literature from current and previous studies undertaken in the field of cognitive development and psychological testing will be outlined, as well as some of the socio-cultural factors that may impact on the cognitive development of children from adverse circumstances.

2.1. Item selection in the construction of an index

A scale or index is an ordinal measure of variables that is based on more than one item. Each item is carefully selected to fit the characteristics of the index, thereafter scores are assigned to patterns of responses (Babbie, 2001). It is therefore very important to select reliable items for an index. Face or logical validity is the first criterion used when selecting items to be included in an index (Babbie, 2001), for example, the gross motor skills index in the KDI needs to include items that indicate aspects of gross movement. Unidimensionality is another criterion which requires a composite measure to represent only one dimension of a concept. In other

words, items reflecting fine motor skills should not be included in a measure of gross motor skills (Babbie, 2001).

The researcher constructing the index needs to decide whether a variable will be measured generally or specifically, and therefore the nature of the items included will determine this criterion. Items selected for an index must be assessed for the amount of variance they provide. There are two options to guarantee variance, which is another criterion for index construction. Several item responses that divide people equally in terms of the variable need to be selected (Babbie, 2001), for example half religious and half atheist. This will enable the researcher to assess the items and characterize a person as either being religious or atheist.

The second option is to select items that differ in variance (Babbie, 2001), in other words, one item may identify half of the respondents as religious, while another item may identify a few respondents as religious. The items that are included in an index need to empirically be related to one another in order to assist researchers to predict how the questions will be answered (Babbie, 2001). The selected items in an index must be related to one another in some way. There are two types of relationships; the bivariate relationship referring to a relationship between two variables, and a multivariate relationship which deals with more than two variables. After reliable items have been selected, scores are assigned for each response thereby creating a composite measure from the selected items (Babbie, 2001).

2.2. Nominal and interval measurement: What is the difference?

The level of measurement of a measurement system gives insight into how different symbols of that system relate to one another (Reaves, 1992). There are four different levels of measurement, namely nominal, ordinal,

interval and ratio, however for the purpose of this thesis only nominal and interval measurements are discussed. "Nominal measurements use symbols that are names of a property, names that put each object into a category but do not contain any information about the amount of the property" (Reaves, 1992, p.75). An example of nominal measurement would be the property of shape, i.e. rectangle, square, circle, triangle. Nominal measurements do not take into consideration the concept of „more' or „less' as this is meaningless. The objects being measured are placed into categories that differ from one another (Reaves, 1992), which is different from interval measurement. Interval measurements are measurements in which the categories are placed in an order depending on the amount of a property they have, and the intervals between the categories are equal everywhere (Reaves, 1992). It is important for a measurement not to be affected by other sources of variability in order for it to be classified as being a good measurement.

2.3. Dealing with missing data during the analysis process

In quantitative research, prior to beginning the process of analysis, the researcher needs to search for patterns in the data to enable him/her to make sense out of the data (Babbie, 2001). Analyzing data and looking for patterns is particularly important when encountering data that has missing values in it. A way to deal with missing data in a study with very few cases, the researcher may choose to exclude the cases with no values from the analysis process. If this is done, then the researcher needs to make sure that the data available for analysis is sufficient and that the exclusion does not affect the representativeness of the sample. Another method of dealing with missing data that includes "Yes" or "No" responses would be to treat the blank answers as "No" (Babbie, 2001). A third way of dealing with the issue of missing data would be to carefully analyze the data to yield an interpretation of their meaning, and then score the cases accordingly. These methods are however conservative

techniques which stand the chance of affecting the index and may reduce the likelihood of it relating to other variables in the way that the researcher had initially hypothesized (Babbie, 2001).

2.4. Internal reliability and different methods of establishing reliability

The reliability of a test depends on "three key elements, namely: stability, internal reliability, and inter-observer reliability" (Singh, 2007, p.253). Internal Reliability refers to the internal consistency of an assessment tool. A reliable assessment tool should essentially yield the same results each time it is administered, however this reliability does not necessarily ensure accuracy of the results (Babbie, 2001). "Measures of internal consistency provide information about the extent to which items in a test, domain, or sub-domain „hang together' or are inter-correlated" (Cicchetti & Rourke, 2004, p.17). Internal consistency is "estimated by determining the degree to which each item in a scale correlates with each other item, and an average inter-item correlation is estimated" (Durrheim & Terre Blanche, 2002, p.90). Assessing for reliability does come with its problems at times because if only one observer is the source of the data, "we cannot guard against the impact of that one observer's subjectivity" (Babbie, 2001, p.141). Reliability is an important aspect in positivist measurement as it addresses issues related to objectivity (Durrheim & Terre Blanche, 2002). One observer may interpret what they see differently to another observer. In addition, when working with young children the concern with reliable data is exacerbated and this may also have an impact on the overall reliability of the scale (Babbie, 2001).

There are different techniques created to deal with the issue of reliability. These include the test-retest method which involves making the same measurement more than once, with the aim of getting similar results each time. Using established measures is another technique, which involves the use of measures that have already been proven to be reliable. Another technique includes the research supervisor who ensures that the interviewer has

recorded the correct data, by verifying the information recorded with some of the respondents (Babbie, 2001). The split-half method is another technique which involves dividing the test in half, then comparing the score from the one half with the scores from the other half. If the test is reliable it should yield similar scores on both halves (Reaves, 1992).

Factor analysis is a method used to elicit patterns among the variations in values of several variables. It does this by generating factors that correlate highly with some of the real variables that are independent of each other (Babbie, 2001). In other words, "it is a technique used to identify a relatively small number of factors that can be used to represent the relationship among sets of many interrelated variables" (Durrheim & Terre Blanche, 2002, p.362). It involves three steps, namely: computing of the inter-correlations between variables; extraction of factors; and then rotation of the factors so that a clearer picture of the factors is obtained (Durrheim & Terre Blanche, 2002). "The output of a factor analysis programme consists of columns representing the several factors generated from the observed relations among variables plus the correlations between each variable and each factor, which are called factor loadings" (Babbie, 2001, p.450). Factor analysis plays a role in establishing reliability in the sense that when there is a high correlation between the factors and the variables of the scale, this could suggest that the derived factors have internal consistency and enhance the reliability of the measurement tool (Babbie, 2001).

2.5. Infant Assessment of Cognitive Development

Infant developmental testing is accomplished by systematically observing the infant and gathering extensive information about the infant's skills from the primary caregivers. This method is more effective in gathering

reliable data from young children. In-depth detailed observations are made by a trained professional who utilizes established criteria in that test situation. The tests used to establish the infant's developmental progress include many test items that evaluate the skills the infant has. When testing infants, standardized instructions are followed by the tester for the administration of items, and standardized toys are given to the infant for behavioural observations (Kopp, 1994). If an infant is found to have developmental delays, then a diagnosis can be made. Diagnoses, which are medically recognized labels, are useful as they provide information about the infant's future (Kopp, 1994).

Psychological testing, specifically in young children from African populations has been a longstanding debate for many years. Some of the issues that have arisen from these debates relate to the appropriateness of psychological testing in populations from previously disadvantaged environments; the type of assessments that are used on African children and whether these assessments are culturally fair and appropriate for this population (i.e. verbal versus nonverbal assessments; language issues, generalizability of test scores); and whether the norms attained from testing are accurate and relevant to populations from culturally diverse backgrounds. Psychological assessments that measure cognitive functions have previously been based on Eurocentric norms and adapted from Western and European tests (Kwate, 2001). These norms did not take into consideration the cultural factors that impact on cognitive abilities in African populations. Therefore, the use of standardized and appropriate assessment tools is of vital importance for professionals working in the field of child development and cognitive assessment, as there is a dire need for accurate measurements of cognitive impairment to be established in order for early intervention to occur.

2.6. Piaget's Stages of Cognitive Development

Piaget described four stages of cognitive development which all children go through in the same order, and each stage reflects general modes of thought. "Intelligence influences all acts of thinking such as perception, language, and morality" (Lee & Gupta, 1995, p.6). The cognitive developmental stages relevant to this study are the sensorimotor stage and the preoperational stage.

The *sensorimotor stage* (from birth to 2 years) is when the child becomes aware of their environment through his/her physical actions. This stage leads to the acquisition of language and thought. The sensorimotor stage consists of six sub-stages. The first sub-stage involves reflexes and spontaneous movements (0 – 1 month). The reflex of "sucking is central to this stage because it brings the infant in contact with disturbance" (McCarthy & Reid, 2002, p.66). During the feeding process, the infant learns to compensate and adjust if he/she loses grip of the nipple (McCarthy & Reid, 2002). The second sub-stage is called first habits and primary circular reactions (1 – 4 months). Looking is first a reflex then becomes an acquired construction or primary circular reaction, which can be defined as "the active reproduction of interesting results that were first produced by chance" (McCarthy & Reid, 2002, p.68). The third sub-stage is called secondary circular reactions (4 - 8 months), and involves the coordination of vision and grasping. At this stage the "infant goes beyond the mere continuation of a result" (McCarthy & Reid, 2002, p.70).

The fourth sub-stage is called the coordination of means and ends (8 – 12 months), and involves the emergence of true acts of intelligence. The infant "relates means and goals more and more independently of environmental stimulation" (McCarthy & Reid, 2002, p.72). The fifth sub-stage is called tertiary circular reactions (12 – 18

months), where the infant makes compensatory accommodation of moving objects in his/her environment without disturbance (McCarthy & Reid, 2002). The sixth sub-stage is called inventions through sudden comprehension (18months – 2 years), and involves an "awareness of relationships rather than by trial-and-error groping. The infant operates on his/her environment in a premeditated way, mentally combining important environmental clues" (McCarthy & Reid, 2002, p.74). The child begins to grasp important features of a situation, and constructs the means to solve problems (McCarthy & Reid, 2002).

The *preoperational stage* (2 to 7 years) is when the child acquires logical thought (Lee & Gupta, 1995), and begins using symbols and language more extensively. At this stage of a child's life, thinking and reasoning are intuitive; the child learns without utilizing reasoning (Sadock & Sadock, 2007). The child's concepts are primitive in the sense that he/she can name objects, however is unable to name the classes of objects. This occurs because at this stage, he/she is unable to think logically or deductively (Sadock & Sadock, 2007).

A child in the *preoperational stage* has the sense of what is good and bad, however, he/she is unable to deal with moral dilemmas or issues. In addition he/she has the sense of *immanent justice*, a notion which suggests that children automatically believe that if they have done something wrong, they will be punished. Piaget argued that children at this stage are egocentric: "they see themselves as the center of the universe; they have a limited point of view, and they are unable to take the role of another person" (Sadock & Sadock, 2007, p. 134). The *concrete operational stage* (7 to 12 years) is when the child begins to reason logically. The *formal operational stage* (12 years and older) is when the child's abstract reasoning skills develop and improve (Berk, 2003).

2.7. The objective of the KDI as a Psychological Assessment

The KDI was developed for the African context and was based originally on the Kaufman Assessment Battery for Children (K-ABC-II) (Kaufman, 1983). The Kaufman Assessment Battery for Children (K-ABC-II) (Kaufman, 1983) is a widely used psychological test that assesses the quality and level of cognitive and neurological functioning in children from diverse cultural contexts (Holding *et al.*, 2004). The K-ABC-II has limited demands on verbal responses and enables the researcher to use it with children who speak different languages. However, language compatibility, or the use of nonverbal tasks, does not necessarily render a test culturally appropriate or sensitive.

The K-ABC-II also "provides a structured approach to assessment while incorporating teaching items to increase familiarity with the test material" (Holding *et al.,* 2004, p.247). These important features of the K-ABC-II are what advocate its use in African populations, with only minor cultural adaptations. Thus, although widely considered to be a culturally sensitive test, there were several subtests which required some modification to render them culturally appropriate. One such modification was the translating of the test into the local language spoken by the population in order for the test administrators to give instructions to the participants (Holding *et al.,* 2004). Other modifications that were required included the materials, task structure, and test demands. "These modifications extended the applicability of testing to children in a rural region of Kenya and resulted in greater task engagement" (Holding et al, 2004, p.252).

Initial piloting of the adapted version of the K-ABC-II in Kenya, suggests that this test will make a valuable contribution in terms of the monitoring of child development and identifying, at a very young age, children who require professional intervention (Holding *et al*, 2004). The purpose of the study in Kenya was to "evaluate the

application of neuropsychological test procedures developed for use in North America and Europe, to children in a rural region of Kenya" (Holding *et al.,* 2004, p.246). The aim was to ascertain whether these methods could be adapted to a non-Western culture in a way that would make the test valid and reliable. The outcome of the study in Kenya was the development of the KDI for African children from a rural district of Kenya called Kilifi.

"The KDI is an extension of an assessment previously applied in studies of early adverse health experiences, the Kilifi Developmental Checklist (KDC) "(Holding *et al*, 2004, p.1). The KDC focused on the assessment of psychomotor functioning in children between the age range of 9 – 60 months. The items were taken from various sources including the Griffiths Mental Development Scale, The Kenyan Screening Test for children aged 6 months to 6 years, The Portage Early Education Programme, and the Movement Assessment Battery for Children (Holding *et al*, 2004). "The KDI incorporates the psychomotor items from the KDC, and extends the range of skills and functions assessed to include a more detailed assessment of language, social and emotional functioning, and early executive functioning" (Holding *et al*, 2004, p.1).

The KDI assessment battery was developed to sample a range of skills and functions for infants and toddlers, by using materials and activities that are culturally appropriate; that the children are familiar with; and that parents can participate in. The three subtests of the KDI are: Locomotor (Walking, Sitting, Lying, Standing, Jumping, Hopping and Climbing); Eye-Hand Coordination (Playing with the Ball, Vision: Ring & Red Tassel, Button, Block Tower, Container & Cubes, Coin Box, Vision: Bead Threading, and Paper & Pen); and Executive Function (A-NOT-B and Self Control). In assessing physical and cognitive skills, there is a reliance on both parental reports and direct observation of the child (Kilifi Development Inventory, 2008).

2.8. Issues relating to the use of psychological assessments in multicultural contexts

Many forms of psychological assessments have been incorrectly used in that they have not been developed or standardized for the African context (Foxcroft & Roodt, 2005). Most African countries were once British colonies, and therefore modern psychological measures in Africa have tended to follow the same patterns of importing tests from European and North American backgrounds without careful and due consideration of the multicultural contexts that exist in Africa. Psychological assessments that have not been standardized for the African context are said to lack cultural equivalence and they are unable to accurately assess giftedness or deficit with regard to cognitive functioning in African children because of their Eurocentric nature (Foxcroft & Roodt, 2005). "Tests which are European-centered like the Wechsler Intelligence Scale for Children – Third Edition (WISC-III), are said to be heavily loaded with „White' cultural influences, and the test scores are meant to measure the mastery of that very same White culture" (Kwate, 2001, p 223). When used to test African children, the WISC-III could have dire consequences on their education, as there is a likelihood that African children would not be accurately or fairly assessed using this instrument (Kwate, 2001).

