Ephraim S. Kalanje

Identification of First Graders at Risk of Reading and Writing Difficulties

Creating a Group-Based Screening Tool in Kiswahili in Tanzania

Ephraim S. Kalanje
Teacher Educator, Morogoro Teachers' College, Tanzania, from 2000

Academic experience
Secondary School Education - Pugu Secondary School, 1970
Teachers' Certificate, 1971 and Fine Art Certificate, 1981- Butimba Teachers' College, Mwanza
Physical Education Teaching Certificate - Dar-es-Salaam Teachers' College, 1988
Special Education Certificate - Summer School, University of Oslo, Norway, 1989
Primary School Management and Administration - MANTEP/ADEM, 1992
Bachelor of Education - Åbo Akademi University, Vasa, Finland, 2000
Master of Education - Åbo Akademi University, Vasa, Finland, 2002

Worked as Primary School Teacher, Head Teacher and Teacher Educator:
Mbozi and Mbeya, 1973-1979
Dar-es-Salaam Teachers' College, 1981-1989
Solomon Mahlangu Freedom College, SOMAFCO, Morogoro, 1990-1991
Kigurunyembe Practicing School, Morogoro, 1992-2000

Cover: Tove Ahlbäck

Åbo Akademi University Press
Biskopsgatan 13, FI-20500 ÅBO, Finland
Tel. int. +358-20-786-1468
E-mail: forlaget@abo.fi
http://www.abo.fi/stiftelsen/forlag/

Distribution: Oy Tibo-Trading Ab
P.O.Box 33, FI-21601 PARGAS, Finland
Tel. int. +358-2-454 9200
Fax int. +358-2-454 9220
E-mail: tibo@tibo.net
http://www.tibo.net

IDENTIFYING FIRST GRADERS AT RISK OF READING AND WRITING DIFFICULTIES

Identifying First Graders at Risk of Reading and Writing Difficulties

Creating a Group-Based Screening Tool in Kiswahili in Tanzania

Ephraim S. Kalanje

ÅBO 2011

ÅBO AKADEMIS FÖRLAG – ÅBO AKADEMI UNIVERSITY PRESS

CIP Cataloguing in Publication

Kalanje, Ephraim S.
Identifying first graders at risk of reading and writing difficulties : creating a group-based screening tool in Kiswahili in Tanzania / Ephraim S. Kalanje. – Åbo : Åbo Akademi University Press, 2011.
Diss.: Åbo Akademi University.
ISBN 978-951-765-590-3

ISBN 978-951-765-590-3
ISBN 978-951-765-591-0
(digital)
Oy Arkmedia Ab
Vasa 2011

Abstract

IDENTIFYING FIRST GRADERS AT RISK OF READING AND WRITING DIFFICULTIES

CREATING A GROUP-BASED SCREENING TOOL
IN KISWAHILI IN TANZANIA
ABSTRACT

EPHRAIM S. KALANJE

Early identification of beginning readers at risk of developing reading and writing difficulties plays an important role in the prevention and provision of appropriate intervention. In Tanzania, as in other countries, there are children in schools who are at risk of developing reading and writing difficulties. Many of these children complete school without being identified and without proper and relevant support. The main language in Tanzania is Kiswahili, a transparent language. Contextually relevant, reliable and valid instruments of identification are needed in Tanzanian schools. This study aimed at the construction and validation of a group-based screening instrument in the Kiswahili language for identifying beginning readers at risk of reading and writing difficulties. In studying the function of the test there was special interest in analyzing the explanatory power of certain contextual factors related to the home and school.

Halfway through grade one, 337 children from four purposively selected primary schools in Morogoro municipality were screened with a *group test* consisting of 7 subscales measuring phonological awareness, word and letter knowledge and spelling. A questionnaire about background factors and the home and school environments related to literacy was also used. The schools were chosen based on performance status (i.e. high, good, average and low performing schools) in order to include variation. For validation, 64 children were chosen from the original sample to take an *individual test* measuring nonsense word reading, word reading, actual text reading, one-minute reading and writing. *School marks* from grade one and a *follow-up test* half way through grade two were also used for validation. The correlations between the results from the group test and the three measures used for validation were very high (.83-.95).

Content validity of the group test was established by using items drawn from authorized text books for reading in grade one. Construct validity was analyzed through item analysis and principal component analysis. The difficulty level of most items in both the group test and the follow-up test was good. The items also discriminated well. Principal component analysis revealed one powerful latent dimension (*initial literacy factor*), accounting for 93% of the variance. This implies that it could be possible to use any set of the subtests of the group test for screening and prediction. The K-Means cluster analysis revealed four clusters: *at-risk children, strugglers, readers* and *good readers*. The main

concern in this study was with the groups of *at-risk children* (24%) and *strugglers* (22%), who need the most assistance. The predictive validity of the group test was analyzed by correlating the measures from the two school years and by cross tabulating grade one and grade two clusters. All the correlations were positive and very high, and 94% of the at-risk children in grade two were already identified in the group test in grade one.

The explanatory power of some of the home and school factors was very strong. The *number of books at home* accounted for 38% of the variance in reading and writing ability measured by the group test. *Parents´ reading ability* and the *support children received at home for schoolwork* were also influential factors. Among the studied school factors *school attendance* had the strongest explanatory power, accounting for 21% of the variance in reading and writing ability. Having been in *nursery school* was also of importance.

Based on the findings in the study a short version of the group test was created. It is suggested for use in the screening processes in grade one aiming at identifying children at risk of reading and writing difficulties in the Tanzanian context. Suggestions for further research as well as for actions for improving the literacy skills of Tanzanian children are presented.

Key words: Beginning readers; literacy skills; group-based screening tool; validation; longitudinal study; children at risk; reading and writing difficulties, home- and school factors; Kiswahili; Tanzania.

Acknowledgements

Various people were involved at different stages of conducting this study. Here, I would like to remember some of them. I thank Prof. Sven-Eric Hansen, not only for being the mastermind and overseer of the TEPT and DEPT projects, but also for believing in me as I ventured for the bachelors, masters and later the doctoral program within the projects. He provided me with the opportunity and the experiences that have enabled me to be a researcher and to actualize this study. I am particularly greatly indebted to Professor Ulla Lahtinen for her able and tireless supervision despite the frequent frustrations she suffered as I sometimes lost focus or lacked motivation. She was always there, in a motherly way, to make me stand up and walk again. My special thanks go to the reviewers of my thesis, Prof. Jan-Eric Gustafsson and Prof. Mikko Aro for constructive comments and suggestions in the final stages. I am also indebted to Prof. Åke Olofsson, and Prof. Karin Taube for their expert advice in the process of creating the screening instrument.

I am also grateful to Prof. Anna-Lena Östern, my mentor, from Åbo Akademi University, who, together with my supervisor, formed a team of experts who were critical, but also offered me invaluable advice on how to conduct this study. Doctor Tom Wikman, alias 'kaka', deserves my compliments for reading and offering his suggestions on some parts of the thesis and for the light side of living in Vasa. During the doctoral program I have enjoyed support from several persons connected to the project. My sincere thanks go to all of them. Therese Ahlström is especially remembered for her support regarding practical matters during the program, Kristina Ström for valuable comments during the seminars, the 'village girl' Tarja Grahn-Björkqvist for technical support and layout of this study, and John Shepherd for tireless proof reading and corrections of my English text.

At this stage I salute my colleagues, all the doctoral students, with whom we toiled together but also at times enjoyed being together as we set out to climb the challenging educational ladder. The discussions we shared in Finland and Tanzania made a great contribution to what this study is now.

I would also like to thank Prof. Mwajabu Possi, then from the Tanzania School of Journalism and the late Prof. Zacharia Mochiwa from University of Dar-es-Salaam who encouraged me to go on with the study after reading the proposal in its infancy stages. Mr. Theobald Mvungi from the Tanzania Institute of Education deserves the same compliments. Here, I also mention Mr. Salmon Mwanyika and Mr. Bartholomew Makoa from Morogoro Teachers' College. All of them provided me with useful information concerning the Kiswahili language. I am indebted to Janet Townsend from the UK, then an expatriate tutor at Morogoro Teachers College, who critically read the thesis in its advanced stages and provided me with invaluable suggestions for improvement.