Some of the consequences of the invalidity of these imported instruments is seen when it is said that African children obtain group test scores that average one standard deviation below that of European or American children (Kwate, 2001). However, from a scientific and ethical perspective, these differences do not in any way suggest that African children have lower levels of intellectual ability (Kwate, 2001), but rather that the test instrument has been culturally biased against them. Researchers have argued that there are numerous socio-cultural factors that contribute to the low scores that African children tend to obtain on various western-based psychological assessment techniques (Kwate, 2001). Some of these socio-cultural factors include differences in environment and life experience such as parent-child interaction, culturally specific child rearing practices, and

formal, as well as informal learning and education. In other words, the way in which a child is brought up and the education they receive, may affect the nature, form and content of cognitive functioning, but not necessarily their ability. In addition, Schiele (1991) suggests that intelligence tests do not take into consideration the different ways in which African people construct knowledge. Rather indigenous knowledge systems incorporate and use different forms of knowledge systems to those utilized within Western and capitalist societies as acknowledged in the theories of cognitive development proposed by Vygotsky, Gardner and other leading theorists in this area.

The cultural contexts that exist in Africa have a large impact on the development of a child, both mentally and physically. According to Vygotsky, cognitive development in children relies a lot on the interactions they have with adults and peers, as well as the instructions they get from their caregivers (Lee & Gupta, 1995). As children grow, their cognitive capacity changes, allowing them to organize, plan, communicate, calculate and co-operate with others. During this process, they develop „cultural tools' which they use to deal with their environment.

Vygotsky's aim in his theory of development was to understand the way in which these cultural tools are acquired by the child. He argued that "concepts, language, voluntary attention and memory are functions which originate in culture, (i.e. in the interactions between people)" (Lee & Gupta, 1995, p.13). These functions are in turn acquired by the child through development, as they interact with other people and are taught by adults. A child's personal knowledge, thought and behaviour is shaped by the internalization of these cultural tools (Lee & Gupta, 1995).

Feuerstein (1979), in his Mediated Learning Experience Theory (MLE), states that as human beings we have the capacity to modify our cognitive functioning and cognitive structure; and this enables us to adapt to the changing environment (Tzuriel, 2001). Cognitive modifiability refers to the changes in cognitive functioning that occur in a child. This process depends on the Mediated Learning Experience, which consists of "an interactional process in which an adult interposes himself/herself between the child and a set of stimuli, and modifies them by affecting their frequency, order, intensity, and context" (Tzuriel, 2001, p.24). The MLE results in the child's curiosity, vigilance, and perceptual acuity being aroused.

The child and the adult (as mediator) both interact during the MLE to improve and/or create cognitive functions required for temporal, spatial, and cause-effect relationships. The child then internalizes the MLE processes until they become a mechanism of change within the child, and this ultimately enables him/her to use them to benefit from learning experiences (Tzuriel, 2001). According to Feuerstein, mediation is not dependent on the language modality or content. Anyone and everyone can interact with the child in order to assist in their cognitive development. He also argues that when it comes to assessment in multicultural contexts, individuals from different cultures may show initial low performance on cognitive abstract tests compared to individuals from mainstream culture. However, they are able to improve after a short intervention entailing mediated learning experiences (Tzuriel, 2001).

Rogoff's Theory of Development (1998), on the other hand, argues that development occurs on three interacting planes, rather than two as suggested by Vygotsky. These include the individual child; the people he/she lives with in the community; and the socio-cultural context in which these people produce and share knowledge. "The process of participation (i.e. through observation, social interaction, and direct teaching) within a

community serves to delineate a child's development capacity, to the extent that development itself is defined by the child's evolving understanding of the socio-cultural context in which he/she lives" (Edwards, 2005, p. 39). Development is therefore a process of people's changing interactions and participation in the socio-cultural activities in their communities (Edwards, 2005).

This study takes these socio-cultural factors into consideration when trying to understand early cognitive development in children who grow up in adverse circumstances; even those of whom do not have parents and are being looked after by multiple caregivers. The knowledge gained from being exposed to the elders in a community and interacting with other children from the neighbourhood brings a sense of community for the child. Within that community, knowledge is shared and the basis of the child's cognitive development relies a lot on that shared knowledge. When standardizing psychological assessments, researchers in the field have to develop culturally appropriate norms that account for learning and cognitive abilities that develop in culturally specific ways, so that the child's abilities are assessed and understood in his/her context.

Bronfenbrenner's Ecological Systems Theory suggests that a "child develops within a complex system of relationships affected by multiple levels of the surrounding environment" (Berk, 2003, p.27). The child's development is molded by his/her biological disposition as well as the environmental factors that he/she experiences. He described five systems that influence a child's development in many ways. The *microsystem* includes relationships that the child has with his parents, peers and immediate surroundings. The relationships are bidirectional, which means that parent's behaviour affects the child, and the child's personality, attitude and capacities affect the parent's behaviour (Berk, 2003), for example, if a parent is emotionally unavailable and does not respond appropriately to the child's needs because of stress and an inability to fulfill parental

responsibilities, this contributes to the child not being able to form a secure attachment and establish a healthy self image (Manegold, 2002). This may cause the child to withdraw or be overly demanding towards the parent, which may result in the parent being aggressive or hostile towards the child.

The *mesosystem* consists of the connections between the microsystems such as the home, school, neighbourhood and church (Berk, 2003). An example is when the parent shows interest in the child's school work by communicating with the teacher to monitor his/her progress. In addition, the parent can continue monitoring the child's academic learning process by assisting him/her at home with homework and other activities. These interactions are said to support the development of the child.

The *exosystem* is made up of social settings that do not directly involve the child but affect his/her experience in the immediate environment. These include the parent's workplace or health care service (Berk, 2003). For example, unemployment places the child and the caregiver at risk in different ways. When a parent is unemployed he/she may have financial difficulties that may result in the child not being able to attend school. This ultimately will affect the child's academic development especially when the parent is stressed and cannot contribute to the child's learning process at home.

The *macrosystem* includes major historical events, cultural and religious values, legal and political practices, and customs shared by a specific cultural group (Bukatko & Daehler, 1995). There are substantial differences in the prevalence and mode of child rearing techniques across cultures. What is important to note is that parents from different cultural backgrounds are said to "behave towards their children the way they think good,

responsible parents should behave as described by cultural norms" (Lonner & Malpass, 1994, p.115). Children perceive the parental acceptance-rejection concept in different ways. For instance, if a child experiences hostility and aggression from a parent they may feel unloved or they may think of their parent as being a bad person. "Caregivers' own parenting and attachment history is known to have a major impact on parenting styles, attitudes, and attribution of infant behaviours; and provides an objective for intervention" (Manegold, 2002, p.45).

The *chronosystem* includes the changes that occur in the child's life such as losing a parent/s or siblings to the HIV/AIDS epidemic, resulting in them being looked after by other people. These events modify the relationships that exist between the child and his/her environment, and create new conditions that may affect the child's development (Berk, 2003).

2.9. Risk Factors in the African Context

Risk factors are "those characteristics or experiences that increase the possibility of the occurrence, the severity, the duration, or the frequency of later disorder" (Manegold, 2002, p.39). In Africa, HIV/AIDS is a major disease affecting people who live in adverse circumstances such as poverty, unemployment, and inequality. HIV is transmitted from mother to the child either during pregnancy, during birth, and from breastfeeding. Approximately two-thirds of these infants will not become HIV positive due to the use of Azidothymidine (AZT). For the one-third of the infants that do become infected, signs of the virus become apparent in the first year of life. A child suffers from a depleted immune system, resulting in him/her becoming vulnerable to other more serious illnesses (Manegold, 2002).

"Neurocognitive aspects of HIV infection in a child may involve progressive encephalopathy (PE), which is characterized by the following symptoms: impaired brain growth, progressive motor dysfunction, loss or plateauing of developmental milestones, as well as cognitive and language delays" (Herbert, 2003, p.361). These neurological complications caused by HIV infection usually occur during the development of the central nervous system, and this has a negative impact on the child's cognitive development. Learning difficulties and behaviour problems become apparent, and the child may also suffer from psychosocial difficulties (Herbert, 2003).

2.10. Advantages of Standardized Psychological Assessments

The purpose of many studies currently being undertaken in South Africa is the "construction of new culture-reduced tests; and adapting, refining, and norming appropriate tests that have been constructed and proven to be valid and reliable in other countries" (Foxcroft & Roodt, 2005, p.103). One of the reasons why adaptations to already existing assessment measures are of importance is because they "enhance fairness in assessment by allowing persons to be assessed in the language of their choice" (Foxcroft & Roodt, 2005, p.58). In addition, they take the cultural context of African children into consideration. The bias associated with being assessed in a second or third language is removed, and this increases the validity of the test results.

Another advantage to adapting and standardizing already existing psychological assessments is that the costs are reduced and time is saved. Adapted assessment measures enable the facilitation of comparative studies between languages and cultural groups, which allows for many nationalities and groups to know and learn more about each other. In addition, newly developed measures can also be compared to already existing norms, interpretations and information (Foxcroft & Roodt, 2005).

Measures are considered to be equivalent cross-culturally when "individuals with the same or similar standing on a construct but belonging to different groups, obtain the same or similar scores on the different language versions of the items or measures" (Foxcroft & Roodt, 2005, p.59). If they do not, then the items are said to be biased and the two versions of the test are non-equivalent (Foxcroft & Roodt, 2005). When individuals from a multicultural context are assessed, the researcher needs to realize that their reasoning is embedded in a particular context as well as in the knowledge they already have. For the persons being assessed, "interpreting language is more than a matter of interpreting word meaning" (Lee & Gupta, 1995, p.17).

A model that has provided professionals interested in the field of psychological assessment with an alternative approach to assessing the abilities of children from previously disadvantaged contexts and culturally diverse backgrounds is the Dynamic Testing Model. The goal of this model is to see whether and how the child may change if given an opportunity to do so. The way in which it proposes to do this is by reducing the obstructive effects that environmental variables can have on a child's performance, and thus quantify a person's true potential for growth from whatever point of cognitive development they may be (Grigorenko & Sternberg, 1998). In other words, instead of testing previously acquired knowledge alone, "Dynamic Testing assesses the capacity to master, apply, and reapply knowledge taught in a dynamic testing situation" (Grigorenko & Sternberg, 1998, p.76).

Dynamic testing is said to provide a basis for teaching and development of cognitive skills. Its concept deals with concrete methodologies and has been widely used for testing the ability of job-seeking candidates, mentally retarded individuals, patients with brain damage, as well as children, adults and aging individuals (Grigorenko & Sternberg, 1998). The contributions of Dynamic Testing in the field of psychological assessment

have been positive in the sense that it may provide more information related to the learning potential of individuals, and may be considered to be a fairer testing procedure for those who are unfamiliar with the testing situation, or those from different cultural backgrounds as opposed to static testing procedures (Schulte, 2004). However it was not suitable for this study as it has poor qualities for screening purposes.

2.11. Debates around Psychological Assessments and Multiculturalism

There are a number of debates that revolve around the issue of multiculturalism and psychological assessments. These include: *etic versus emic perspective; assimilation versus pluralism; group personality versus individual difference; and racism.*

The etic and emic perspectives influence the way assessment devices are selected and interpreted. Traditionally, many psychologists favoured an *etic* perspective which emphasizes the "universals among human beings by using examination and comparison of many cultures from a position outside those cultures" (Dana, 1993, p.21). This perspective assumes the "commonalities in all human experience and behaviour as psychological universals" (Lonner & Malpass, 1994, p.130). If we assume the universality of human behaviour and experience, we then ignore the influence that culture has on explaining our psychological well-being (Lonner & Malpass, 1994). Scientists working from an absolutism position believe that "there is an underlying common (true) nature to all human beings that can be identified, described, and used to explain the products of their activity" (Lonner & Malpass, 1994, p.130). Cultures are mere masks of basic human truths, and if variables such as cultural norms and expectations are removed, „true' human nature will be revealed (Lonner & Malpass, 1994).

The imposed *etic* perspective was adopted where the middle-class Anglo-American group was used as a standard to compare to other groups. For example, the original norms for the Minnesota Multiphasic Personality Inventory (MMPI-1) described a 35-year-old white, rural, married, semi-skilled person with grade eight (Dana, 1993). The description in the MMPI-2 was somewhat different; however, the standard was still predominantly that of middle-class. This standard suggests that any cultural differences are to be treated as statistical differences that describe a shift away from normality (Dana, 1993).

The *emic* perspective on the other hand is "culture-specific and examines behavior from within the culture, using criteria relative to the internal characteristics of that culture" (Dana, 1993, p.21). This perspective emphasizes that people from non-Anglo-American groups should be understood as they are, without being compared to the standards of people from Anglo-American groups. This perspective requires a researcher to really get to know a particular cultural group by becoming a long-term resident in that community, and learning everything about the culture of that group (Dana, 1993).

According to Lonner and Malpass (1994), scientists working from the relativism perspective believe that human behaviour and its explanations are determined by socio-cultural and historical factors that exist in society. These factors are in turn created by the very people who formulated them; therefore, what is valid in one culture is not at all valid in another culture.

Assimilation versus pluralism refers to the cultural orientation of the person being assessed. This means that the person assessing should have sufficient knowledge about the cultural orientation of the client before any

assessment takes place. It was presumed that "homogenization of different ethnic groups would produce a prototypical American" (Dana, 1993, p.22). However, it was discovered that the American character is made of people who create meaning in their lives by remembering their past sufferings and future aspirations. This cultural identity is seen as a source of strength for minority groups who continue to carry on with their lives even when faced with discrimination and limited opportunities (Dana, 1993).

The cultural *pluralism* perspective "accepts individual differences as personal assets and aspires to maintain separate institutions for distinct social groups within a single political unit" (Dana, 1993, p.22). The problem for culturally different populations lies in how they can maintain essential aspects of their unique and historic identities, while still being active citizens of the community. Dana (1993) argues that when assessing people from culturally diverse backgrounds, culture and its contribution must be separated from the presenting problem and symptom. Moderator variables are then used to assess the extent to which acculturation has occurred or whether the original culture has been maintained. Moderator variables are used to correct cultural differences during assessments (Dana, 1993).

Group personality versus individual differences: Psychologists often act on the basis that "research evidence emphasizes group differences rather than within-group variations, even though it has been acknowledged that within-group differences are greater than differences between groups" (Dana, 1993, p.23). Every cultural group consists of individuals who develop a sense of self from the time of birth. This sense of self enables the person to separate himself or herself from other people, while still being part of that culture. The self relates to society in different ways, and the way in which the self is defined differs from one culture to another (Lonner & Malpass, 1994). Members from a segment of a cultural group would feel ‚different' if placed in another

segment of that cultural group. For example, a person who is disabled might feel different and uncomfortable if placed in a group with people who are not disabled, even if they both belong to the same church denomination.