I should not forget the DEO and Inspectorate offices in Morogoro Municipality for useful information and granting me permission to conduct the study in the sampled schools. I also remember the heads of primary schools, teachers and parents and the children in SUA, Kigurunyembe, Kilakala and Mwande in Morogoro Municipality for arranging a conducive environment for conducting the study in their schools.

I thank Pastor Mikko, Hannele, Liisa and all members of the Vasaan Keidas Seurakunta for spiritual and material support during my study periods in Vasa. I am also grateful to my special friend Kaisa Anttonen, who has supported me in many ways during my Vasa visits. Antti Lahtinen deserves my heartfelt thanks for introducing me to Finnish cultural activities like sauna, ice-fishing and horse-racing. Antti, a friend in need, a friend indeed!

I thank the Ministry of Education in Tanzania and Mr. Wolfram Ngonyani, the principal of Morogoro Teacher's College for support and permission to pursue the program.

For financial support, I extend my sincere thanks to the University Foundation of Ostrobothnia (Högskolestiftelsen i Österbotten), to the Research Institute of Åbo Akademi Foundation (Stiftelsens för Åbo Akademi Forskningsinstitut) and to the Rector of Åbo Akademi University.

Last, but not least, I dedicate this work to my late wife Marie, who was always alongside my struggle as I set out to pursue the educational trail. Although she never got to see the completion of my endeavours, she left her encouraging and understanding spirit to our loving children Ephel, Pamela, Colleta, Nancy and Sidney, who were always there for me, and to our grandchildren Aidan, Paul, Ikunze and Bukeyeneza.

Vasa, 15th of April, 2011

Ephraim S. Kalanje

List of figures

List of tables

List of abbreviations

BEST	Basic Educational Statistics
CV	Consonant-Vowel
DAST	Dyslexia Adult Screening Test
GP	Good Performance
HP	High Performance
IQ	Intelligence Quotient
LDs	Learning Disability (ies)
LP	Low Performance
MLYS	Ministry of Labor, Youth and Sports
MOEC	Ministry of Education and Culture
MoEVT	Ministry of Education and Vocational Training
NCLB	No Child Left Behind
NRFI	National Reading First Initiative
PA	Periphery Average
PCA	Principal Component Analysis
PEDP	Primary Education Development Plan
PSLE	Primary School Leaving Examination
RAI	Rapid Assessment Instruments
RTI	response to intervention
SD	Standard Deviation
SEDP	Secondary Education Development Plan
SLDs	Specific Learning Disability (ies)
SPSS	Statistical Package for the Social Sciences
SUA	Sokoine University of Agriculture
SVR	Simple View of Reading
UNICEF	United Nations Children Educational Fund
USA	United States of America
USDE	United States Department of Education

A whole 80 percent of students in a form three class cannot read and write? You will likely be inclined to say that is impossible, yet Arusha's Njiro Secondary School will easily prove you wrong. And it will be none other than the headmistress affirming that 170 out of a class of 224 failed miserably last year's national form two examinations, chiefly because they cannot even write their names!

(Editor, The Guardian. Wednesday, May 13, 2009).

1 Introduction

1.1 Background to the problem

How do we best assist children with reading difficulties to benefit in schools, and later in life? Obviously, there is not a clear cut answer to this issue. Various strategies are employed in schools today to try to address this problem. Research in the field of learning difficulties and reading and writing difficulties, specifically, has provided us with knowledge and strategies about identification and intervention programs for children at risk of these difficulties. However, most of the research is done in developed countries, in contexts and languages not directly comparable with contexts in developing countries. This study addresses the topic in Tanzania, a developing country, where Kiswahili is the main language. The focus is on early identification of children at risk, for the purpose of providing proper intervention before failure sets in. In this chapter, the importance of research on the topic in general, and especially in the context of Tanzania is shortly discussed.

Early identification of children with reading and writing difficulties

Identification of beginning readers at risk of reading and writing difficulties is an important issue. Many researchers (e.g., Pumphrey & Reason 1991; Taggart, Sammons, Smees, Sylva, Melhuish, Siraj-Blatchford, Elliot& Lunt, 2006) claim that identification can be justified as essential for the estimation of incidence, and as necessary for provision of appropriate intervention. In order for children to benefit maximally from the intervention, it is stressed that the identification should be done in the early years (e.g., Bailet, Repper, Piasta & Murphy, 2009; Cavanaugh, Kim, Wanzek & Vaughn, 2004; Torgesen, Wagner, Rashotte, Rose, Lindamood & Garvin, 1999; Torgesen, 2002; Vellutino, Scanlon, Small & Fanuele, 2006). Many studies have also shown that early intervention can significantly improve the reading outcomes of children at risk of reading disabilities (Bailet et al., 2009; Denton & Mathes, 2003; O' Connor, Fulmer, Harty & Bell, 2005; Simmons, Coyne, Kwok, MacDonagh, Harn & Kame'enui, 2008).

Early identification and intervention should take place before a sense of failure sets in (e.g., Lesaux & Siegel, 2003; Maugham & Yule, 1994; National Research Council, 1998). On this issue, Ziolkowska (2007) has raised concerns and questioned, *"Why wait for a child to fail if failure can be anticipated?"* (See also MacMaster, Du & Pétursdóttir, 2009). The current model of early identification focuses on early preventive measures, and is more successful compared to the traditional model, which focuses on waiting for the child to fail before remedial programs are initiated (Berninger, Stage, Smith & Hildebrand, 2001; Berninger, Abbot, Vermeulen, Ogier, Brooksher & Zook, 2002; Coyne, Kame'enui, Simmons & Harn, 2004).

Response to intervention (RTI) has been seen as a successful model for the early identification of children at risk of reading difficulties. Within the typical RTI model, all children receive periodic screening (i.e. universal screening) for risk of reading difficulties. Pupils identified as at risk on the basis of this screening are provided with short term intervention. Those who fail to respond to intervention are considered to be truly at risk of reading difficulties, and are provided with more specialized intervention (Bailet, et al., 2009; Gunning, 2009; Fletcher, Coulter, Reschly & Vaughn, 2004; Fuchs, Fuchs & Speece, 2002; Vaughn, Fletcher, Francis, Denton, Wanzek, Wexler, Cirino, Barth & Romain, 2008).

Reliable and valid tools for identification

Reliable and valid tools are essential in the identification of children with learning difficulties (Rothenberger, 2005). Many tools are available in different countries, measuring specifically children's reading and writing ability and performance. From the late 1960's, Beery and his colleagues had already noticed that the number of such tools surpassed those that measure other areas of the curriculum (Beery, Barrett & Powell, 1969). However, it has been found that many of the tools lack the support of a theoretical framework with current theories of reading and writing (Haats, 2002; Hurry & Doctor, 2007; Wenner, 1995). In research studies, it has been noticed that many tests indicate a reading or spelling age, but do not inform specifically about the individual child's literacy difficulties. This is in contradiction with current initiatives, for instance, taken in countries such as the United States, which require teachers to use evidence-based practices within the policy of No Child Left Behind (NCLB), (Graham & Bailey, 2007; Learning Print Associates, LPA, 2007; Odom, Brantlinger, Gersten, Horner, Thompson & Harris, 2005). In order to successfully prevent reading difficulties, it is emphasized that at risk students first have to be identified accurately (see Phillips, Lonigan & Wyatt, 2009).