Racism: "Although many culturally different groups have their own unique history pertaining to discrimination, prejudice and racism, most ethnic minority groups have been historically depicted as culturally deficient on both socially desirable and devalued attributes" (Dana, 1993, p.23). According to Lonner and Malpass (1994), white European people believed that black African people were genetically, culturally, and socially inferior to them. They discriminated against them as they were viewed to be different, and so this could be used to benefit them economically, socially, politically and psychologically. By using tests that are not culturally appropriate, some of the attitudes of racism are re-enacted with negative impacts on the African individuals being tested and the decisions made on the basis of the biased testing procedures.

Therefore, when trying to assess individuals from culturally diverse backgrounds, it is important to take their cultural values, practices and beliefs into consideration as these factors impact on the assessment procedures and cooperation of the individuals with the researcher. When research participants are not recognized as being as important as their counter-parts (American or European individuals), this may underestimate the role that these people play in society. Currently used psychological tests have tried to take into consideration the past experiences of minority groups, especially with regard to racism and discrimination, however there is still a large gap that needs to be bridged when it comes to the development of norms that account for these cultural aspects. People who fall within minority groups need to be understood for who they are and what they have to offer with regard to knowledge production in their society.

2.12. Situating this study within the larger studies

The Kaufman Assessment Battery for Children (K-ABC-II) is a psychological assessment tool used to assess cognitive and neurological development in young children from North America and Europe (Kaufman, 1983). It has been considered appropriate for the African context as it makes limited demands on verbalization, and therefore can be used on children from different cultural backgrounds (Holding *et al*, 2004). However, due to the fact that this assessment tool was developed for North American and European children, certain subtests had to be standardized for the African child and appropriate norms developed before its use in the African context. The K-ABC-II consists of the following subtests: Face Recognition, Gestalt Closure, Magic Window, Word Order, Triangles, Number Recall, Hand Movements, and Matrix Analogies, Arithmetic, Picture Vocabulary, Pegboard, Visual Search, and Pragmatic Errors (Kaufman, 1983).

A study in Kenya (Holding Study) used the adapted version of the K-ABC-II to develop a new neuropsychological assessment battery (KDI) for a rural population in the Kilifi District. The objective of the Kenyan study was to assess the school-age sequelae of several malarial diseases in the Kilifi District, as malaria is endemic in that coastal region of Kenya (Holding *et al*, 2004). The focus was to "produce material appropriate for children aged 5-7 years, as this age is a key time in the child's life in relation to disease factors and socio-cultural influences" (Holding *et al*, 2004, p.247). Eight subtests of the K-ABC-II were chosen and adapted as they were considered to adequately assess cognitive processing skills in children of school-going age; i.e. children in the age range of 5 – 7 years. An additional subtest was included in this study, namely "a modified version of the Child Behaviour Questionnaire for Parents" (Holding *et al.,* 2004, p.248).

This assessment battery was called the Kilifi Developmental Inventory (KDI), and was developed to measure cognitive development by looking specifically at the following: Locomotor functioning, Eye-Hand Coordination, and Executive Functioning. The KDI is an extension of the Kilifi Developmental Checklist (KDC), which is an assessment battery, previously applied in studies of early adverse health experiences (Holding *et al*, 2004, p.1). The KDC focused on the assessment of psychomotor functioning in children between the age ranges of 9 – 60 months.

The KDI was developed in three phases: "The first phase focused on investigating the cultural applicability of the test content and procedures. Phase two involved subtest selection and item modifications, while the third phase investigated sensitivity and external validity by assessing the relationship of performance with biological and social risks, and with measures of academic ability and behaviour" (Holding *et al.,* 2004, p. 247). The subtests of the KDI that measure Locomotor functioning include: Moving, Standing, Lying, Walking, Jumping, Hopping and Climbing. Those that measure Eye-Hand Coordination include: Playing with the Ball, Vision (Ring & Red Tassel), Button, Block Tower, Container & Cubes, Coin Box, Bead Threading, and Paper & Pen. To assess Executive function, the A-NOT-B and Self Control subtests were used (Holding *et al.,* 2004).

There is a large research programme currently underway in KwaZulu-Natal, North West Province, and Western Cape called the National Early Learning Standards Age Validation Study (NELDS). It began in 2008 and aims to age validate the Kilifi Developmental Inventory (KDI) for the African child growing up in adverse circumstances. It focuses on the physical and psychosocial functioning of South African children from poor and vulnerable settings between the ages of 6-48 months (Kvalsvig *et al.,* 2009). Its objective is to undertake a national and international literature and policy review on similar programmes and processes. Recommendations

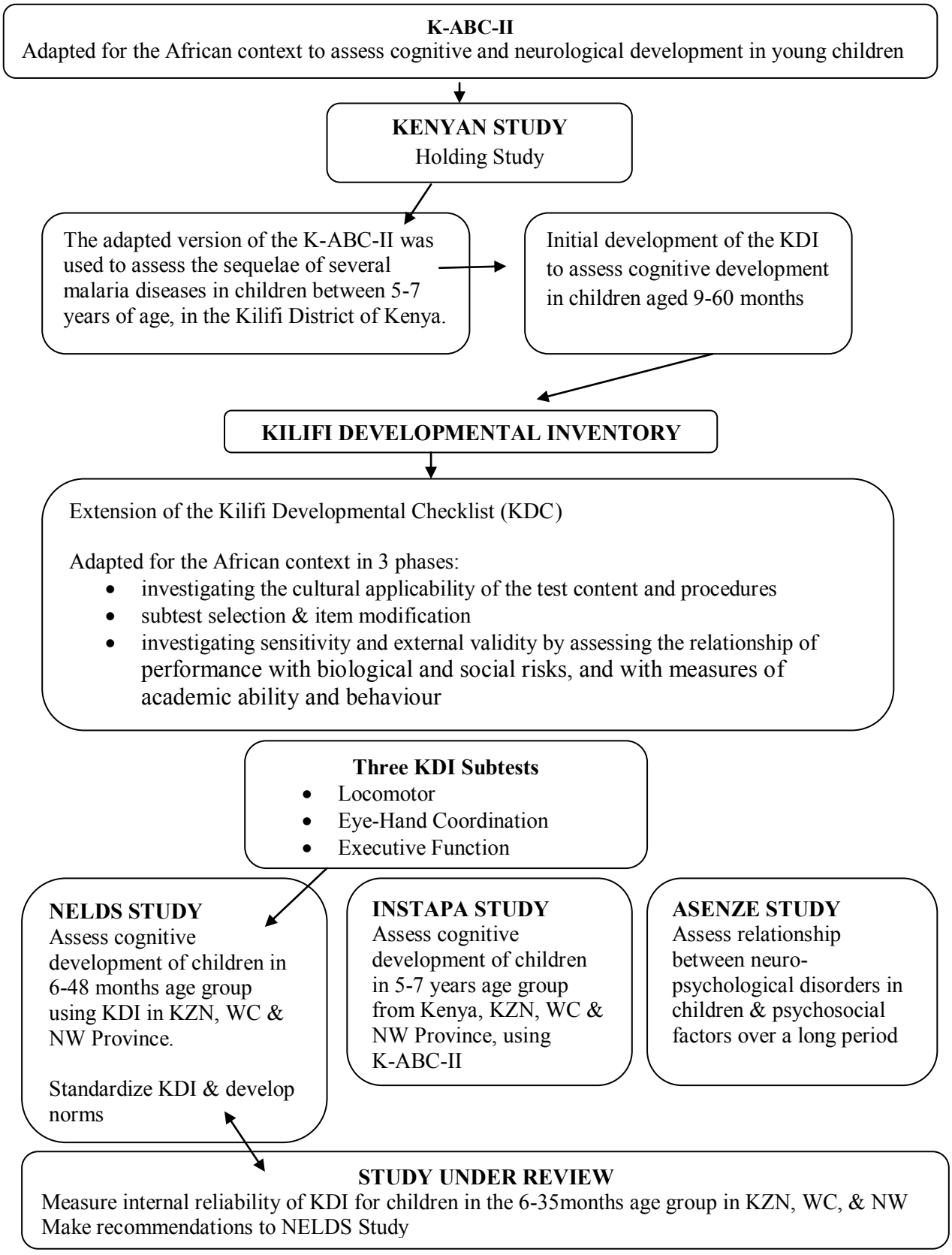
to the relevant government departments will be made for adjustment to the National Early Learning Standards Age Validation Study (NELDS) in terms of age validation and socio-demographic factors (Kvalsvig *et al.*, 2009).

For the purpose of adapting the KDI for the South African context, the researchers have modified the language to suit the target populations, i.e. *IsiZulu* (KwaZulu-Natal), *Afrikaans/IsiXhosa* (Western Cape), and *Setswana* (North West Province); however, the original test construct was largely retained as it appeared to be culturally appropriate. Furthermore, modifications were made to the test content, test items, and the format. In order to check for construct equivalence, factor structures of the tests from the cultural groups (KwaZulu-Natal, Western Cape, and North West Province) were compared as suggested by Foxcroft and Roodt (2005), for interrogating the structure of the adapted test. The KDI will help contribute to the early detection of neuropsychological disorders and developmental delays in children. This will assist in filling a current gap in diagnostic tools that can be used by a psychologist.

The study under review forms part of the larger NELDS study, and focuses specifically on the Kilifi Developmental Inventory (KDI) and its properties by measuring its internal reliability. If accurate measurements of cognitive impairments for children in South Africa between the age group 6 – 35 months are achieved, then the results will give the researcher an idea of whether this assessment tool is appropriate for this target cultural population, thus enabling researchers from the larger study (NELDS) to standardize it and develop norms for this population. This study therefore sets out to achieve a statistical investigation of internal reliability of the KDI. However, a more in-depth understanding of the larger study will assist the reader to contextualize this smaller study.

There is another study taking place concurrently with the NELDS Study in KwaZulu-Natal, Western Province, and North West Province, and Kenya; namely the INSTAPA Work Package 6 Study. It began in 2008 and aims to standardize some of the subtests of the Kaufman Assessment Battery for Children (K-ABC- II) and develop norms for the South African child between the ages of 5-7 years. For the purposes of adapting the K-ABC-II, the language was modified (*IsiZulu* in KwaZulu-Natal, *Afrikaans/IsiXhosa* in Western Cape, *Setswana* in North West Province, and *Kiswahili* in Kenya) to suit the local language of the target population in order for the examiners to administer the test easily in the language of the population. In addition, the subtests were examined item by item to assess for floor and ceiling effects, to check which drawings were scorable, and assess the clarity of the instructions (Holding *et al.,* 2004). This study was stimulated by the need to develop more accurate assessment techniques so that diagnosis of developmental difficulties in children could be made early in a more reliable and valid manner (Holding *et al.*, 2004). See Diagram A below for an explanation of how the study under review is situated within the NELDS and INSTAPA Work Package 6 Study.

Diagram A: Situating the current study within the two larger studies, namely NELDS and INSTAPA.



2.13. Research from developing contexts

There is a study currently taking place, namely the ASENZE Study, in a rural village called the Valley of 1000 Hills in KwaZulu-Natal. Its objective is to intervene in promoting better physical and psychosocial functioning in South African children. It aims to establish how children with neuropsychological disorders are able to function cognitively and socially, and how this is influenced by factors such as HIV/AIDS, socioeconomic and environmental conditions, as well as psychosocial aspects (caregiver characteristics, parent mental health and substance abuse, and family function) (Kauchali, 2008).

A sample of 2000 5-6 year old children from the HIV/AIDS prevalent area were screened for disability before school entry and re-assessed 24 months after school entry. The study offered HIV testing to children and their caregivers, and referrals to professionals were made for treatment if any medical conditions were identified. The study aims to investigate cross-sectionally and longitudinally, "the relationships among neuro-developmental disorders and the above-mentioned risk factors, including as outcomes, child cognitive and psychosocial functioning, as well as school functioning" (Kauchali, 2008, p.1). The study will therefore address these relationships; assess the impact of referrals, treatment, and access to programmes over time. An early detection of neuro-developmental disorders will enable the professionals to intervene before the affected children start attending school (Kauchali, 2008). Information from this cross-sectional and longitudinal investigation will enable researchers from the NELDS and INSTAPA studies to understand how the early detection of cognitive deficits as well as appropriate interventions can benefit, and possibly minimize the effects that psychosocial factors would otherwise have on the children in the future.

2.14. Conclusion of the literature review

The literature presented in this chapter suggests that before any measurement tool can be deemed appropriate for use on a population; its psychometric properties have to be assessed first. Reliability is a very important aspect in statistics especially when it comes to assessment tools. When an assessment tool is reliable, this means that it has internal consistency and the same results are obtained every time the same construct is measured. It is somewhat difficult for an assessment tool to be perfectly reliable; however there are techniques used in statistics that can manage technical errors that may arise. When assessing infants specifically, one may encounter difficulties as sometimes the data collected may not be reliable or there may be missing values in the data. Techniques such as the exclusion of the missing values; treating the blank values as „No' responses; or analyzing the missing data to find patterns or meaning, may assist in managing the issue of missing data.

Assessment tools that have not been adapted or standardized for populations from multicultural backgrounds may elicit inaccurate results. There have been many debates around the issue of psychological assessments in multicultural contexts, which have raised many questions about the ethical dilemmas involved in testing. It is clear, from the literature, that for a research team to adapt an already existing assessment tool; many cultural factors need to be taken into consideration. The research results from the Kenyan and the ASENZE study will enable the researcher to think critically about the possible factors that could significantly influence the normal pattern of cognitive development in young children.

CHAPTER 3

3. Methodology

In this chapter, the researcher approached the study from an objective perspective and situated it within the positivist paradigm. The classical test theory was adopted, and steps from the theory were used to establish the reliability of the KDI. The sample was taken from three areas, KwaZulu-Natal, Western Cape, and the North West Province. Children between the ages of 6 – 35 months were selected using stratified purposive sampling techniques. One hundred and twenty children were assessed using the KDI. The writer did not do the assessments herself and relied on data collected by the research teams who were trained and experienced in infant assessment procedures and the rigours required for research. Before the data was analyzed, it was cleaned up and prepared. There were missing values in the data, however this issue was managed accordingly by using statistical techniques appropriate for managing missing data. The SPSS version 18.0 system was used to analyze the data in order to assess the internal reliability of the KDI, and the five percent significance level was chosen as criterion for all the statistical comparisons. A factor analysis was conducted on each KDI subtest, i.e. Locomotor, Eye-Hand Coordination, and Executive Function; to elicit possible underlying factors as well as the relationship between these factors. An Analysis of Variance (ANOVA) was conducted for the Locomotor and Eye-Hand Coordination subtest to assess whether there were any differences in subtest item scores for Age, Gender and Region.

3.1. Research Paradigm and Theoretical Perspective

This study is situated in the positivist paradigm, where what is being studied "consists of a stable and unchanging external reality" (i.e. cognitive mechanisms) (Durrheim & Terre Blanche, 2002, p.6). It is quantitative in approach and employs a methodology that enables manipulation of reality (Durrheim & Terre

Blanche, 2002). Classical Test Theory and some of its steps were adopted as an approach to psychometrics, in order to establish the reliability of the KDI. The Classical Test Theory is based on the assumption that "an obtained test score reflects both true score and error score" (Weiner, Freedheim, Graham, Schinka & Naglieri, 2003, p.44). Test scores are usually expressed in the formula: Observed Score = True Score + Error. In this formula, the observed score is the score that was obtained after the assessment was done. The true score represents the "hypothetical amount of the designated trait specific to the examinee, a quantity that would be expected if the examinee was tested an infinite number of times without any confounding effects of practice or fatigue" (Weiner *et al*, 2003, p.44). Measurement error, on the other hand is the difference between the true score and observed score. Error does not correlate with true score and other variables, and it is normally and uniformly distributed about the true score (Weiner *et al*, 2003). The influence of error is random; therefore the average measurement error across many testing situations is expected to be zero. Internal reliability is a psychometric function of random measurement error, which is equivalent to the ratio of the true score variance to the observed score variance (Weiner *et al*, 2003).

The Classic Test Theory (CTT) follows three set steps in attempt to establish the reliability of a measurement tool, depending on the source of inconsistency or measurement error that is relevant to a particular testing situation. There are four types of inconsistencies that CTT focuses on: (1) internal inconsistencies among items or tasks within a measurement tool; (2) inconsistencies over time; (3) inconsistencies across different forms of the measurement tool; and (4) inconsistencies within and across raters (Bachman, 2004). There are two general approaches to establishing reliability for these four measurement errors. The first approach involves calculating the correlation between two sets of scores. The second approach involves calculating the coefficient Alpha (α), which is based on the variances of the different sets of scores. The variances that are used are the unbiased estimates of the population variance (Bachman, 2004).

"The three set steps that are used within CTT to establish reliability are: (a) identifying the relevant source of measurement error; (b) designing a study to collect two sets of independent and parallel scores; and (c) estimating the reliability by calculating either the appropriate correlation coefficient or coefficient alpha" (Bachman, 2004, p. 160). To estimate the reliability of a measurement tool with internal inconsistencies among the items or tasks, CTT suggests the use of split half scores and item variances. For a measurement tool with inconsistencies over time, first the stability of the test scores over time need to be estimated and two sets of scores need to be attained. This is followed by calculating the correlation between the two sets of scores and interpreting it as test-retest reliability. Thereafter, a coefficient alpha is calculated (Bachman, 2004).

If there are inconsistencies across different forms of the measurement tool, the equivalence of the test scores across the different forms of the measurement tool need to be estimated. This is done by constructing two forms of the test that are equivalent in content and the measured abilities. The two tests are then given to the same group of test takers and the two sets of test scores are treated as parallel measures (Bachman, 2004). With regard to inconsistencies within and across raters, two rater's ratings would be different resulting in the one's ratings being higher than that of the other. These two sets of ratings would be highly correlated; however they would not be statistically parallel making the CTT unable to use them to estimate reliability (Bachman, 2004).

3.2. Research Questions

This study essentially involved the statistical analysis and interpretation of data collected by the research team working with the adapted version of the KDI, for the South African population. Three subtests were the focus of consideration: Locomotor function, Eye-Hand Coordination, and Executive Function. This study aimed to respond to the following specific questions, and proposes methods used to answer them.

- Do each of the three subtests have adequate internal reliability as indicated by the KR-20 scores?
- Is there consistency of the underlying dimensions present in the scale, as determined by the factor analysis?
- What is the difference between each subtest item score of the KDI with regard to Age, Region and Gender, as determined by the ANOVA?

3.3. Sampling

For the purpose of cost-effective data collection in a short period of time, this study was conducted in KwaZulu-Natal (language isiZulu), Western Cape (languages Afrikaans/isiXhosa) and the North-West Province (Setswana). The sites were all sites where the Child Development Research Unit (CDRU) researchers have existing research collaborations, and there were trained child development assessors familiar with the rigour required for research endeavours. At each of the three sites, a stratified purposive sample was used to access participants in each of the two age categories (6-18 months, 18-35 months). The rationale for selecting these two age categories is because this study aims to assess early cognitive development while the child is in the sensorimotor stage of cognitive development (0-2 years), and the preoperational stage (2-7 years). Dividing the age into two groups made it easier for the researcher to analyze the test scores according to these stages of development. A specific quota of male and female children from each age group, were selected from the three sites to create a balance in the sample (Kvalsvig *et al.*, 2009).

Stratification occurred in terms of type of geographic setting (urban, peri-urban/rural), age (6-18 months, 18-35 months), and gender (male, female). This stratification into "developmental niche areas" (urban, peri-urban/rural) was utilized in the previous work undertaken for UNICEF, where the time available for data

collection was also too limited to allow for more extensive sampling (Kvalsvig *et al.*, 2009). These "niche" areas were targeted to attain samples from four different socio-economic groups: peri-urban/rural (KwaZulu-Natal); township (North West); and informal settlement & middle-class (Western Cape). As this study aims to assess children from different socio-economic backgrounds, these niche areas will give the researcher an idea of how the socio-economic status of the child may contribute to their cognitive abilities (Kvalsvig *et al.*, 2009).

There were 10 male and 10 female participants in each age group per site. Participants were purposively selected in each category, dependent on ease of contact, and informed consent from each caregiver. The recruitment of the participants from the rural area was done on a door-to-door basis after permission from the tribal authorities was attained. Similarly in the informal settlement/township area, the community leaders were approached for permission before the door-to-door recruitment began. In the middle-class area, recruitment was initially done through pre-schools, thereafter through a snowball sample of the area. For each participant, the main caregiver was interviewed (Kvalsvig *et al.,* 2009). However, the data from the primary caregiver was not used in this study. The sample was classified into different categories (i.e. Male/Female; Age group – 6-18 months and 18-35 months, and Socio-economic Status).

The sample for this study is presented in Table 1 below. Three sites (KwaZulu-Natal, Western Cape, and North West Province) were used for the study. The total sample size was 120, which comprised 40 children from KwaZulu-Natal, 40 children from Western Cape and 40 children from North West Province. There were a total of 20 children per age group (6-18 months and 18-35 months). The intention was to identify at least 300 young participants in the specified age groups with a wide range of abilities, and administer the full battery of tests to them, as well as obtain biographical and living condition information from their parents/primary caregivers.

However, due to unavailability of other participants, the total number in this sample was 120 at the time that the data needed to be processed. The larger study will continue to boost sample size (Kvalsvig et al., 2009).

Table 1: Stratified Sample per Site (n = 40) (Kvalsvig *et al., 2009)*

KWAZULU-NATAL PROVINCE		NORTH WEST PROVINCE		WESTERN CAPE PROVINCE	
6-18months	18-35months	6-18months	18-35months	6-18months	18-35months
Male = 10	Male = 10	Male = 10	Male = 10	Male = 10	Male = 10
Female = 10	Female = 10	Female = 10	Female = 10	Female = 10	Female = 10
n = 20	n = 20	n = 20	n = 20	n = 20	n = 20

3.4. Scoring of the KDI

The KDI comprises three subtests: Locomotor functioning, Eye-Hand Coordination, and Executive Functioning (See Appendix D). The tasks that assess Locomotor functioning include: Moving, Walking, Lying, Sitting, Standing, Climbing, Hopping and Jumping. The tasks used to assess Eye-Hand Coordination include: Playing with the Ball, Vision (Ring & Red Tassel), Button, Block Tower, Container & Cubes, Coin Box, Bead Threading, and Paper & Pen. To assess Executive function, the A-NOT-B and Self Control tasks were used (Holding *et al.,* 2008).

The scoring of the KDI (See Appendix D) is as follows: for the *Locomotor* tasks (i.e. Moving, playing with the ball, and lying & standing), the score is either Yes (1) or No (0), which signifies whether the child can perform the task or not. For the *Eye-Hand Coordination* tasks (Vision: Ring & Red Tassel, Button, and Paper and Pen), the score is either Yes (1) or No (0). The first part of the Bead Threading task of the *Eye-Hand Coordination* subtest has three items that are scored as either Yes (1) or No (0). The second part of the Bead Threading subtest consisted of one item made up of three trials. A score in seconds is given, depending on how long the child took to complete the three trials in 30 seconds; and the number of threading the child is able to do is recorded. In the Block Tower and Container and Cubes tasks of the Eye-Hand Coordination subtest, the child gets a tick if he/she could perform each task. Thereafter the number of ticks is added up and a total score is given (Kilifi Developmental Inventory, 2008).

The first part of the Coin-Box task of the *Eye-Hand Coordination* subtest received a score similar to that of the Block Tower and Container and Cube tasks. For the second part of this task a score in seconds was given depending on how long the child took to complete the three trials. After a percentile analysis was conducted, an average time of 17.5 seconds was computed from the mean scores. If the child completed all three trials in 17.5 seconds or less, a score of 1 was given. A child received 0 if he/she completed the three trials in more than 17.5 seconds. For the Self Control task of the *Executive Function* subtest, a score was recorded in seconds, depending on how long the child took to perform the task. In the A-NOT-B task of the *Executive Function* subtest, the child had 10 trials in which he/she looked for the treat. If he/she found the treat, a tick would be given; if not a cross was given. Thereafter, the total number of correct attempts at finding the treat was added up to give an overall score. The number of incorrect attempts was also added up and a total score was recorded (Kilifi Developmental Inventory, 2008).

3.5. Preparing the data

The data collection process involved using the KDI inventory, which consists of multiple ways of measuring different cognitive developmental functions. The inventory consists of three subtests; locomotor, eye-hand coordination, and executive function. These different subtests consist of sets of items that have been shown to be reliable measures of locomotor control, eye-hand-coordination, and executive function. The tasks either required the data collector to rate success or failure on a specific item (0 = failure, 1 = success); or they required timing, or counting number of successes. In order to ensure that no particular items contributed too heavily to the final scores on subtests, these items were all converted into dichotomous scores of 0 = failure, 1 = success. This was required for items EH14-16, EH17-18 and the items from the executive function subtest.

EH14-16

These items involved tasks where specific actions done during that task needed to be coded as present. If these actions were present, the examiner ticked the box and then added up the number of ticks for each item. In order to convert these items into dichotomous measures, cut off points were included as indicated by the KDI Manual. These were said to approximate the 75th percentile of the different age groups (Kilifi Developmental Inventory, 2008). If a child of age 9 months scored 6 on item EH14 then the result was coded as 1 for a success. If a child of age 20 months scored 10 on EH15 it would be coded as 0 for a failure. Using this system the scores for items EH14-16 were coded as dichotomous variables. Table 2 below shows how these item scores were converted.

Table 2: Converted scores for items EH 14-16

Age-bands	EH14	EH15	EH16
6-8	4	1	2
9-11	6	6	2
12-14	7	10	3
15-17	8	11	3
18-20	8	13	3
21-23	9	13	6
24-26	9	13	6
27-29	10	13	6
30-32	10	13	6
33-35	10	13	6

EH17-18

Items EH17 and EH18 involved running three trials in which the infant was required to insert three coins into a box that had been rotated, using different hands for each trial. The amount of time it took the child to perform the task was measured and averaged over the three trials. A cut-off time of 17.5 seconds was used to

dichotomize the results, as per the recommendations set out in the KDI manual (Kilifi Developmental Inventory, 2008). For children who took longer than the cut-off time, a code of 0 was given for failure, and those that scored less than the cut-off time, a code of 1 was given for success.

The executive function subtest

The executive function subtest involved two items, which were self control tasks (SC1, SC2). These items measured how long the child can resist grabbing a treat in front of them, after being told not to take the treat. No recommendations on how to code these items were found in the KDI manual. There proved to be a great deal of missing data in these items making it difficult to estimate a defendable cut-off point. It was therefore decided that raw scores would be used. The A not B item in the executive function subtest involved repeated tests of whether the child could remember which cup the treat was under during different time delays. A potential 10 trials could be done and the examiner noted how many of these trials the child guessed correctly. If less than 10 trials were done, the amount of correct guesses were converted to be out of ten by dividing by the total trials done and multiplying by ten. This meant that the executive function subtest consisted of raw timed scores for the self control task and a score out of 10 for the A not B task.

Missing data

After the item scores had been converted it became apparent that certain items were missing large quantities of data. Three ways to deal with this were attempted, with varying results. Complete case analysis was conducted whereby the participants with data missing in the above items were not used in the analysis. This proved to be unsatisfactory because it threw away responses to other items, which contributed to the overall internal reliability of the scale. It also did not help if the missing data was systematically biased as this meant that the complete case analysis would be biased (Gelman & Hill, 2007). Simple random imputation of missing values is the next method that was used, which involved imputing random zeros and ones into the missing data values.

This enabled the items to be included in analysis and for greater use of the data to be achieved, however this ultimately proved unsatisfactory in that it affected the internal reliability of the scale.

When factor analysis was done on the subtests of these items, it revealed that they no longer loaded on the first factor, which defines the measure that the subtest is measuring. This is outlined in more detail in a later section. A random regression imputation (Gelman & Hill, 2007) was attempted but due to the amount of missing data and the lack of adequate predictors to model the data, the results were less than satisfactory. It was then decided that items EH17 and EH18 be removed from the scale, which increased the internal reliability of the scale and made it possible for all the items in the eye-hand coordination subtest to get loaded onto the first factor. Items SC1 and SC2 were kept and mean-coded, replacing the missing values with the mean score of the item. This was also unsatisfactory as will be seen when the reliability of the executive function scale is examined. Table 3 below presents an outline of how the missing data was dealt with.

Table 3: Management of missing data

Univariate Statistics

	N	Mean	Std. Deviation	Missing		No. of Extremes[a]	
				Count	Percent	Low	High
EH17	49	27.0816	29.00419	71	59.2	0	3
EH18	43	39.5814	41.10974	77	64.2	0	6
SC1	96	13.5000	31.59214	24	20.0	0	12
SC2	92	12.1196	28.27372	28	23.3	0	13

a. Number of cases outside the range (Q1 - 1.5*IQR, Q3 + 1.5*IQR).

3.6. Data Collection and Analysis

The leader of each ten-person data collection team from the three provincial sites and the data collection manager from each site were trained in Durban over a period of two weeks. The training included practice assessments with children in the appropriate age ranges at an ECD site. They then trained the other members of their teams under the supervision of a psychologist at each site (Kvalsvig *et al*., 2009, p.10). The data for this study was collected by the trained researchers from the larger study and the author of this thesis was not involved in the data collection process. Permission to have access to the collected data was received from the Principal Investigators (Dr. S. Kauchali and Dr. J. Kvalsvig).

The main aim of this study was to establish the internal reliability of the KDI. The five percent significance level was used for all statistical comparisons. The reliability coefficients for the age group 6-35 months were calculated across the three subtests of the KDI namely: Locomotor; Eye-Hand Coordination and Executive Function. The researcher used Kuder-Richardson (KR-20) in the SPSS Version 18.0 system to assess the internal consistency of the KDI scores. Kuder-Richardson (KR-20) measures the internal consistency of a test and is based on the proportion of correct and incorrect responses to each of the items on the test. The KR-20 was deemed suitable for this study as the data collected for the Locomotor and Eye-Hand Coordination subtest was dichotomous data. The data from the Executive Function subtest was interval, and therefore had to be converted to dichotomous data by setting a cut-off point, as explained in the previous section. The internal reliability of each subtest was conducted as well as a factor analysis to check whether the items loaded on one factor, which would indicate whether the items are measuring the same latent variable. In the subtests that were affected by missing data, results from different methods for dealing with missing data will be shown to justify the ultimate decision that was made.