The prevalence of children with reading and writing difficulties

A reliable identification instrument is a prerequisite for obtaining reliable prevalence figures. These again are of high importance for estimating the need for countrywide measures to be taken, e.g., related to teacher education, curriculum development and intervention strategies. The prevalence of children with reading and writing difficulties varies according to many factors, e.g., definition, context and language. It can be explained by referring to the broader learning disability (LD) group which encompasses several specific disabilities including reading and writing difficulties. The LD group, which by definition excludes the physical, sensory and intellectual disabilities, is the largest of all disability groups in most schools the world over (Chapman, 1992). In the United States, 75 to 80 percent of students with LD have reading or language deficits

(Rathvon, 2006; Podhajski, Mather, Nathan & Sammons, 2009). This is also observed in Malaysia, a developing country, where it is estimated that a similar approximation figure of 80 percent of children with LD have their primary difficulties in learning to read. The prevalence of children with LD in Malaysia was reported to be 9.1 percent in 2003 and 7.7 percent in 2007 (Lee, 2008).

The number of children facing difficulties in reading and spelling seems to vary as a function of the nature of the writing system, (i.e. orthography). Lyytinen and Erskine (2006) found that in highly transparent orthography, such as Finnish; roughly 6% of children have difficulties with acquisition, while 3%, mostly with familial background, have severe difficulties. In opaque orthography, such as English, the number of delayed early learners is relatively larger, with more than 10 % of young readers of English facing problems in achieving sufficient accuracy and fluency of reading and spelling (Lyytinen & Erskine, 2006).

The Tanzanian situation

Tanzania, like some other developing countries, e.g., Malaysia and Zambia, lacks theory-based literacy assessment tools (Lee, 2008; Ketonen & Mulenga, 2003). However, the schools have children in difficulty of reading and writing and in need of identification and assistance. Tanzania, in particular, lacks such tools, probably as a result of the general lack of research and service in the area of specific learning difficulties (SLDs). According to Mboya and Possi (1996), as well as Kahigi (2003), the SLD field has not been catered for in Tanzania compared to other disabilities such as the fields of physical, sensory and intellectual impairment. Some authors have noticed that the prevention of reading and writing difficulties in Tanzania has not been taken care of (Kyando, 2007; Malekela, 2000).

In fact, concern for the provision of quality education even to disadvantaged groups, including disabled children, is very well stipulated in Tanzanian education and training policy (Ministry of Education and Culture, MOEC, 1995). The policy has gone even further to include the need for screening talented students. Currently, there are *special schools for talented children*[1] in Tanzania. However, although it is mentioned in the policy about the need for early identification of children with learning difficulty learning (Ministry of Labor, Youth Development and Sports, MLYDS, 2004), this has yet to be realized practically. There are also no specific programs for children in the mainstream who are facing various kinds of specific learning difficulties, many of them not disabled and having average intelligence.

Due to lack of research in this field, we cannot be sure of the prevalence rate in Tanzania. Although figures may not exist, however, we can speculate it to be high, following the presence of many problems that hinder attainment of quality education (MOEC, 1995; Mosha, 2000; Munishi, 2000). In a recent extensive

[1] These are public secondary schools specified for children who excel in the Primary School Leaving Examination (PSLE).

[2] In Tanzania *grades* are referred to as *standards, i.e. standard one, standard two, etc.*

assessment survey, involving 38 out of 133 districts in the country, a total of 42,033 children aged 5-16, in 22,800 households, were assessed in basic literacy and numeracy skills (Uwezo, 2010). It was found that one in five primary school leavers cannot read standard 2 level Kiswahili. It was also found that half of the children who complete primary school cannot read in English. It can be reasoned that as result of their great numbers and lack of identification strategies, children with learning disabilities, and specifically those with reading and writing difficulties, have probably not been well aided in Tanzania.

However, it is equally true that the lack of reliable and validated screening instruments have left most of these children unidentified and unaided.

Tanzania needs to have its own contextually developed literacy assessment tools underpinned by current theories of reading and writing. The country needs to develop both screening tests for early identification of children at risk of developing reading and writing difficulties, as well as theoretically based diagnostic tests for planning proper and relevant intervention. Because of contextual and language and language differences, Tanzania cannot successfully use internationally available tests, whether be they non-theoretical or theoretical. Reid and Wearmouth (2009) contend that it is important to contextualize the assessment process and results from any tests used in relation to the curriculum and the nature of the child's learning situation. The authors have noticed that factors within the classroom and materials that are being used may account for the difficulties the child is displaying as much as the child's own attributes.

Currently, we rarely find such contextualized tools in Tanzania. To my knowledge, after thirty years of teaching experience and discussion with teachers and experts in the field, the identification practice seems to rely only on the experience of teachers and teacher-made tests. I noticed, in most cases, that the tests measure curriculum coverage and lack cognitive linguistic indicators. The assessment done also considers very little about the school and home background factors that surround the child when trying to acquire the skills of reading and writing.

Existence of children with reading and writing difficulties in Tanzania

Are there children with persistent reading difficulties in Tanzania needing organized identification and intervention or requiring specialized assistance? In fact, there are no reasons why such children would not exist among the Kiswahili speaking children in Tanzania. Høien and Lundberg (2000) contend that dyslexia, (i.e. persisting difficulty in reading and writing) is not confined to any one group of language, and that dyslectics are likely to be found in any society with an alphabetic language. Kiswahili is an alphabetic language (see Alcock, Nokes, Ngowi, Musabi, Mbise, Mandali, Bundy & Baddeley, 2000; Kahigi, 2003) and thus the existence of dyslectics in Tanzania is obvious, although no studies about their prevalence have been carried out so far. However, we do have evidence that there are children in Tanzania who complete

primary education without the skills of reading and writing. Experienced teachers, including myself, have encountered children with these problems. Also Kahigi (2003) has indicated the existence of children with reading difficulties in Tanzania. He has shown concern for the children and the desire for them to be assisted. Further, he has suggested worthwhile methods for teachers to use in order to identify these children and help them learn to read and write.

Moreover, reports from available research and statistics indicate that there are problems of reading and writing as well as dropouts in nearly all schools in Tanzania (Ministry of Education and Vocational Training, MoEVT, 2008). According to Kutnick, Jules and Layne (1997) and OXFAM (2007), the problems are frequent, but not well understood, despite the fact that since 1967 Tanzania has been launching various educational projects aiming at improving the educational delivery system (Kapinga, 2004).

Children repeating classes or dropping out of school in Tanzania could also be indicative of some children facing learning difficulties, although not the sole reason (Basic Education Statistics, BEST, 2008; MOEC, 2003; MoEVT, 2008). However, there is evidence that some children in Tanzania complete primary education without the skills of reading and writing (Malekela, 2003).

Reading and writing ability in relation to quality of education in Tanzania

How is the failure to read and write in some children related to low quality of education in Tanzania? It should be understood that there has been general concern about improving the quality of education. According to Kamwela (2000) and Otieno, (2000), the quality of education has been declining since the late 1980s. Many reasons and indicators of low quality education have been suggested in elaborated categories of inputs, outputs and outcomes. Inadequate financial support, lack of well-trained teachers, lack of teaching and learning materials, ineffective methods of teaching, poor teaching and learning environment have been presented as tangible reasons for poor quality education. Intangible reasons given include low ability of students in decision-making, critical thinking, reasoning and problem-solving (Mosha, 2000; Munishi, 2000).

In an effort to address the situation, Tanzania has initiated educational development strategies through the Primary Education Development Plan, PEDP/2002-2006, (Basic Education Development Committee, BEDC, 2001; Davidson, 2004; MOEC, 2001; Sumra, & Rajani, 2006) and Secondary Education Development Plan SEDP/2004-2009, (Tonini, 2009; Sumra & Rajani, 2006). Through the initiative, school curricula and teacher training college curricula have been reviewed. Further still, more classrooms, teachers' houses, libraries, laboratories, and hostels have been built. More children are enrolled to school and more primary school qualifiers are now getting the chance to continue with secondary education. Also, not only are more teachers recruited,

but efforts have also been made to ensure that teachers in the field are engaged in capacity building programs.

While all the above are commendable measures that Tanzania is taking in raising the quality of education, more has yet to be done. Further steps should lead to the scrutiny of the actual teaching and learning process in classroom situations (see Uwezo, 2010). For example, there is not enough evidence-based knowledge about how children learn to read and write and the difficulties they encounter. Apart from the given indicators of low quality education in Tanzania, and the efforts to address them, the general public, and some authors (e.g., Chonjo, 1966; Malekela, 2003), have argued that children's inability to read and write on completing primary education also contributes to poor quality education. Teachers, in this case, need to be oriented towards evidence-based practices, at classroom level, that can lead to the identification and development of proper interventions with those who lag behind in reading and writing and learning in general.

Dropout and repetition rates in Tanzanian schools

Low school performance among pupils, in general, and the failure of some pupils to read and write at the end of their seven years of primary education have been attributed to ineffective teaching and truancy on the part of pupils. Teachers have been considered to be inadequately prepared in teacher training colleges. According to Malekela (2003), the low quality of teacher education in Tanzania has a negative impact on the performance of students in schools. Concerning truancy, Malekela and Ndeki (2001) found that about 30 percent of the children who start standard one do not complete the seven year cycle of primary education. Furthermore, MoEVT (2004-2008) found that only 40 percent of children complete the primary cycle. Due to low school achievement these students are subjected to dropout. However, the statistics given for dropout are not explicit about low achievement as the reason. The dropout rate (DR) by reason, for four years (during 2000-2007) is given in Table 1:

Table 1: Dropout rate by reason in (%)

Year	*Truancy*	*Pregnancy*	*Death*	*Other*
2000	86	3.9	3.7	6.7
2003	79.7	5.4	6.7	8.2
2005	77.7	6.0	5.6	11.1
2007	66	5.5	5.6	22.3

The table indicates that truancy is the most leading and persisting reason for dropout in Tanzania. The rates ranged from 66% to 86% between 2000 and 2007. The other reasons: pregnancy, death and other reasons such as illness, illness of parent or guardian, lack of school needs and other, have comparatively

lower rates (BEST, 2008; MOEC, 2003; MoEVT, 2008). A remedy for truancy has been urgently called for in Tanzania (BEST, 2008).

When the transition rates (grades 1 to 7 for the years 2003/2004 to 2007/2008) are studied, dropout rate by grade is found to be highest in grades 4-5 /7.5% and 5-6 /3.4% (MoEVT, 2008, p. 20). This is indicated in Table 2.

Table 2: Dropout rate by grade from 2003/04 to 2007/08 in (%)

	1-2	***2-3***	***3-4***	***4-5***	***5-6***	***6-7***
	DR	**DR**	**DR**	**DR**	**DR**	**DR**
2003/04	2.0	3.9	2.6	5.0	4.1	4.8
2004/05	2.4	3.0	2.4	6.6	3.3	2.8
2005/06	1.1	1.7	3.0	7.4	2.0	4.9
2006/07	3.3	1.5	1.8	7.2	1.4	3.7
2007/08	1.6	1.0	1.7	7.5	3.4	3.2

These statistics support Malekela & Ndeki (2001), who found that low achievement for some children causes a high dropout rate. Students who show low achievement at the end of the year (i.e. those who fail the grade 4 final examination) are required to repeat a class. However, some decide to drop out.

School attendance is also a factor that influences school achievement or performance among children. Children, mostly from low socioeconomic environments, may have low attendance for various reasons, which include lack of adequate food, sickness, engaging in petty business, running errands, attending to house chores, lack of school materials, lack of school uniform, fear of bullies, fear of punishment and walking long distances to school. (See Elbers, Lanjouw & Lanjouw, 2003; Mswia, Lewanga, Moshiro, Whiting, Wolfson, Hemed, Alberti, Kitange, Mtasiwa & Setel, 2003).

The above mentioned are part of the factors responsible for low performance in many subjects offered in schools, including reading and writing. Although the underlying reasons for truancy are not given, it cannot be ruled out that constant failure in some school subjects, including failure in reading and writing, has also the potential to frustrate students and therefore lead them to acts of truancy, consequently leading to poor attendance and dropout from school. It is worth noting here that the reverse could also be true. Truancy for any other reason could cause poor learning of school subjects, including reading and writing.

The high rate of children repeating classes between grades[2] one and four suggests failure to advance mostly as a result of poor performance in subjects. Here, also failure to read and write can be speculated to be one of the underlying

[2]In Tanzania *grades* are referred to as *standards, i.e. standard one, standard two, etc.*

factors of poor performance in other subjects. The repeating rate by grade is given in Table 3 for the years 2003/04 to 2007/08.

Table 3: Repetition rates by grade 2003/04 to 2007/08 in (%)

	1-2	***2-3***	***3-4***	***4-5***	***5-6***	***6-7***
	RR	**RR**	**RR**	**RR**	**RR**	**RR**
2003/04	9.6	5.0	4.8	11.3	0.7	0.7
2004/05	9.1	6.2	5.0	14.7	0.3	0.1
2005/06	10.6	6.2	5.1	10.9	0.2	0.0
2006/07	7.1	5.6	4.2	9.4	0.1	0.1
2007/08	9.7	5.5	5.4	8.3	0.4	0.3

The transition figures in the table for these five years indicate that pupils are repeating more in grade one: between 7.1% and 10.6%, and also in grade four: between 8.3% and 14.7%.

Failure to read and write among students, on a wider scope, however, cannot be attributed to ineffective teaching and truancy or the other cited factors only. Through my own teaching experience I have encountered children with quite good school class attendance, some coming from homes of comparatively high socioeconomic status and attending good schools with good, able and committed teachers. Many of them are average children, without a notable disability such as sensory, intellectual, behavioral disorder, speech and language problem or chronic disease. Yet still these children find it difficult to overcome their reading and writing difficulties. This is supported by Hamilton and Glascoe (2006), who contend that significant and persistent reading difficulties can occur despite adequate instruction, intelligence, and socioeconomic status. However, the authors point out the contribution of environmental factors (e.g., poverty, low parental education, less stimulating home environment and inadequate instruction) and organic factors (e.g., mental retardation, low IQ score, and hearing impairment) to reading difficulties.

In primary schools in Tanzania, specifically in the lower grades, the problem of reading and writing among quite average children is evident. Many children face difficulties in learning to read and write in the beginning years; however, most of them overcome the difficulties over time and master the skills, while others are left to lag behind with their difficulties persisting. Many children will also remain with difficulties in reading despite adequate intervention (Hamilton & Glascoe, 2006).

Ultimately, children with persisting reading and writing difficulties, according to my own experience, face a redundant, boring school life. Catts (2006) observed that the motivation towards learning of these children becomes low and they

develop a negative attitude towards their school life, with no good expectations. Hence, they become very limited in what they can do, which consequently leads them to academic failure. It is therefore not surprising that the children can resort to acts of truancy as discussed above. The motivation for and acceleration towards truancy is further enhanced by their frustrated parents and guardians (Chapman, 1992), who deem it worthless to continue spending money on individuals who are not profiting from or gaining anything at school (Malekela & Ndeki, 2001). Finally, the children drop out or move into sectors which are academically less demanding. This could also be part of the reason for youngsters later into life being out on the streets unoccupied and engaged in undesirable behavior, when they have never been able to learn to capitalize on their other strong points.

In well-to-do families, mostly in urban areas, the parents may use a large amount of money in private tuition, popularly practiced nowadays in Tanzania. Most of this tuition is drill-oriented, subjecting learners to rote learning, which is basically cramming (Sambo, 2001). The instruction in these classes is mostly not based in any diagnostic analysis. Those children, who manage to complete primary school, if they have not already dropped out, will only add to the population of adults in the country who lack the skills of reading and writing. In this way, the adult literacy rate is inflated.

Adult literacy rate in Tanzania

Tanzania is known to have been performing well in adult literacy as a result of literacy campaigns that were implemented between the 1960s and 1970s. By 1986 the adult literacy rate was 90%. However, there was a sharp drop by 1992. The rate was 84%, with an annual drop of 2%. The reasons for this situation include adult classes ceasing to operate, weak mobilization and management, lack of funding and unmotivated and incompetent instructors. Some 32% or more of the adult population are now thought to be illiterate. Late enrolment in school is also noticed as a reason (Kadege, Keregero, Mlekwa & Mushi, 1992). This is resulting in a fast decrease in adult literacy. The recent findings by Uwezo (2010) that many children are completing primary education without the skills of reading and writing may result in permanent low rates of adult literacy. According to Bhalalusesa (2003), there is a lack of accurate and reliable data on adult literacy in Tanzania. However, she reasons that the adult literacy rate may remain low as schools continue to throughput children who cannot read and write.