Factor Analysis is a "statistical technique that is used to identify a relatively small number of factors that can be used to represent the relationship among sets of many interrelated variables." (Durrheim & Terre Blanche, 2002, p.362). A factor analysis was done to determine the consistency of the underlying dimensions present in the KDI. The factor analysis was conducted on the items of the three subtests; locomotor, eye-hand coordination, as well as executive function, to reveal if all the items were loading on the first factor. An Analysis of Variance (ANOVA) was conducted on the locomotor and eye-hand coordination subtests to assess whether there was a difference in the subtest item scores for Age, Region and Gender.

3.7. Ethical considerations

Since this study formed a part of a larger study, namely the NELDS Study; many of the direct ethical concerns were not negotiated directly by this author. Rather she relied on the members of the two research team's negotiations and carefully considered the methods through which they had ensured the rights of the individuals. By considering their written proposals and the forms they had used before gaining access to individual children, families and communities, the researcher was able to satisfy herself that the necessary ethical procedures had been adhered to.

Wassenaar's (2002) structure regarding ethical procedures required for research was explored. Using this structure, the following steps have been taken to ensure that this research was ethical and sensitive to the vulnerable child population on whom this study is based. Ethical issues arising from working with children in vulnerable settings include: informed consent from the parent/primary care-giver allowing their child to participate in the research, confidentiality, beneficence, nonmaleficence (Wassenaar, 2002). Children in this

research were considered to be vulnerable because they were very young; therefore, they were not able to provide informed assent.

In terms of assessing the degree of ethical adherence to Informed Consent, the researcher considered the rights of the participants by obtaining consent for participation. The NELDS research team submitted the draft informed consent forms and information sheets to the University of KwaZulu-Natal Humanities Research Ethics Committee for ethical approval. Due to the fact that the Informed Consent Forms had the personal details and signatures of the participants' parents, these were not included in the appendix. These documents contained the required assurances of confidentiality, voluntary participation, and freedom to leave the study without fear of penalty. The author of this thesis did not ensure ethical consent directly, but relied on the informed consent obtained by the NELDS Research team. The NELDS study obtained ethical clearance from the University of KwaZulu-Natal Ethics Committee (Kvalsvig *et al*, 2009).

The issue of confidentiality was also adhered to by the larger NELDS study. The assessors were reminded of the importance of never to discuss information obtained from the data collection process with anyone other than a member of the research team. The data was stored in a locked cupboard, and the electronic copies of the data did not contain names of the participants or information pertaining to the families (Kvalsvig *et al*, 2009). The subjects' names, contact details and addresses were not included in the data set that was given to the writer. Even so, various procedures were taken to protect anonymity of the subjects by using pseudonyms. The data set was password protected and only senior members of the research team had access to these passwords. The size of the database also required much of the data to be stored in encrypted form.

With regard to the ethic issue of non-maleficence, the research did not do harm to the participants or any group of persons being assessed, beyond taking up the time of the child and caregiver within their own household context. It did, however, aim to benefit the participants by referring those children needing intervention to relevant professionals (Wassenaar, 2002). The potential benefits were explained as not being directly related to the child or family, but rather being indirect in the sense that if the child did require further intervention after being assessed, referral for intervention was indicated and discussed. No payment of any sort was made to the child or family (Kvalsvig *et al*, 2009). It was also very important to gain the trust of the parent whose child was participating in the research. This required the research team to be honest, reliable, and also to communicate well in order to maintain a good relationship with the parents. Parents were also informed of their right to withdraw their child from the study at any point (Kvalsvig *et al*, 2009). This study was also authorized by the University of KwaZulu- Natal Faculty of Humanities, Development and Social Sciences (HDSS) Ethics Committee.

3.8. On Reflexivity

Being involved in this study has made me realize the important part I've played in this research, no matter how small. It has been an eye-opening experience and has taught me how to be aware of the ethical implications involved when working with research participants, especially children from a disadvantaged background. Working as a team with the staff from Valley of 1000 Hills, as well as colleagues from UKZN, has made this experience interesting as it has enabled me to be a part of a large scale national research study, and also to complete a Masters degree in Clinical Psychology. This study has contributed to a broader understanding of research tools, research design and how to be open-minded in thinking about the impact that the results of such a study has on African child populations.

Due to time constraints and an inability to be available on-site during the data collection period, I could not be directly involved in this process. I was therefore, not been able to have a feel for the entire research procedure or an understanding of all the possible problems that could have been encountered by the research staff while collecting the data. I am however, grateful that the research team was willing to provide me with access to the research data in order to complete my dissertation requirements.

The data analysis process was both intellectually stimulating and challenging. There was a lot of data to manage, and the cleaning procedure was quite time consuming. The missing values in the data exacerbated the challenges as some analytical procedures could not be done properly; however, these challenges were dealt with appropriately with the help of an experienced statistician, Mr. O. Bodhlyera, who is employed at University of KwaZulu-Natal, Pietermaritzburg, as a statistician; and Stephen Olivier, a Research Masters Student at University of KwaZulu-Natal, Pietermartizburg. Being part of the data collection process would have enabled the researcher to ascertain whether the missing values (i.e. no scores recorded) were due to zeros, or due to the infants not being cognitively competent to do the task as a result of their age (not being at the appropriate developmental age to complete the task) or as a result of cognitive impairment as a result of other factors such as disease, and inadequate opportunity to explore the environment and acquire those required skills. This is a serious limitation to the current research and although the results are based on the available data, it would have been very useful to access file notes to ascertain this. There are psychometric techniques that were used to manage the issue of missing data which were used during the analysis process, and these have been explained in the data preparation section of this study.

The write-up of this dissertation has taken a long time, as there have been many demanding factors during my internship year. However, weekends and afternoons were set aside to continue with it in order for a timeous submission. Overall, this research has made me more aware of the importance of establishing the reliability of assessment tool and assessing the psychometric properties of a scale, in order for it to elicit accurate and consistent results should it be administered on another population at a later stage.

CHAPTER 4

4. Results

In this chapter, the results from the analysis are presented and discussed. A five percent significance level was used for all statistical comparisons. Results were then compared with those from the larger study (NELDS), and the previously conducted study in Kenya. A discussion incorporating literature from other studies is included to describe the relevance of this study in relation to child development and psychological assessments in the African context.

Table 1: Descriptive Statistics describing the distribution of the sample according to gender

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Female	63	52.5	52.5	52.5
	Male	57	47.5	47.5	100.0
	Total	120	100.0	100.0	

The number of males and females selected in this sample are near even.

Table 2: Descriptive Statistics describing the distribution of the sample according to region

Region

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	KZN	40	33.3	33.3	33.3
	POTCHEFS	40	33.3	33.3	66.7
	WCI	20	16.7	16.7	83.3
	WCM	20	16.7	16.7	100.0
	Total	120	100.0	100.0	

Forty children were selected from KwaZulu-Natal, North-West Province (Potchefstroom), as well as Western Cape Province, bringing the total size of the sample to 120.

Table 3: The mean and standard deviations (SD) of the sample

Statistics

Age (Months)

N	Valid	120
	Missing	0
Mean		18.6583
Std. Deviation		8.78549
Range		30.00
Percentiles	25	11.0000
	50	17.5000
	75	25.7500

The children's age ranged from 6-36 months, with mean = 18 months, and Standard Deviation = 8.79

4.1. Internal Reliability

Table 4: Internal Reliability of the Locomotor Subtest

Reliability Statistics

KR-20 Alpha	N of Items
.942	35

Due to the binary nature of the items of the locomotor subtest, a Kuder-Richardson 20 (KR-20) test of internal reliability was done. The result was a value of alpha = 0.942. This compares favourably to the reported internal consistency score reported in the KDI manual of alpha = 0.92.

Table 5: Internal Reliability of the Eye-hand Coordination Subtest

Complete case analysis

Reliability Statistics

KR-20 Alpha	N of Items
.732	25

When using only the participants with complete scores in the Eye-hand Coordination (EH) subtest, the missing data in EH17 and EH18 ruled out a substantial number of participants and the subsequent internal reliability measured by the KD-20 indicated alpha = 0.732. The reliability of this subtest is less than satisfactory. This is probably because complete case analysis throws out many cases limiting the cancellation of error associated with longer tests.

Random Imputation

Reliability Statistics

KR-20 Alpha	N of Items
.890	25

Using a random imputation of missing values, the internal reliability has gone up to alpha = 0.89. This estimates that 89% of the variance is true variance. This means that the eye-hand coordination subtest of the KDI is internally reliable.

Table 6: Internal Reliability of the Executive Function Subtest

Reliability Statistics

KR-20 Alpha	N of Items
.594	3

The executive function subtest does not have adequate internal reliability as indicated by alpha = 0.594. This could be due to the fact that two of the items in this three item subtest had missing values which could not be modelled, and therefore were given mean imputations.

4.2. Factor Analysis

Locomotor Subtest

Factor analysis was done on the items of locomotor subtest to reveal if all the items were loading on the first factor. The factor analysis revealed that the first factor explained 35% of the variance. It also revealed that almost all the items in the locomotor subtest loaded on the first component of the factor analysis. The exceptions were LM01, LM24, and LM 27-28 (See Appendix A: SPSS outputs for Locomotor Factor Analysis). The fact that these components do not load on the first factor is not too worrying because of the high internal validity.

Eye-hand Coordination Subtest

Factor analysis was done on the items in the eye-hand coordination subtest to ascertain whether the items were loading on the first factor. This showed that the first factor explained 31% of the variance (See Appendix B:

SPSS outputs for eye-hand coordination Factor Analysis). It also showed that all the items loaded on the first factor except EH17 and EH18, which are the items that had the missing items imputed. This suggests that the imputation may have biased the values on these particular items. Seeing that these items had substantial missing data and did not contribute to the first factor latent variable, they were removed.

Removal of items EH17 and EH18

Reliability Statistics

KR-20 Alpha	N of Items
.910	23

The removal of items EH17 and EH18, has increased the alpha value to alpha = 0.91. This compares well with the reported internal reliability of alpha = 0.93 reported in the KDI Manual (Kilifi Developmental Inventory, 2008). Removal of the items not implicitly measuring the latent variable helps reduce error when using the findings from test takers.

Executive Function Subtest

The factor analysis indicates that all the items loaded on the first factor, which explains 58% of the variance within the items (See Appendix C: SPSS output for executive function factor analysis). Although there were only three items, it was difficult to know how to interpret the results especially with regards to the low internal reliability of the subtest. The items ABSC1 and ABSC2 correlated very well with each other, whereas the A-not-B item correlated only a little with these two items.

4.3. Analysis of Variance (ANOVA)

Locomotor Subtest

Univariate Tests

Dependent Variable: Total_LM

	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Contrast	3771.397	3	1257.132	79.605	.000	.729
Error	1405.500	89	15.792			

*The F tests the effect of age group. This test is based on the linearly independent pair-wise comparisons among the estimated marginal means.

The overall ANOVA shows that there are significant differences within the data $F = 79.605$, $p < 0.05$, Eta squared = .729. This means that there are significant differences in the data and a post hoc analysis needs to be done. (Eta squared is a measure of how big an effect size is)

Table 7: Tests of Between-Subjects Effects for Region, Age and Gender

Dependent Variable:Total_lm

Source	Type III Sum of Squares	df	Mean Square	F
Corrected Model	6458.092[a]	30	215.270	13.631
Intercept	28013.246	1	28013.246	1773.873
Region	307.057	3	102.352	6.481
Gender	.015	1	.015	.001
age group	3598.880	3	1199.627	75.964
region * gender	2.063	3	.688	.044
region * age group	553.134	9	61.459	3.892
gender * age group	45.368	3	15.123	.958
region * gender * age_group	64.675	8	8.084	.512
Error	1405.500	89	15.792	
Total	40303.000	120		
Corrected Total	7863.592	119		

a. R Squared = .821 (Adjusted R Squared = .761)

Dependent Variable:Total_lm

Source	Sig.	Partial Eta Squared
Corrected Model	.000	.821
Intercept	.000	.952
Region	.001	.179
Gender	.975	.000
age_group	.000	.719
region * gender	.988	.001
region * age_group	.000	.282
gender * age_group	.416	.031
region * gender * age_group	.845	.044

The test for between-subjects effects shows that the differences are within the independent variables region Eta^2 = .179, $p < 0.05$; age group Eta^2 = .719, $p < 0.05$; and their interaction Eta^2 = .282, $p < 0.05$. No significant differences were found in gender.

Table 8: Multiple Comparison Test for Region

Multiple Comparisons

Total_lm
LSD

(I) region	(J) region	Mean Difference (I-J)	Std. Error	Sig.
KZN	POTCHEFSTROOM	-3.0000*	.88860	.001
	WCI	-3.7250*	1.08831	.001
	WCM	-3.2750*	1.08831	.003
POTCHEFSTROOM	KZN	3.0000*	.88860	.001
	WCI	-.7250	1.08831	.507
	WCM	-.2750	1.08831	.801
WCI	KZN	3.7250*	1.08831	.001
	POTCHEFSTROOM	.7250	1.08831	.507
	WCM	.4500	1.25667	.721
WCM	KZN	3.2750*	1.08831	.003
	POTCHEFSTROOM	.2750	1.08831	.801
	WCI	-.4500	1.25667	.721

Based on observed means.
The error term is Mean Square (Error) = 15.792.
*. The mean difference is significant at the .05 level.

The multiple comparison test shows that the only significant difference in regions is found between KZN and the other regions; where KZN has a mean of 14.2750, $p < 0.05$; Potchefstroom has a mean of 17.2750; WCI has a mean of 17.7273, $p < 0.05$; and WCM has a mean of 17.5500, $p < 0.05$.