The importance of reading and writing ability

Indeed, children need to be identified and assisted early enough in acquiring the skills of reading and writing, since the skills are vital to their advancement in both their school and later life. According to Høien and Lundberg (2000), reading and writing is a task that every educational stakeholder, in their varied capacities, must be concerned about. Regular teachers in the lower grades,

especially, must be fully committed to see that every child entrusted to them can read and write before they advance into the upper grades. It is crucial for the children to be able to break the code, and obtain meaning from it, because much of the learning in school depends on information delivered through the written word. The ability to read at a proficient level enables children to be successful in other school subjects. According to Lyon (1997), reading skills have been found to affect a child's general knowledge, spelling skills, writing skills and vocabulary development.

The importance and special position of reading and writing among school subjects has been further discussed and acknowledged by Høien and Lundberg (2000). They state that instruction in reading and writing is one of the most important responsibilities. Moreover, they contend that it is a special subject to the extent that students who fail in learning to read will have a hard time in school, because so much of what they will learn in other subjects will depend upon their ability to read and write. The United States Department of Education, USDE, (2001, p. 3) categorically stated that, '*Reading is the key that unlocks virtually all other learning.'* Other authors (e.g., Hamilton & Glascoe, 2006; Holopainen, 2002; Zeffiro & Eden, 2000) see reading to be an important prerequisite for subsequent achievement, and failure to acquire competence affects other fundamental cognitive skills. It has also been argued that reading is essential for success in society. Furthermore, it has been found that the ability to read is highly valued and important for social and economic advancement (Casey & Howe, 2002; Haats, 2002; Snow, Burns& Griffin, 1998).

Kiswahili is the national language and also the medium of instruction in public primary schools. It is taught as a subject in public primary and secondary schools and in some private schools as well (MOEC, 1995). Hence, it is necessary for children to be enabled to read and write in Kiswahili and be able to communicate well through the language.

Summary

Screening with reliable and validated instruments ensures accurate identification of children in need of specialized support. This is effective when the process is done early rather than waiting for the child to fail. Tanzania needs tools of this nature because there are children, with average intelligence and without specific disabilities, who are struggling to acquire the basic skills of reading and writing. These children exist in Tanzania from the evidence that Tanzania uses Kiswahili language, which is an alphabetic language, although not necessarily dyslexic in its strictest sense. Further evidence is that some children in Tanzania complete the primary education cycle without the skills of reading and writing. The high dropout and repetition rates and poor school attendance, especially in the lower grades, may also suggest that some children are facing learning problems, including reading and writing difficulties. This situation need to be addressed because reading is important for the children's learning of other subjects at school, and for their later socioeconomic life. The child's failure and not being

helped frustrate both children and parents. The struggling children and their parents need to be relieved of this frustrating experience. The children, in particular, need to be saved from being future parents who cannot support their own children in acquiring the skills of reading and writing because they too are illiterate. This only extends the vicious circle.

1.2 Motives for the study

This study is, firstly, motivated by the need to identify beginning readers at risk of reading and writing difficulties. In Tanzania literacy education is mainly conducted for children between age 5 and 8 that is, between nursery school and standard two (MOEC, 1995). This is a critical period in terms of acquiring the skills of reading and writing. If children pass through these stages without acquiring the skills, it is very likely that they are headed towards facing a lot of learning problems in various subjects in the higher grades (Hamilton & Glascoe, 2006; Høien & Lundberg, 2001). Experience has shown that problems are solved best when they are well understood and detected earlier. The old adage, **'***The early bird catches the worm*,' applies. In realizing this, the National Reading First Initiative (NRFI) part of the 2001 NCLB Act, in the case of the USA, for example, not only emphasizes how to address the reading achievement gap between higher and lower performing students by reducing it, but also that NRFIs primary goal has been geared to ensure that every student can read at grade level or above not later than the end of grade three (Gersten & Santoro, 2006). Children with reading and writing difficulties need to be exposed earlier, for the sole reason of implementing an effective educational intervention as a support (Bailet et al., 2009).

Secondly, this study is motivated by the need for Tanzania to develop a culture of consistent early screening and identification of students with learning disabilities or other needs with reliable and validated instruments. This need is well stipulated in the National Policy on Disability (MLYDS, 2004, p. 4):

There is a need for early identification of children with disabilities. Early identification followed by appropriate intervention has a chance of eliminating the occurrence of a disability or minimizing its impact later in life. Unfortunately, Tanzania does not have a national programme for early intervention which would assist the identification of children with disabilities in their respective communities. The fact that parents hide their children with disabilities pushes them further away from accessing appropriate services.

In order to promote reading and writing among school children, therefore, Tanzania needs to build a culture of early identification by using reliable and valid instruments in order to arrange proper intervention. The fear is that if Tanzania does not do so, it might continue to have numbers of unknown children who complete primary school without the ability of reading and writing. Researchers have recommended that the screening process should be conducted

three times, a year, in the same grade. For example, Beech and Singleton (1997) specified that screening should be done three times a year: firstly, by a teacher-completed check-list or a test of pre-reading and early literacy skills, very soon after entry to school; secondly, by a follow-up test of reading later in the year, and thirdly, by an annual re-testing as part of monitoring progress. At present there is no consistency in the process in Tanzania, or it may not take place at all.

The teacher-made tests administered at different times in the school year, or even the external examination taken at the end of grade four, mainly assess curriculum coverage. They are not objective enough to serve the purpose of a screening tool. These tests lack psychometric properties for decision-making (see Leung, Lindsay & Lo, 2007). Furthermore, teacher-made tests need to be supplemented by evidence-based tools. On this, Montgomery (1990) elaborated and emphasized that the instrument to be developed should meaningfully supplement the teachers' experience and their tests such that together they could be useful to the diagnostic profile of the individual pupil.

Thirdly, this study is motivated by the need to create awareness among various educational stakeholders in Tanzania (parents, teachers and other educationists). It is important for them to be aware of the existence of a group of children who face real problems in reading and writing (or have other special needs) in the regular classroom and what problems they can face in schools and later in life. The stakeholders, particularly regular teachers, need to have ample knowledge about the group in terms of how it is defined, what causes the problems, and how to organize assessment and identification modalities. They also need to know the type of effective educational adaptations and intervention required for the group (Kalanje, 2002). Here, the old adage, *'A problem well-stated is a problem half-solved,'* applies. In so doing, the stakeholders would be well versed in policy-making, planning and decision-making for children identified with disabilities or other special needs.

Fourthly, more important still, this study is motivated by current initiatives for the establishment and development of inclusive schools, following international calls and the demand for Education for All (EFA) and School for All (SFA), (Savolainen, Matero & Kokkala, 2006). This includes local initiatives and programs on education e.g., PEDP/2002-2006 (BEDC, 2001; Davidson, 2004; MOEC, 2001; Sumra, & Rajani, 2006) and SEDP/2004-2009 (Tonini, 2009; Sumra & Rajani, 2006). According to Mmbaga (2002), the implementation of the inclusive schools project in Tanzania began in 1996. Various schools in some districts (e.g., Kibasila, Wailes and Temeke primary schools in Dar-es-Salaam Region, and others up country), have been set up as pilot projects schools. (See also Ainscow & Lewis, 2005).

In fact the urgency is even greater when it is realized that current demands are that all students with disabilities are provided with instruction aligned to age appropriate standards, regardless of the nature or severity of the disability. In other words, students must receive meaningful access to the general curriculum.

As such, this goes beyond earlier concepts of inclusion, which tend to focus on where students are taught (Gersten& Santoro, 2006). However, in no way can this venture be successful if students are not screened early and reliably enough in order to understand their needs. Before students in special schools or units are integrated into or included in the regular classroom, teachers need to understand the learning problems and needs of those in regular classrooms (see Cook, Cameron & Tankersley, 2007).

Parallel to the establishment and development of inclusion is the realization of the child-friendly school in Tanzania (Ainscow & Lewis, 2005; United Republic of Tanzania, URT, 2008, pp. 13-14; UNICEF, 2009, pp. 14-15). This set-up idealizes the inclusion of all children with their diversity and needs and enables all of them to learn effectively in a friendly atmosphere and to grow and develop to their latent capabilities. A child-friendly school ought to make sure that there are deliberate plans to enable all students to acquire the skills of reading and writing. A reliable and valid screening instrument is therefore necessary in a child friendly school set up.

Fifthly, the study is motivated by the need to help children with problems and frustrations in learning, and also the parents' and teachers' frustration resulting from their children failing at school. The frustration is worsened by ridicule, and unjustifiable labeling of the children by their peers and other people in the community. Frustrating terms and names such as *minimally brain injured, slow learners, dyslexic or perceptually disabled are given to the children* (Hallahan & Kauffman, 1994). In Tanzania, the children may be referred to as *mentally retarded, dull, un-teachable, uneducable or children with emotional and behavioural disorders* (Kalanje, 2002). Mmbaga (2002) noticed another frustrating label, '*watu wangu*'. This is used by some teachers in Tanzania. Literally, the phrase means, '*my people.'* It is an expression used sarcastically, and in desperation, by teachers to refer to under–achievers who interact less with teachers in the classroom. The expression is synonymous with '*failures'*.

The tendency of teachers to prefer interacting more with high–achievers, and to neglect assisting the low–achievers, who often include those with reading and writing difficulties, has been noticed by Cohen and Manion, (1994). Kalanje (2002) has also noticed this tendency in Tanzania. Montgomery (1990), on the other hand, saw that the failure on the part of the teachers to see the difficulties of children in their classes as real, may blind them into thinking that they do not have any children with learning difficulty, but rather only lazy or careless ones.

The sixth motive is related to the establishment of a Children's Book Project in Tanzania (1991). Teachers and the general public have for a long time been, and are still, being called upon and encouraged to write books for children. Children on their part are encouraged to develop a culture of reading books. The schools are encouraged to set up school and class libraries for children to use. Without the skills of reading and writing, however, some students may never enjoy

reading and benefit from the books. The reading and writing culture will be enhanced if all children are assisted to learn the skills of reading and writing.

1.3 Aim of the study

The major focus of this study is on the identification of children at risk of reading and writing difficulties. Specifically, the aim is to create a theory-based group test in Kiswahili, to be used for identifying children in the first grade who are at risk of developing reading and writing difficulties, in a Tanzanian context. The validity of the grade one group test is studied by comparing with an individual test and through a follow-up test in grade two. The test is created and its function is presented based on a sample of first graders from the municipality of Morogoro in Tanzania.

In studying the function of the test there is special interest in the explanatory power of certain contextual factors related to the home (e.g., literacy) and school (e.g., school attendance). These contextual effects are of interest because the study is conducted in Africa and there is little knowledge about the roles of schooling and home related determinants outside Europe and the USA, or generally in developing countries.

1.4 Structure of the study

The study is arranged in seven chapters. The first is the introduction, where the background to the problem, the motives and the aims are presented. The second, third and fourth chapters deal with theoretical perspectives. Specifically, the second chapter involves the general understanding of the basic concepts of reading and writing. The third chapter involves the assessment of reading and writing ability and screening instruments. The fourth chapter focuses on contextual aspects, ending with a conclusion. The fifth chapter concerns the methods. It presents the specific tasks necessary for the construction of the instrument. It also highlights the nature and design of the study. The sample, procedure and questions of validity and reliability are also described in this chapter. At the end of the chapter data analysis issues and procedures and ethical issues are presented. The results of the study are presented in Chapter six and finally a discussion is presented in Chapter seven.

2 Understanding the basic concepts of reading and writing

This chapter attempts to analyze concepts and current theories basically related to the learning of reading and writing and the challenges encountered in the process. It concentrates on discussing the current understanding of the basic concepts of reading and writing and how children come to acquire reading and writing amidst the challenges of understanding the orthography of the language. Deep (opaque) orthographies and shallow (transparent) orthographies pose different challenges in some aspects and similar ones in other aspects as one learns a language. Also this chapter presents what reading and writing difficulties and risk factors entails. It is envisaged that the topics in this chapter will serve to form the initial theoretical basis needed for the construction of a screening instrument for identifying children at risk of developing reading and writing difficulties.

2.1 Review of some early and current definitions of reading

In this section, some definitions of reading are reviewed and their conceptual foundations are analyzed. However, the review is limited to only those definitions that form the basis for the current understanding of reading and writing.

From early times, reading and writing have been conceptualized and defined in different ways (Dallman, Rouch, Chang & Deboer, 1974; Downing, 1973; Stauffer, 1969). The definitions have ranged from simple to complex and from contrasting to complementing (Dallman et al., 1974). Reading is also viewed as a complex developmental challenge. It is interwoven with many other developmental accomplishments such as attention, memory, language and motivation (Snow et al., 1998). Not understanding this may lead to very simplified and unsatisfactory definitions of reading and writing.

The variations in conceptions on reading have, firstly, made scholars define it differently and this has led to misunderstandings in other related issues as a result of differing premises. Dallman et al., (1974) have observed both the complementing and the contradicting aspect in some definitions. Downing (1973), on the one hand, has blamed the ambiguity of the term reading as a source of error that has been affecting the conducting of cross-national comparisons of reading achievement. Secondly, following differences in conceptualization, reading and writing have been viewed and treated as two separate unassociated skills when teaching and learning.

As indicated above, many definitions of reading have been proposed over the years. Downing (1973, p. 32) provides two contrasting definitions on reading. The first one is by Elkonin:

Reading is a creation of the sound form of the word on the basis of its graphic representation. Therefore, a good reader is one who knows how to create the correct sound form not only of a known word, but also unknown word. Despite the fact that people often advance the comprehension of a word as a criterion of its correct reading, nevertheless, understanding is not an essential part of the process of reading.

The above definition puts emphasis on reading as the creation of the sound form of the word according to its graphic model. This definition neglects and excludes the comprehension aspect in reading. In contrast Downing (1973, p. 32) has also cited another definition by Tinker and McCullough (1975), which places emphasis on comprehension and avoids any mention of speech sounds:

"Reading involves the recognition of the printed or written symbols which serve as stimuli for the recall of meanings built up through past experience, and the construction of new meanings through manipulation of concepts already possessed by the reader. The resulting meanings are organized into thought processes according to the purposes adopted by the reader. Such an organization leads to modified thought and/or behavior, or else leads to new behavior which takes its place, either in personal or in social development."

Definitions which emphasize meaning such as the above have been more popular with Thorndike's (1917) influential statement after observing children's oral reading paragraphs. He concluded that:

"Understanding a paragraph is like solving a problem in mathematics. It consists in selecting the right elements of the situation and putting them together in the right relations, and also with the right amount of weight or influence or force for each. The mind is assailed, as it were, by every word in the paragraph. It must select, repress, soften, emphasize, correlate, and organize, all under the influence of the right mental set or purpose or demand."

For some more early definitions of reading that represent both the similarities among some points of view and the differences among others, (see Dallman et al., (1974, pp. 13-14). Generally, the definitions either take the *recognition of the printed word* stance or the *meaning* stance, or both. Some definitions are very inclusive and broad as to encompass the reading of signs, the clouds, or the flight of birds; thus the definitions move beyond the reading of written text.

Gough and Tunmer (1986) expressed their view of reading in their Simple View of Reading (SVR) model:

"Reading = Decoding x Comprehension."