Table 9: Multiple Comparisons Test for Age group

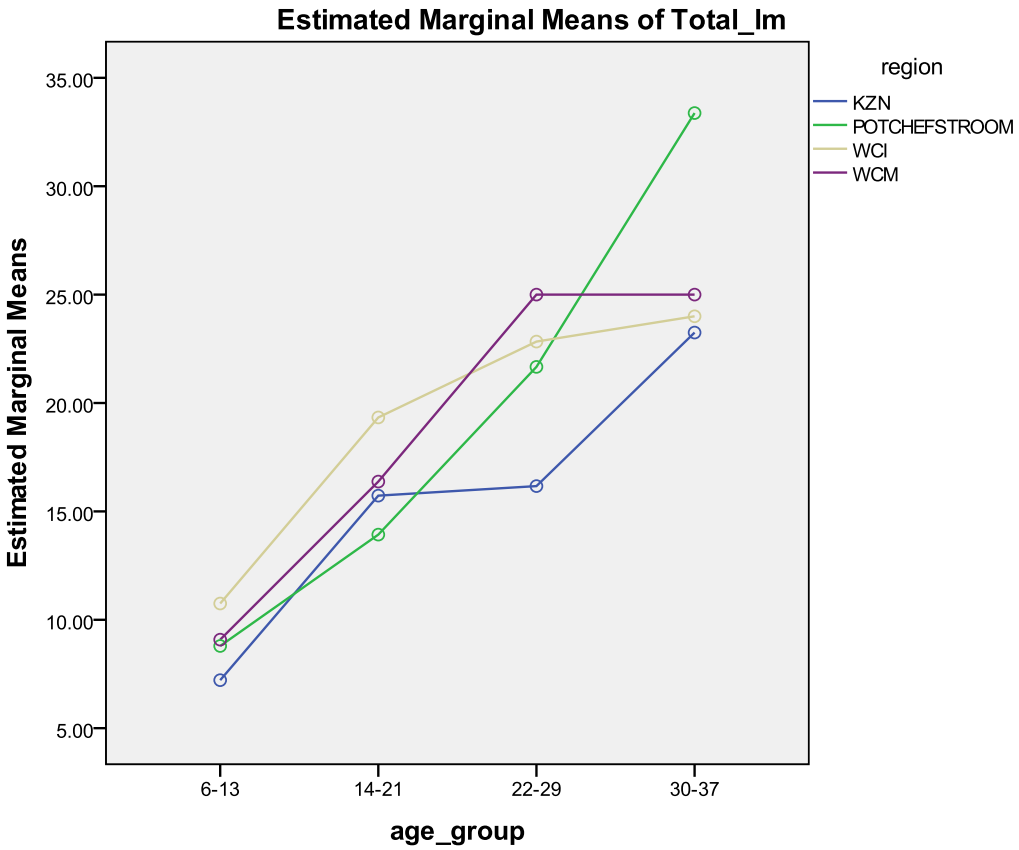
Multiple Comparisons

Total_lm
LSD

(I) age_group	(J) age_group	Mean Difference (I-J)	Std. Error	Sig.
6-13	14-21	-7.0194*	.91295	.000
	22-29	-12.5333*	1.02606	.000
	30-37	-18.9250*	1.08831	.000
14-21	6-13	7.0194*	.91295	.000
	22-29	-5.5139*	1.04722	.000
	30-37	-11.9056*	1.10828	.000
22-29	6-13	12.5333*	1.02606	.000
	14-21	5.5139*	1.04722	.000
	30-37	-6.3917*	1.20317	.000
30-37	6-13	18.9250*	1.08831	.000
	14-21	11.9056*	1.10828	.000
	22-29	6.3917*	1.20317	.000

Based on observed means.
The error term is Mean Square(Error) = 15.792.
*. The mean difference is significant at the .05 level.

The multiple comparisons for Age Group show significant differences in all age groups to the alpha level of 0.05. The 6-13 months Age Group mean = 8.9286; the 14-21 months Age Group mean = 15.6944; the 22-29 months Age Group mean = 21.2083; and the 30-37 months Age Group mean = 27.6.



This plot shows an increase in the locomotor subtest scores with Age, and the different scores for the three regions that participated in the study. It also shows differences between the Regions, and the interaction effect between the Age and Region. Potchefstroom (North West Province) in particular changes places drastically as the Age Group increases.

Eye-Hand Coordination

Table 10: Tests of Between-subjects Effects

Tests of Between-Subjects Effects

Dependent Variable:Total

Source	Type III Sum of Squares	df	Mean Square	F
Corrected Model	2670.000[a]	30	89.000	7.326
Intercept	16506.410	1	16506.410	1358.783
Region	549.772	3	183.257	15.085
Gender	1.461	1	1.461	.120
Age_Group	1082.937	3	360.979	29.715
Region * Gender	4.999	3	1.666	.137
Region * Age_Group	244.817	9	27.202	2.239
Gender * Age_Group	6.660	3	2.220	.183
Region * Gender * Age_Group	28.330	8	3.541	.292
Error	1081.167	89	12.148	
Total	22752.000	120		
Corrected Total	3751.167	119		

a. R Squared = .712 (Adjusted R Squared = .615)

Tests of Between-Subjects Effects

Dependent Variable:Total

Source	Sig.	Partial Eta Squared
Corrected Model	.000	.712
Intercept	.000	.939
Region	.000	.337
Gender	.730	.001
Age_Group	.000	.500
Region * Gender	.938	.005
Region * Age_Group	.026	.185
Gender * Age_Group	.908	.006
Region * Gender * Age_Group	.967	.026

The overall ANOVA shows differences between Region, where Eta squared = 0.337, $p < 0.05$. Age group shows significant differences where Eta squared = 0.5, $p < 0.05$. The interaction between Age group and Region also shows significant differences, where Eta squared = .185, $p < 0.05$. No significant differences were found for Gender.

Table 11: Multiple Comparison Test for Region

Multiple Comparisons

Total
LSD

(I) Region	(J) Region	Mean Difference (I-J)	Std. Error	Sig.
KZN	POTCHEFSTROOM	-4.6250*	.77936	.000
	WCI	-5.9750*	.95451	.000
	WCM	-4.9250*	.95451	.000
POTCHEFSTROOM	KZN	4.6250*	.77936	.000
	WCI	-1.3500	.95451	.161
	WCM	-.3000	.95451	.754
WCI	KZN	5.9750*	.95451	.000
	POTCHEFSTROOM	1.3500	.95451	.161
	WCM	1.0500	1.10218	.343
WCM	KZN	4.9250*	.95451	.000
	POTCHEFSTROOM	.3000	.95451	.754
	WCI	-1.0500	1.10218	.343

Based on observed means.
The error term is Mean Square (Error) = 12.148.
*. The mean difference is significant at the .05 level.

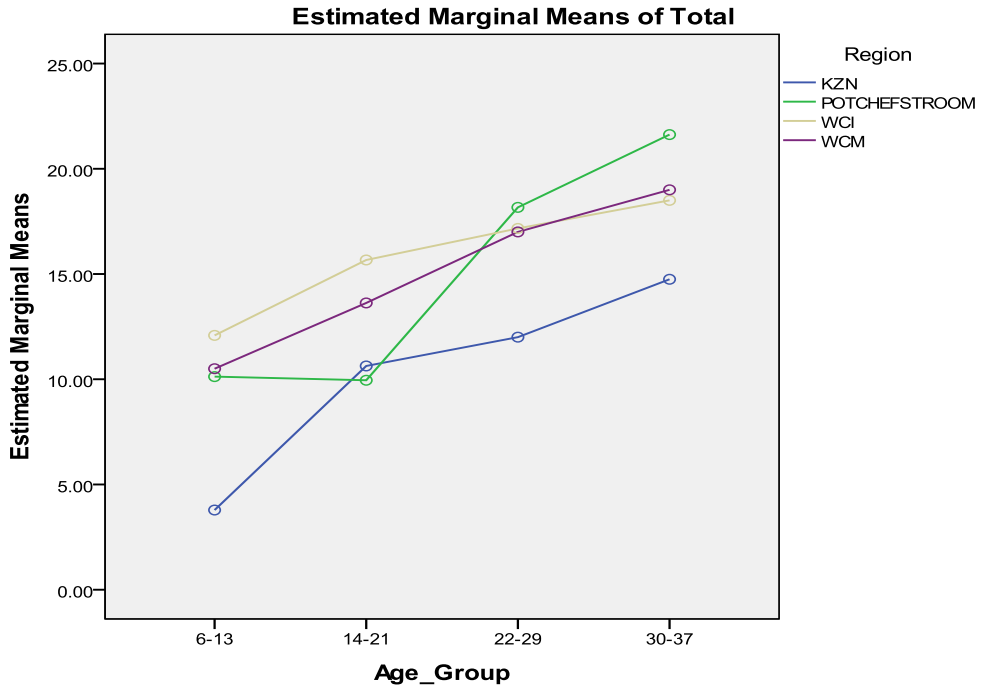
The multiple comparisons between different Regions show that KZN mean = 9.225, which is significantly different from the rest of the regions at $p < 0.05$. Potchefstroom mean = 13.85, WCI mean = 14.9091; and WCM mean = 14.15.

Table 12: Multiple Comparison for Age group

(I) Age_Group	(J) Age_Group	Mean Difference (I-J)	Std. Error	Sig.
6-13	14-21	-3.4250*	.80071	.000
	22-29	-7.8417*	.89992	.000
	30-37	-9.9750*	.95451	.000
14-21	6-13	3.4250*	.80071	.000
	22-29	-4.4167*	.91848	.000
	30-37	-6.5500*	.97203	.000
22-29	6-13	7.8417*	.89992	.000
	14-21	4.4167*	.91848	.000
	30-37	-2.1333*	1.05525	.046
30-37	6-13	9.9750*	.95451	.000
	14-21	6.5500*	.97203	.000
	22-29	2.1333*	1.05525	.046

Based on observed means.
The error term is Mean Square(Error) = 12.148.
*. The mean difference is significant at the .05 level.

The multiple comparisons between the different Age Groups show significant differences between all of the Age Groups at $p < 0.05$. The 22-29 months Age Group mean = 16.17, the 30-37 months Age Group mean = 18.3, the 6-13 months Age Group mean = 8.325, and the 14-21 months Age Group mean = 11.75.



The means plot above shows the increasing of Eye-Hand coordination scores with Age and the differences between the three regions. The interaction between Region and Age group is less apparent in this graph than the previous one, but is still there.

4.4. Discussion: Results from the current study

As mentioned earlier in the thesis, the questions that this study sought to answer are:

- Do each of the three subtests have adequate internal reliability as indicated by the KR-20 scores?
- What is the relation between each subtest score of the KDI with regard to Age, Gender and Region?
- What are the underlying factors elicited from the factor analysis, and what is the relationship between these variables?

Each subtest of the KDI was analyzed separately to assess the internal reliability. The results from the data analysis show that the Locomotor subtest of the KDI had a good internal reliability when used on the South African population that was selected for this study. What this means is that the items are reliably measuring the same latent variable. In other words, when one item is answered one way, the item that should theoretically be responded to in one way should be reliably answered in this way, suggesting that the items are internally consistent. The Eye-hand coordination subtest scores had missing values in the data which affected the internal reliability of this subtest. However, after using the random imputation procedure, which removed the unreliable items (E17 and E18) the internal reliability of the Eye-Hand coordination subtest improved. The unreliability of these items is most likely due to the missing data, and future research could show that they are useful items which can measure the latent variable of Eye-Hand coordination.

With regard to the Executive Function subtest, the results show that this subtest has low internal reliability, and therefore requires further investigations with regard to the items used to assess this cognitive function. In light of the low internal reliability of this subtest it was not used when looking at the differences between Regions, Age groups and Gender in the ANOVA. This is because the error associated with an unreliable measure rules out the possibility of making sound inferences about differences between Groups. Only two subtests of the KDI that have shown to be internally consistent to the level of alpha = 0.942 and alpha = 0.91, namely the Locomotor and Eye-Hand coordination subtest respectively, were used in the ANOVA. The Executive Function subtest has proven to be unusable due to missing values and a low internal reliability.

The results from the factor analysis suggest that the first factor in the Locomotor subtest explained 35% of the variance. One could speculate that the underlying factor in this subtest is static and dynamic balance. It also revealed that almost all the items in this subtest load on the first component of the factor analysis, with LM01, LM24, and LM 27-28 as exceptions. The fact that these components do not load on the first factor is not too worrying because of the high internal reliability of the subtest. If in a repeated test of the reliability of the KDI these items showed similar results, a better case could be made for these items to be unsuitable for the South African context.

With regard to the Eye-hand Coordination subtest, the first factor explained 31% of the variance. The underlying factor could be fine motor skills that involve coordinating eye and hand movements. It also showed that all the items load on the first factor except EH17 and EH18, the items that had their missing items imputed. This tells us that the imputation may have been biased to the values on these particular items. Seeing that these items had substantial missing data and did not contribute to the first factor latent variable, they were removed.

For the Executive Function subtest, the factor analysis showed that all the items load on the first factor, which explains 58% of the variance within the items. The underlying factor could possibly be working memory and inhibition. Because there are only three items in this subtest, it was difficult to know how to interpret the analysis especially with regards to the low internal reliability. The items ABSC1 and ABSC2 correlate very well, where as the A not B correlates only a little with these two items.

The ANOVA results suggest that there were differences found in the two subtests (Locomotor and Eye-hand coordination) with regard to Regions and Age groups, however there were no differences for Gender. For the locomotor subtest, the participants seemed to perform better the older they were, which is expected according to the normal pattern of cognitive development. However, the interesting aspect is that the participants from KZN and Potchefstroom (North-West Province) performed differently to those from Western Cape Province. The scores of the participants from KZN increased as the age increased, however, not in a linear fashion as seen in the scores of the participants from NW. Instead, the scores for the participants between the age group 14-21 months and 22-29 months did not increase as expected, which raises concern.

One could speculate that the participants in KZN performed poorly as compared to those from NW, and there could be a number of reasons why this was the case. The participants sampled from KZN were mainly from rural areas with poor resources and adverse circumstances, which could give an idea of the factors that could have negatively affected the participants' performance. The participants sampled from NW were mainly from urban areas and had better access to resources; therefore this may have had a positive effect on their performance. The participants from WC were sampled from peri-urban areas, which may also have adequate

access to resources. This could have affected the participants in a positive way, but not in the same manner as that of NW.

With regard to the Eye-hand coordination subtest, there were differences between Age and Region, but also no differences for Gender. The participants from WC and NW performed well in this subtest as indicated by an increase in scores as the age increased, which is also expected according to the normal pattern of cognitive development. However, the scores for the participants from NW between the age group 6-13 months and 14-21 months appear not to have increased, which raises some concern especially because the sample was taken from urban areas. One would presume that children from an adequately resourced environment should perform better, however this does not seem to be the case. One would have expected a similar increase in scores as indicated by the locomotor subtest scores. With regard to the scores from KZN, they were lower than those of the other regions, and this could be due to factors related to socio-economic status of the area the participants were selected from. The poorly resourced area in KZN could have had a negative effect on the performance of the participants, resulting in lower scores, even though there is a gradual increase as the age increases.

4.5. Strengths and Limitations of this study

One of the KDI's greatest limitations is that it does not give an explanation to the possible reasons for the high number of missing data. One could speculate that the missing values were due to the participants from the different areas not being able to complete the tasks as a result of being young. The researcher did attempt to look for a pattern in the missing data, and thought that perhaps the problem could have been due to the examiners not recording the scores appropriately for some of the tasks, which affected the overall reliability of the data itself. The researcher also thought that perhaps the missing values in the data were due to non-completion of tasks as a result of withdrawal from the study, or as a result of cognitive impairment or poor

performance due to socio-cultural factors. Due to the fact that the KDI had inadequate internal reliability as an assessment tool, further investigations and modifications of the items need to be undertaken.