Many researchers have been influenced by this definition (e.g., Adams, 1990; Høien & Lundberg, 2000). They have sided with the proposition that *decoding* and *comprehension* work together in normal skilled reading, although the skills are different in type. It has been observed that if one of the factors is zero, then the product is zero. It has been argued that good reading ability needs both.

Although generally accepted, the model has recently been questioned and challenged. Duke et al., (2006), have noted that the model has excluded some important variables such as vocabulary knowledge, motivation, and the cultural background of the reader and the speed of processing included in reading fluency. The proponents of the SVR model have defended it by elaborating that the implication was not that reading was that much a simple process, but rather the information-processing aspect of reading could simply be explained by the product of decoding and linguistic comprehension.

It is now understood that reading involves more than word recognition. Hence, comprehension is an essential part of reading because it is the actual reason for reading (Armbruster, 2001; Felton, 2001).Without comprehension, no reading takes place. Reading ought to be active and purposeful. Earlier, Moyle (1968, pp. 21-27) also wrote in support of the idea. Høien and Lundberg (2000, pp. 4-5) also in recent years have asserted that reading ought to be both decoding and comprehension. These authors observe the fact that decoding is the technical side of reading: seeing a string of letters and knowing that they represent, say, the word *nation.* Decoding, they contend, involves the ability to exploit the alphabetical principle, or code, in order to decipher written words. This component involves both laborious and time-consuming processes (such as sounding out letters and syllables) and the instantaneous, automatic word recognition that characterizes the good reader.

The comprehension component, on the one hand, in contrast, requires more in the way of cognitive resources. Comprehending a text includes such processes as connecting the text to one's own experiences and frames of reference, drawing conclusions from the text, formulating interpretations of it, and the like. Lipson and Wixon (1997) likened this kind of thought process, in principle, to the same kind of process that one engages in when listening to another person read aloud.

Most researchers in recent years (e.g., Snow, et al., 1998) also view reading as an active skill-based process of constructing meaning and organizing knowledge from oral, visual and written text. This is an inclusive definition which takes into consideration the visually impaired, by including Braille. Current authors perceive reading as a process aiming at the interchange of information and ideas. Through the application of their pre-knowledge, the use of their comprehension and decoding skills, good readers become fluent readers when reading. Authors maintaining this definition of reading, however, single out other types of reading such as music notation, pictograms, or the analogous reading in computer science that relates to acquiring data stored in the computer. Such kind of reading is non-textual.

Reading has also been related to the system of writing in the language. A writing system is the encoding machinery that records units of sounds into a set of symbols. Encoding, in this case, is what we know as spelling. Hence, spelling is the encoding process of turning units of sounds into symbols. In other words, while writing is a process of presenting speech in a more permanent visual form,

reading involves decoding those symbols back into sounds to formulate words. Thus, reading and spelling (writing) are reversible processes, and need to be taught in association or simultaneously in order to make obvious this reversibility (see Kahigi, 2003; McGuinness, 2004).

For some years now, there has been a change in conception and in practice concerning reading and writing. Since 1920, according to Lipson and Wixson (1997, pp. 11 –12), there has been an integration of reading and writing. This integration views reading and writing as closely related processes that also ensure that the code nature of the writing system is not obscured. Should reading and writing continue to be taught in isolation as if they were unrelated entities, i.e. using different words, different approaches and methods or teaching on different days, then learning to read and spell may remain a laborious and confusing task.

The premise on which this study is based is that reading is more than simply knowing what each letter of the alphabet means, as suggested in some definitions. Neither does this study include Braille nor does it consider other inclusive definitions such as those that include reading of signs, musical symbols, the clouds or the flight of birds. In the context of this study, reading is limited to the reading of written material. It is recognized that understanding what surrounds us may be essential to reading. However, this study adopts the definition which emphasizes reading as fluent mastery of a strategy according to which the child knows and uses correspondences between individual graphemes and phonemes and is also able to fluently decode pseudo-words. The latter ability is often used for assessing the accuracy of the phonological representations or phonological decoding skill (Holopainen, 2002; Siegel & Ryan, 1989).

2.2 Reading acquisition: A context for reading difficulties.

Understanding how normal reading is acquired is an important starting point, leading to an understanding of the difficulties that children face as they are set to learn to read (Holopainen, 2002). In order to create the measures that can tap into the reading difficulties of children, it is important to understand how children learn to read in the first place, and from that to discover the problems they face. Hence, in this section, some models through which it can be understood how reading skills are acquired are described and discussed.

There are a variety of models which provide theories about how reading is acquired. The dual-process model suggests that word recognition as studied in English orthography involves at least two relatively independent mechanisms (Castles & Coltheart, 1993; Jimenez, 2000; Stuart & Coltheart, 1988). The first is described as the phonological indirect or non-lexical route. It is also referred to as the sub-lexical procedure. This involves the use of phonological information. This process is used in the sounding out of unfamiliar words by blending grapheme-phoneme correspondences in order to identify a word. The

reader uses knowledge of grapheme-phoneme correspondences to translate the printed word to an internal phonological representation and this may ultimately be used to retrieve the meaning of the word. The second mechanism, which is described as the visual or direct, lexical route, uses 'direct' mapping from the visual word form onto word meaning, i.e. words are recognized as 'wholes', thus establishing the word's specific orthographic patterns.

The stage models of literacy development, similarly, describe the development of word recognition (Frith, 1985; Marsh, Friedman, Welch, & Desberg, 1980). In an initial logographic stage, children recognize words based on any salient visual and contextual features. In the phonological alphabetic stage of reading development, children acquire knowledge of letter-sound relationships that can be used to derive pronunciation for printed words. The final stage is fluent orthographic reading. The stage models of reading processes were formulated by Seymour and McGregor (1984), and later by Høien and Lundberg (1988). In relation to the models, the development of word recognition skills is described in four distinct stages: pseudo-reading, logographic reading, alphabetic reading and orthographic-morphemic reading.

Another model involves the use of analogical processes which have also been important in learning to read (Goswami & Bryant, 1990). According to the author, analogy in reading refers to the use of the spelling-sound pattern of one word, for example, 'light' as a basis for decoding a new word, for example, 'fight'. An example in Kiswahili would be 'mwanga' ('light') being the basis for decoding a new word 'twanga' ('pound' as with pestle and mortar). The similarity in spelling allows the inference that the pronunciation of the two words can also be analogous. The systematic relationship between letters and sounds form the basis for the analogous prediction. The proposition is that readers in the late stages of reading development could best use analogical processes.

However, the current models of reading acquisition are challenged and criticized (see e.g., Share, 2008), in the sense that they are mostly found based on data collected among children with English as their first language. English has a low consistency of grapheme-phoneme correspondence and is therefore characterized as an opaque orthography. To minimize this bias, more studies are needed that show the relationship between enabling skills and languages with more consistent orthographies (see Aro, 2004; Blachman, 1997; Wimmer& Goswami, 1994; Share, 2008).

2.3 The effect of orthography on reading and writing ability

Languages differ in the way they are written. This has an effect on trying to learn to read and write in that language. In this section, a discussion on the effects of orthography on reading ability is made, concentrating on the two major distinctions: transparent and opaque orthographies.

Orthographies refer to the differing ways of writing languages, which include the complex or simple linguistic information that is coded in the script. When one is able to understand the orthography of the language, then mastery of reading and writing becomes easier. According to Vellutino, Fletcher, Snowling and Scanlon (2004), orthographic awareness refers to the sensitivity to constraints on how the letters in written words are organized (for example, in English, *vid* is legal, *xqr* is illegal). Hence, McGuiness (2004) defines orthography as standardized spelling and knowledge of legal and illegal writing.

Orthographies are identified as graphic-syllabic (e.g., Japanese and Chinese); alphabetic-consonantal (e.g., Arabic and Egyptian) and alphabetic (e.g., Finnish and English). However, they all have a phonetic base in common. The representational levels are syllabic, consonantal and alphabetic. In each of these levels orthographies can be grouped following the coding of morphological information to script or whether the orthography is entirely phonetic. Orthographies that are purely phonetic (like Greek and Finnish) have consistent grapheme-phoneme correspondence. English and Korean, on the other hand, are extreme examples of languages in which there is a morphological inconsistency in grapheme-phoneme correspondences (cf. Aro, 2004).