With regard to the results attained from the data, it is clear that the translated versions of the KDI have some good qualities of a research tool. As it was largely a non-verbal assessment tool, this aspect makes it appropriate for the cultural population sample that it was used on and this enhances its strengths. Even though some of the subtests warrant further investigation (i.e. EF subtest) and require additional tasks to be administered in order to measure a wider range of Executive Function skills, it was suitable for the population between the ages of 6– 35 months.

CHAPTER 5

5.1. Discussion: Results from the literature

The larger study taking place in KwaZulu-Natal, North West Province, Western Cape; namely the National Early Learning Standards Age Validation study (NELDS), and the INSTAPA Work Package 6 Study, elicited very interesting results. The KDI was one of the three developmental measures devised independently of the NELDS programme, and it was included in this study to give alternative ways of assessing developmental standards.

The KDI was administered to children in the 6 – 17 month and 18 – 35 month categories, and measured Locomotor skills, Eye Hand Co-ordination and Self-control function (Kvalsvig *et al.*, 2009, p. 11). In addition to this assessment, The Language Test and The Hand Movements Task were administered to supplement the results from the KDI data. The Language Test was designed for and administered on participants between the ages of 36 – 48 months. The Hand Movements task measures working memory and is only suitable for the 36-48 months age group. Even though it was conceptually unrelated to any of the NELDS items, it still managed to give a measure of neurological functioning (Kvalsvig *et al.*, 2009).

Results from the larger NELDS show that for both the Locomotor (LM) and Eye-Hand (EH) Coordination scores there was a "significant difference between the scores for the rural area and the other three areas, with lower scores from the rural area. The LM and EH Coordination subtest scores were different in that, there were more infants who completed the LM tasks as opposed to the EH tasks. Even though both these subtests measure similar constructs (movement), the one is specific to gross motor (LM) and the other is more specific to fine-

motor (EH). The NELDS was able to analyze the difference between the scores with regard to area, and this enabled the researcher to understand the impact that areas/locations have on cognitive development and abilities. However, the study under focus was only able to ascertain that both subtests (LM and EH) have adequate internal reliability.

In the self-control trials, the middle class area had a higher score than the other three areas. "These differences cannot be tied directly to the effect of the environment because the three data collection teams were moderated separately, but they nevertheless suggest that the effects of the environment are present and measurable before the age of three" (Kvalsvig *et al*, 2009, p.11-12). With regard to the EF subtest which measures self control, working memory and goal-directed behaviour, the results showed that the subtest itself had inadequate reliability, and therefore inaccurate results could have been elicited.

As the differences between the middle class area and the other three areas could not be attributed to one specific factor, this problem could be attributed to the subtest's inadequate reliability. Other factors that could contribute to these differences could be that the infants from middle class areas are familiar with the testing process; have sufficient support from their environment in stimulating cognitive development (i.e. interactions between caregiver & infant through play and education), however the researcher cannot be certain of this. Therefore, this subtest requires further investigation in order to improve its qualities, for the above-mentioned reasons.

The results from the Language Test show that there were significant effects that language had on age and place; and the rural areas had lower scores for this particular test. The study under focus did not use this test during the

analysis process as the target age-group was between 6-35 months, and full language development has not been established. However, when looking at the impact of language development on the over-all cognitive development of a toddler, one would be able to see how language deficits in the age group 36- 48 months could signify significant neurological deficits.

According to Skinner, children should have mastered a vast vocabulary and the use of grammatical constructions by the age of 4 – 5 years. They have an innate ability to acquire language through operant conditioning (reinforcement and imitation) from as young as 12 months of age, with the help and responsiveness of the caregiver (Berk, 2003). It could therefore be suggested that infants from rural areas may not receive sufficient responsiveness from caregivers to assist in language development. This however, does not mean that with adequate support from caregivers, the community and early learning practitioner, children will not improve on these language competencies. The ability to engage in language is a competency that keeps getting better, according to the *zone of proximal development* from Vygotsky's Theory of Development (Vygotsky, 1933/1979). In cognitive development, like all other aspects of child development, if the skill has not been established at that particular age, it could mean that it is due to it being a new skill; or the skill itself is still in its emerging stage (Kvalsvig *et al*, 2009).

From the Hand Movement task, results showed that the rural areas had the lowest scores while the townships scored the highest for this task. The aim of using these three independent tasks was to show that there are age related skills and competencies that differ markedly from area to area, even within South Africa (Kvalsvig *et al*. 2009).

The results from the Parent Rating Scale suggest that there was a poor performance on the desired result 5: Children are Learning About Mathematical Concepts. Suggestions for the inclusion of some items and the exclusion of others were made. A second round of analysis was done on the Parent Rating Questionnaire for the three age groups (6 – 17 months, 18 – 35 months and 36 – 49 months), and were separated into niche area (rural, township, middle class and informal settlement) as well as Gender.

Results suggest that there were "considerable variations across the niche areas in the performance of the tasks, with the children from the rural areas having the most difficulty" (Kvalsvig *et al*. 2009, p.12). Another interesting aspect was that the girls tended to perform better than boys on the Social and Communication Skills. Boys on the other hand, performed better on the Mathematical Concepts. This could suggest that boys are more mathematically inclined, whereas, the development of language and social interaction is established at a higher rate for girls, when looked at from a cognitive developmental perspective.

Research in Kenya also elicited some interesting results from their study which assessed the effects that HIV/AIDS infection has on cognitive development, with specific reference to Executive Functioning (EF). Because there were no studies showing the effects of HIV infection in infants younger than 4 years of age, previous studies have focused on children between ages 4 – 17 years. According to many reports related to EF in infants younger than 4, it is believed that EF is only really fully established by the age of 4, hence no research has been done in this regard (Holding *et al*, 2004). In addition, subtests that assess EF in measurement tools do not include sufficient tasks to measure EF in infants younger than 4 years adequately. According to other studies, it is reported that EF in healthy (HIV non-infected) children as young as 6 months can be assessed because by this age they manifest abilities to use their working memory, have inhibition capacities and goal-

directed behaviour. These skills continue to improve with age, suggesting a rapid growth of prefrontal brain regions (Holding *et al*, 2008).

With a focus on EF in HIV/AIDS infected children (assessed with the KDI), the results from the Kenya study aimed to show that infants infected and affected by HIV have cognitive impairments, specifically EF deficits, when compared to infants from the reference group. The assessment was done on infants between the ages of 6-35 months from the Kilifi District in Kenya, who spoke Kiswahili. The EF subtest (A-NOT-B) was used to assess working memory. Two variables were measured: *total correct trials,* which were the sum of all the trials in which the infant selected the correct location; and *perseveration error,* which was the total number of times that child continued to select the same side after the object had been removed to the other side. The EF subtest (Self Control) was used to assess inhibition (Holding *et al*, 2008).

Table 13: Means (and Standard Deviations) of the Various Background Variables
SD-Standard Deviation (Holding *et al*, 2008, p.13).

Variables	Reference population	HIV exposed	HIV infected	*P*
Sample size [girls]	319 [159]	17 [6]	31 [14]	
Age (*SD*)	18.84 (8.43)	17.72 (8.77)	21.10 (8.86)	.300
Maternal education (*SD*)[a]	3.40 (3.50)	4.35 (4.07)	4.80 (3.55)	.069
Weight-for-Age (*SD*)	-1.24 (1.08)	-1.28 (.79)	-2.12 (1.36)	.000

Table 13 above shows that for both Working Memory tasks, children who are HIV infected scored significantly lower than the reference population, as indicated by the mean. Even though there were significant differences between the two groups, these differences were relatively small. The HIV (exposed) but un-infected group did not differ significantly to the reference group, for both working memory tasks, however they did find it difficult to maintain self control/inhibition (Holding *et al*, 2008).

The study under focus did not specifically assess infants who were HIV infected. The sample was taken from populations who came from adverse environments plagued by an array of factors (HIV, malnutrition, poverty, and socio-economic status). All these factors contribute to possible deficits in the cognitive development of a child, in as much as HIV infection does. The focus on HIV infected children does however alert the community to the debilitating impact that this disease has on cognitive development.

Table 14: Descriptive Statistics for the different groups in the study (Holding *et al*, 2008)

Task	Reference population					HIV infected			HIV exposed		
	N	*M*	*SD*	*L. B*	*U. B*	*N*	*M*	*SD*	*N*	*M*	*SD*
Inhibition	198	0.00	0.97	-0.14	0.14	14	0.00	0.98	9	-0.52*	0.23
Working Memory											
Total correct	265	-0.00	0.98	-0.12	0.10	19	-0.18*	0.94	11	0.02	1.34
Perseverative errors	250	-0.00	0.98	-0.11	0.13	18	-0.18*	0.94	11	0.00	0.89

Table 14 above shows an important observation, noted from the data. It seems that a high percentage of infants were not able to reach the baseline for the task. A chi-square test was done to elicit reasons for non-completion of tasks for both HIV-infected (39%) and HIV un-infected (35%) infants. The results showed that both groups scored significantly high for non-completion of tasks than the reference population. A Zero-inflated Poisson (ZIP) regression model was used to assess factors that contributed to non-completion of tasks. Results suggest that HIV Status, Weight for Age, and Chronological Age were the main factors associated with non-completion.

This could suggest that infants who were exposed to HIV or whose health was compromised by disease and malnutrition, resulting in poor weight for their age, had difficulty regulating their emotions and perseverated during difficult tasks (Holding *et al*, 2008). This shows the importance of assessing children from adverse environments who encounter difficulties in their childhood, so that early detection of cognitive impairment can be made. These assessments have managed to bridge this gap, even though amendments still have to be made on certain aspects of the measurement tools to improve their quality in detecting cognitive deficits.

CHAPTER 6

6. Conclusion

This study has focused on the Kilifi Development Inventory (KDI), a measurement tool that is able to detect early cognitive impairment in children between the ages of 6 – 35 months. The questions that this study sought to answer were: Do each of the three subtests have adequate internal reliability as indicated by the KR-20 scores? What is the relation between each subtest score of the KDI with regard to age, gender and region? What are the underlying factors elicited from the factor analysis. Some of the subtests of the KDI assessment tool, namely Locomotor and Eye-Hand Coordination had adequate internal reliability, whereas the Executive Function subtest had inadequate internal reliability. Therefore, this assessment tool as a whole still requires some investigation as its internal reliability is not as adequate as expected.

The study under focus assessed a specific age group (6–35 months), while the other studies also looked at the older age groups. The ANOVA results suggest that there are differences between each subtest score of the KDI with regard to Age and Region, however there were no differences with regard to Gender. The older the child becomes, the better they perform in the tasks. Participants from the Western Cape and North West Province seemed to score higher as their age increased, while participants' scores from KwaZulu-Natal did not increase as expected, when their ages increased. Factors such as HIV infection/exposure and disease; malnutrition; poverty; limited access to healthcare services or education facilities; and poor social support are factors that may have a impacted negatively on the performance of these participants.

The next step towards the adaptation of the KDI, after further modifications to subtest items are made, would be to age validate the NELDS by undertaking a national and international literature and policy review on similar programmes. Thereafter, the research team will determine whether the indicators in the standards provided by the NELDS are appropriate for the 0-4 years age group, before a scientifically appropriate assessment tool is designed, which would be based on the developmental indicators provided by the NELDS for children from this age range. The research team will ascertain whether the indicators proposed for each age group are valid, and whether the validity holds across all child-rearing environments. Thereafter, recommendations to the appropriate government departments will be made on adjustments to the NELDS, in terms of age validity and socio-demographic factors. The researcher hopes that this study has assisted the reader in understanding the importance of assessing the psychometric properties of a psychological assessment tool in order to establish its reliability and elicit accurate results; as well as the need for standardized psychological assessments to enable early detection of cognitive deficits.

References

Abubakar, A., Holding, P., Van de Vijver, F., Van Baar, A., Newton, C., & Andrews Espy, K. (2008a). Effects of Prenatal HIV Exposure on Early Executive Functions of Infected and Uninfected Children in Kilifi, Kenya. *Tropical Medicine and International Health,* 13, 1-8.

Babbie, E. (2001). *The Practice of Social Research* (9th Ed). Wadsworth/Thomson Learning: USA

Bachman, L.F. (2004). *Statistical Analyses for Language Assessment.* Cambridge University Press.

Balnaves, M., & Caputi, P. (2001). *Introduction to Quantitative Research Methods: An investigative approach.* Sage Publication: London.

Banich, M. (2004). *Cognitive Neuroscience and Neuropsychology* (2nd Ed). Houghton Mifflin Company:Boston.

Berk, L. (2003). Theory and Research in Child Development (6th ed). *Child Development*. Illinois State University.

Bukatko, D., & Daehler, M.W. (1995). *Child Development: A Thematic Approach* (2nd ed). Houghton Mifflin Company:Boston.

Cicchetti, D., & Rourke, B.P. (2004). *Methodological and Biostatistical foundations of Clinical Neuropsychology: Studies on Neuropsychology, Development and Cognition* (2nd ed).Taylor and Francis Group: London.

Dana, R. (1993). *Multicultural Assessment Perspectives for professional Psychology.* Allyn and Bacon Boston.

Dreyfuss, M.L., Msamango, G.I., Spiegelman, D., Hunter, D.J., Urassa, E.J., Hertzmark, E., & Fawzi, W.W. (2001). Determinants of Low Birth Weight among HIV-infected pregnant women in Tanzania. *American Journal of Clinical Nutrition,* 74, 814-826.

Durrheim, K., Terre Blanche, M. (2002). *Research in Practice: Applied Methods for the Social Sciences.* University of Cape Town Press.

Edwards, S. (2005). Constructivism does not only happen in the individual: Socio-cultural Theory and Early Childhood Education. *Journal of Early Child Development and Care,* 175, 37-47.

Foxcroft, C., & Roodt, G. (2005). Psychological Assessment. *An Introduction to Psychological Assessments in the South African Context,* (2nd ed). Oxford: SA.

Gelman, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models.* New York: Cambridge University Press.

Grigorenko, E.L., & Sternberg, S.J. (1998). Dynamic Testing: The nature and measurement of learning potential. *Psychological Bulletin,* 124, 75-111.

Herbert, M. (2003). *Typical and Atypical Development: From Conception to Adolescence.* BPS Blackwell.

Holding, P., Abubakar, A., Kalu, R., Mbonani, L., Kabunda, B., Obiero, E., & Katana, K. (2008). *Kilifi Developmental Inventory (KDI)*, (pp. 1-40)

Holding, P., Taylor, G., Kazungu S., Mkala, T., Gona, J., Mwamuye, T., Mbonani T., & Stevenson, J. (2004). Assessing cognitive outcomes in a rural African population: Development of a Neuropsychological Battery in Kilifi District, Kenya, *Journal of the International Neuropsychological Society*, 10, 246-260. Cambridge University Press. (Electronic version)

Humphrey, J. H. (2004). *Childhood Stress in Contemporary Society.* New York: Oxford.

Jaramillo, J. (1996). Vygotsky's Socio-cultural Theory and Contributions to the Development of Constructivist Curricula. *Journal of Education,* 117, 113 – 117.

Kauchali, S. (2008). The Asenze Research Study: *Health and Psychosocial Need of Children with Development Disorders in a Time of HIV – an eco-social epidemiologic approach to the study of developmental disorders in a resource-poor setting.* (Electronic version)

Keller, H. (2002). Culture & Development: Developmental pathways to individualism and interrelatedness. In W.J. Lonner, D.L. Dinnel, S.A. Hayes, & D.N. Sattler (Eds). *Online Readings in Psychology and Culture* (Unit 11, Chapter 1). http://www.ac.www.edu/~culture/keller.htm

Kopp, C.B. (1994). Infant Assessment. In C.B. Fisher & R.M. Lerner (Eds). *Applied Developmental Psychology.* McGraw-Hill Inc.

Kvalsvig, J., Govender, K., & Taylor, M. (2009). *Research Report on the Age Validation of NELDS related to the Cognitive Development of Children between 0 – 4 years*. Unpublished.

Kwate, N.O.A. (2001). Intelligence or Misorientation? Eurocentrism in the WISC-III. *Journal of Black Psychology,* 27, 221-228.

Lee, K. (2000). *Childhood Cognitive Development: The Essential Readings.* Blackwell Publishers.

Lee, K., & Johnson, A.S. (2007). Child Development in Cultural Contexts: Implications of Cultural Psychology for Early Childhood Teacher Education, *Journal of Early Childhood Education,* 35, 233-243.

Lee, V., & Das Gupta, P. (1995). *Children's Cognitive and Language Development.* Blackwell Publishers.

Lonner, W.J., & Malpass, R.S. (1994). *Psychology and Culture*. Allyn and Bacon:Boston.

Manegold, J. (2002). *The Importance of Caregiver-Child interactions for the survival and Healthy Development of Children: Implications for Intervention.* UKZN:PMB.

McCarthy Gallagher, J., & Reid K. (2002). *The Learning Theory of Piaget and Inhelder.* Brooks/Cole Publishing Company: USA.

Newman, I., & Benz, C.R .(1998). *Qualitative-quantitative Research Methodology: Exploring the Integrative Continuum.* Library of Congress Cataloging-in-Publication Data.

Reaves, C. (1992). *Quantitative Research for the Behavioural Sciences*. John Wiley & Sons Inc.

Sadock, B.J., & Sadock, V.A. (2007). Contributions of the Psychosocial Sciences, in *Kaplan & Sadock's Synopsis of Psychiatry* (10th Ed). Lippincott Williams & Wilkins.

Schiele, J.H. (1991). An Epistemological Perspective on Intelligence Assessments among African-American children. *Journal of Black Psychology,* 17, 23-26.

Schulte, A. (2004). Has Dynamic Testing lived up to its potential? A review of Grigorenko and Sternberg's Dynamic Testing. *School Psychology Quarterly,*19, 88-92.

Singh, K. (2007). *Quantitative Social Research Methods.* Sage Publications: India.

Thelen, E. (2000). Three Month Old Infants can learn Task-Specific Patterns of Interlimb Coordination. *Childhood Cognitive Development*. Blackwell Publishers.

Tzuriel, D. (2001). *Dynamic Assessment of Young Children.* New York: Plenum Publishers.

Wassenaar, D. (2002). *Research in Practice: Applied Methods for the Social Sciences.* University of Cape Town Press.

Weiner, I., Freedheim, D., Graham, J., Schinka, J., & Naglieri, J. (2003). *Handbook of Psychology: Assessment Psychology* (Vol 10). John Wiley & Sons Inc.

APPENDIX A – SPSS OUTPUT-FACTOR ANALYSIS FOR LOCOMOTOR SUBTEST

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.861
Bartlett's Test of Sphericity	Approx. Chi-Square	2976.500
	df	595
	Sig.	.000

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings
	Total	% of Variance	Cumulative %	Total
1	12.423	35.493	35.493	12.423
2	3.749	10.711	46.204	3.749
3	2.103	6.007	52.212	2.103
4	1.859	5.312	57.524	1.859
5	1.532	4.377	61.901	1.532
6	1.262	3.607	65.508	1.262
7	1.167	3.335	68.843	1.167
8	.964	2.755	71.598	
9	.882	2.519	74.117	
10	.837	2.393	76.510	

11	.820	2.344	78.854
12	.739	2.112	80.965
13	.704	2.011	82.977
14	.591	1.690	84.667
15	.529	1.513	86.179
16	.506	1.445	87.624
17	.471	1.345	88.969
18	.438	1.252	90.221
19	.382	1.092	91.312
20	.340	.972	92.284
21	.329	.941	93.225
22	.305	.872	94.097
23	.270	.772	94.869
24	.252	.720	95.589
25	.225	.644	96.233
26	.197	.563	96.795
27	.192	.549	97.344
28	.170	.486	97.830
29	.142	.405	98.235
30	.137	.390	98.625
31	.123	.351	98.977
32	.113	.324	99.300
33	.103	.295	99.595

34	.077	.220	99.816	
35	.064	.184	100.000	

Extraction Method: Principal Component Analysis.

Total Variance Explained

Component	Extraction Sums of Squared Loadings	
	% of Variance	Cumulative %
1	35.493	35.493
2	10.711	46.204
3	6.007	52.212
4	5.312	57.524
5	4.377	61.901
6	3.607	65.508
7	3.335	68.843
8		
9		
10		
11		
12		
13		
14		
15		

16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32		
33		
34		
35		

Extraction Method: Principal Component Analysis.

APPENDIX B – SPSS OUTPUT – FACTOR ANALYSIS FOR EYE-HAND COORDINATION

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings
	Total	% of Variance	Cumulative %	Total
1	7.894	31.575	31.575	7.894
2	2.728	10.913	42.488	2.728
3	1.833	7.330	49.818	1.833
4	1.526	6.106	55.924	1.526
5	1.112	4.448	60.372	1.112
6	1.094	4.375	64.748	1.094
7	.986	3.943	68.690	
8	.894	3.575	72.265	
9	.772	3.087	75.352	
10	.734	2.937	78.289	
11	.693	2.772	81.061	
12	.683	2.733	83.793	
13	.579	2.315	86.108	
14	.535	2.142	88.250	
15	.426	1.705	89.955	
16	.418	1.672	91.627	

17	.375	1.499	93.126	
18	.299	1.198	94.324	
19	.274	1.096	95.420	
20	.265	1.061	96.481	
21	.218	.872	97.353	
22	.196	.783	98.136	
23	.169	.677	98.813	
24	.159	.634	99.447	
25	.138	.553	100.000	

Extraction Method: Principal Component Analysis.

Total Variance Explained

Component	Extraction Sums of Squared Loadings	
	% of Variance	Cumulative %
1	31.575	31.575
2	10.913	42.488
3	7.330	49.818
4	6.106	55.924
5	4.448	60.372
6	4.375	64.748
7		
8		

9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		

APPENDIX C – SPSS OUTPUT – FACTOR ANALYSIS FOR EXECUTIVE FUNCTION

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.565
Bartlett's Test of Sphericity	Approx. Chi-Square	62.065
	df	3
	Sig.	.000

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings
	Total	% of Variance	Cumulative %	Total
1	1.755	58.483	58.483	1.755
2	.851	28.377	86.860	
3	.394	13.140	100.000	

Extraction Method: Principal Component Analysis.

Total Variance Explained

Component	Extraction Sums of Squared Loadings	
	% of Variance	Cumulative %

1	58.483	58.483
2		
3		

APPENDIX D – Kilifi Developmental Inventory

Kilifi Developmental Inventory
Psychomotor Scale /Working Memory and Inhibition
November 08 – South Africa NELDS

STUDY NUMBER [_][_ _][_ _][__][__][__]

CHILD'S NAME: [_____ _____] [__][__]

TIME ARRIVED AT HOME: [_____]

ASSESSMENT DATE: [__][__]/[__][__]/[__][__]

DOB [__][__]/[__][__]/[__][__] **AGE AT ASSESSMENT:** [____] **Months** [____] **Days**

ASSESSOR INITIALS [____][____]

Test session observations

Health
Any unusual events during testing

Codings:

Code each item yes/no (Y/N) (coding box shaded)

If action not observed code 99 – failure to administer

98 – refusal by child

Where specifically indicated some items require a time or quantity (coding box not shaded).

If no score available code 99 – failure to administer

98 – refusal by child

97 – no score as item designated as inappropriate (too difficult) by assessor

Instruction: Brief the person accompanying the child on the tasks and the role they are expected to play during the assessment.

Page numbers refer to KDI manual instructions **START TIME [____:____]**

MOVING	**(page 10)**	**Yes/No**
LM 01	Stands with support	
LM 02	Stands without support	

PLAYING WITH THE BALL	**(page 11)**	
Code/score		**Yes/No**
EH 01	**THROWS** and **CATCHES** ball – at least 3 throws and catches anyhow.	
EH 02	**Releases a ball** purposefully	
EH 03	Throws a ball towards someone	
EH 04	**Catches** a ball using arms and hands	
EH 05	Catches a ball using hands only	
LM 03	**Can kick** a ball from stationary position	
LM 04	Can kick a moving ball	

MOVING	**(page 10 and pages 12- 19)**	
Code/score		**Yes/No**
LM 05	Walks **when held with one hand**	
LM 06	Walks without help	

LM 07	**Jumps** with two feet together, holding on to person's hand (p.12)	
LM 08	Jumps with two feet together unsupported both feet leave the ground	
LM 09	**Stands on one leg**, without support for 10 – 20 seconds (p.13)	
LM 10	Stands on one leg, without support for 21 seconds+	
LM 11	**Walk on tip toes** for 3 or more steps (p.14)	
LM 12	Walk on tip toes for length of mat	
LM 13	**Walks backwards** with support or for less than length of mat (p.15)	
LM 14	Walks backwards alone for length of mat	
LM 15	**Walks** along line **heel to toe**, arms out for 3 steps (p.16)	
LM 16	Walks along line heel to toe, arms out for length of mat	
LM 17	Walks on **side of feet** (p.17)	
LM 18	Can **pull** self onto platform (p.18)	
LM 19	Can **climb** straight onto platform	
LM 20	**Jumps off** platform with support	
LM 21	Jumps off platform and lands on both feet	
LM 22	**Hops on one** leg on spot (p.19)	
LM 23	Hops on one leg for length of mat	

LYING AND STANDING	**(pages 20-24)**	
Code/score		**Yes/No**
LM 24	Lifts upper body while lying on stomach (p.20)	
LM 25	Sits with support (p.21)	
LM 26	Can sit steadily/ without support	
LM 27	No head lag in sitting position	
LM 28	Can reach out and return to sitting position	
LM 29	**Rolls** from side to back (p.22)	
LM 30	Rolls from side to side	
LM 31	**Moves from lying** to sitting pushing up with hands (p.23)	
LM 32	Moves from lying to sitting not using hands	

LM 33	**Moves from sitting** to standing rolling over and up	
LM 34	Moves from sitting to standing not using hands	
LM 35	Can sit down steadily (from a standing position- without hands) (p.24)	

VISION: RING AND RED TASSEL (page 26)		
Code/score		**Yes/No**
EH 06	Reaches for dangling ring	
EH 07	Takes dangling ring	
EH 08	Follows red tassel with eyes/attempts to grasp	
EH 09	Grasps red tassel successfully	
EH 10	Can hold and examine object (ring, bear etc)	
EH 11	Passes object from hand to hand	

BUTTON (page 27)		
Code/score		**Yes/No**
EH 12	Can do up button with adult holding one piece of cloth	
EH 13	Can do up button alone	

BLOCK TOWER (page 28)			
A. Can retain cube in either hand when given	**B**. Retains one cube When second offered	**C**. Picks cube up from mat	**D**. Mature (radial) grasp
E. Can hold 2 cubes in one hand	**F**. Retains 2 cubes when third offered	**G**. Releases one cube on top of another	**H**. Builds tower 3-4 cubes
I. Builds tower 5-6 cubes	**J**. Builds tower 7-8 cubes	**K**. Builds tower 9-10 cubes	**L**. Builds tower 11-12 cubes
EH 14	Number of boxes ticked		

CONTAINERS AND CUBES (page 29)				
A. Rattles box	**B**. Lifts lid of box (not knocking off/over)	**C**. Tries to take cube out of box	**D**. Manages to take 1 cube out of box	**E**. Removes both cubes from box
F. Opens 2 boxes	**G**. Puts 1 cube in box	**H**. Puts 2 cubes in box	**I**. Puts cubes in and out	**J**. Puts lids back, trial

	(encourage)	(encourage)	of box	and error
K. Puts 2 cubes and lid back	**L**. Puts lid back, adjusts lid to box	**M**. Puts 3 boxes together	**N**. Assembles boxes by colour	
EH 15	Score (count number of boxes ticked)			

COIN BOX	**(page 30-31)**				
A. Picks up coin any method R L	**B**. Picks up coin between thumb and finger R L	**C**. Can put coin in the box (slot horizontal) R L	**D**. Can put coin in rotated box (slot vertical) shakily R L	**E**. Can put coin in rotated box: easily R L	**F**. Puts in 6 coins in rotated box R L
Code/score					**Yes/No**
EH 16	Score (count number of ticks)				

EH 17 R 6 Rotated box	**a)** Trial 1(secs) (record time)	
	b) Trial 2(secs)	
	c) Trial 3(secs)	
EH 18 L 6 Rotated box	**a)** Trial 1(secs)	
	b) Trial 2(secs)	
	c) Trial 3(secs)	

BEAD THREADING	**(page 32)**	
Code/score		**Yes/No**
EH 19	Picks up beads with pincer grasp	
EH 20	Drops beads into container	
EH 21	Threads 2 beads onto shoe lace	
EH 22 How many in 30 secs.	**a)** Trial 1 (record number)	
	b) Trial 2 (record number)	
	c) Trial 3 (record number)	

PAPER AND PEN (page 33)		
Score		**Yes/NO**
EH 23	Holds a pen in any way	
EH 24	Holds a pen between finger and thumb	
EH 25	Can scribble using a pen	
EH 26	Can imitate a straight line	
EH 27	Can imitate a circle	

TIME [____:____]

A-Not- B And Self Control Scoring Sheet (pages 35-37) **Start time: [____:____]**

Self Control:

SC1	Time in secs.	

A not B:

1.Delay time to be observed -*Tick delay time you will be using during this session.*

3 secs **less than 12 months**
5 secs **12 to 24 months**
10 secs **25 to 36 months**

2. Left and right refer to the assessor's right or left.
3. Alternate the side you start with i.e. if the first child you started assessing from left, then second child starts from right.
4. Always place the mark where the child searched ✓ for getting the treat and × for missing the treat.

Child seating

Trial	Left Cup	Right Cup
1		
2		
3		
4		
5		

6		
7		
8		
9		
10		

Assessor seating

SCORING:

Trials**:** **[**_____**]**

Total Number correct: **[**_____**]**

Total number of errors (after first set) **[**_____**]**

Self Control 2:

SC2	Time in secs.	

Test behavior: Comment on child's behavior and any unusual events during this task.

End time