It seems that it takes a longer time for children to learn to read an irregular (deep or opaque) orthography compared to those learning a regular (shallow or transparent) orthography. This was shown by Aro (2004) in his two separate studies. In the first one he studied children in grades 1 to 4 in pseudo-word reading, word reading and spelling skills after a year of instruction, in seven languages. The other one was after a year of instruction in thirteen languages. The author compared reading in Finnish (a transparent orthography) and English French, Portuguese, Danish (opaque orthographies). Another author (McGuiness, 2004, pp. 39-40) has also confirmed that it is easier to learn a transparent alphabetic language compared to opaque languages. The reason is that transparent codes are transparently reversible. The processes of encoding and decoding are distinct and easily seen. The sound /d/ is always written d, and the letter d is always read as /d/.

2.4 Reading and writing difficulties

Almost all children face reading and writing problems, in varied forms, when they begin reading. Most of them overcome the problems and come through to master the skills, leaving others to lag behind. For others, however, the problems persist until they complete school and go through to adulthood. These children have some kind of reading and writing difficulties due to a variety of factors.

Reading difficulties, *reading disabilities* and *dyslexia* are common concepts used in the LD field in referring to problems associated with reading. The terms are understood differently by different people. Often, the terms have been used interchangeably, although each of them has a distinctive meaning in the field of reading and writing. All the terms are identified under the larger LD term, which

encompasses different kinds of specific learning problems. In this section the terms are shortly defined and distinguished, thus justifying why the identification of children at risk of reading and writing difficulties in general was chosen in this study and the method of doing it.

Reading difficulties, specifically, is a general term referring to various kinds of reading problems, including dyslexia (also referred to as reading disability) and its other forms. Reading difficulties and the associated problems of writing and spelling have been extensively researched as individual aspects, but also to a great extent as a subgroup "dyslexia," within the much larger learning disability (LD) group (Kirk, Gallagher & Anastasiow, 2000).

There is no single universally agreed upon definition of dyslexia that exists to date. Many definitions are available. These definitions are mostly influenced by the author's or practitioner's professional background and what they have judged as the underlying cause of dyslexia. The following excerpts show a variety of definitions currently in use and suggested causes and of characteristics of dyslexia:

> "*A disorder manifested by the difficulty in learning to read despite conventional instruction, adequate intelligence and social cultural opportunity. It is dependent upon fundamental cognitive disabilities which are frequently of constitutional origin*." (World Federation of Neurology, 1968)
>
> *"A persisting disturbance in the coding of written language, which has its cause in a deficit in the phonological system."* (Høien & Lundberg, 2000).
>
> *"Dyslexia is a specific learning disability that is neurological in origin. It is characterized by difficulties with accurate or fluent word recognition and by poor spelling and decoding abilities. These difficulties typically result from a deficit in the phonological component of language that is often unexpected in relation to other cognitive abilities and the provision of effective classroom instruction. Secondary consequences may include problems in reading comprehension and reduced reading experience that can impede growth of vocabulary and background knowledge*." (This definition was adopted by the IDA Board of Directors, Nov. 12, 2002). It is also used by the National Institute of Child Health and Human Development, NICHD.
>
> *"Dyslexia is a specific learning difficulty which mainly affects the development of literacy and language related skills. It is likely to be present at birth and to be lifelong in its effects. It is characterized by difficulties with phonological processing, rapid naming, working memory, processing speed, and the*

> *automatic development of skills that may not match up to an individual's other cognitive abilities. It tends to be resistant to conventional teaching methods, but its effects can be mitigated by appropriately specific intervention, including the application of information technology and supportive counseling.*" (British Dyslexia Association, BDA, 1968).

From the four definitions given above, it seems to be evident that a child with dyslexia would most likely face a persisting difficulty in learning to read, write and spell; and with language. This is an unexpected difficult because the child may have good instruction, average intelligence, adequate sensory ability and a good socio cultural background (see, e.g., Puolakanaho, 2007). Hence, there is a discrepancy between the child's ability and what the child actually achieves in terms of reading, writing and spelling. Causes and difficulties that characterize a dyslexic child are many and varied. Dyslexia seems to be a genetically and environmentally transmitted linguistic disorder as a result of phonological deficit (e.g., Grigorenko, 2001; Pennington & Olson, 2005; Vellutino, Fletcher, Snowling & Scanlon, 2004).

As discussed earlier, there are no reasons why Tanzania should not have children who face reading and writing difficulties, to the extent of requiring special learning assistance and no reasons why there should not exist children with dyslexia. However, in Tanzania, there are no proper instruments for specifically diagnosing dyslexia or other specific reading *disabilities*.

It is in this context that this study uses the general and inclusive term *"reading and writing difficulties*." The focus is on screening children with reading and writing difficulties in general without specifying the subgroups.

2.5 Enabling skills

What reading and writing entails has been presented in the previous sections. It has also been discussed how reading is acquired and the challenges of orthography in the process. General reading and writing difficulties and some specific difficulties that can make individuals lag behind their peers as they learn to read have been discussed. This section presents and discusses the essential skills required to enable one to read and write.

There is large agreement among researchers in the science of reading, that deficit in the phonological system is an indicator of reading and writing difficulties (e.g., Catts & Kahmi, 2004; Høien & Lundberg, 2000; Kerins, 2006; National Reading Panel, NHI, 2000; Snowling, 2000; Vellutino et al., 2004). Deficits in core phonological processes seem to be central to many reading problems. These processes are often categorized into three areas: phonological awareness, rapid naming and phonological memory.

Phonological awareness involves an individual's awareness of the sound structure of spoken language. The phonological awareness is demonstrated when one can hear, remember and manipulate sound units within words, syllables, syllable fragments (i.e. onsets/rimes and phonemes (Bailet et al., 2009; McGuiness, 2004). Phonological and orthographic awareness are cognitions which are reciprocally related. They work together in sensitizing the irregularities and redundancy nature of the alphabetic writing system (e.g., *at* in *cat*, *fat* and *rat* and *ing* in *walking* and *running*). This relationship is important for success in reading.

Children in difficulty of acquiring phonological awareness and learning to relate alphabetic symbols to sounds will also face difficulties in acquiring orthographic awareness. Awareness and ability to manipulate sounds at the phonemic level is known as phonemic awareness. This awareness is further demonstrated in an individual's ability to analyze and synthesize speech sounds or segments, and blend sounds. Ample evidence is now available that demonstrates the contribution of phonological awareness to success in decoding (Vellutino et al., 2004).

Children with weak phonological awareness showed improved reading performance after intervention designed to improve phonological awareness (Kerins, 2006; Torgesen, Morgan & Davis, 1992). Deficit in phonological processing as the underlying cause of reading difficulties has been well-supported by studies conducted in different languages, including languages with transparent orthographies (Goswami, 2003; Lee, 2008; Wimmer, Mayringer & Landerl, 1998; Ziegler & Goswami, 2005). Bailet et al. (2009, p. 337) concludes by implying that "*Functional assets or deficits in these skills, measured in preschool and kindergarten, are strongly predictive of subsequent reading* (see also Anthony & Francis, 2005; Wagner, Torgesen & Rashote, 1994).

Phonological awareness is an important predictor of early reading ability in both opaque and transparent orthographies. Researchers have conducted studies in both orthographies to find out or verify the role of phonological awareness in those orthographies. For example, Lundberg, Olofsson and Wall (1980) investigated the possibility of predicting reading and spelling on measures of phonological awareness on Swedish children who they followed from kindergarten (6 -7years) through grade two (8-9 years). The results indicated that performance in phonics manipulation tasks in kindergarten was a strong predictor of spelling and reading in grade two. Phonological awareness was the most powerful predictor of reading ability among the tests administered. Bradley and Bryant (1983) extended the study of Lundberg, et al., (1980). They studied English children and concluded that knowledge of nursery rhymes enhances phonological awareness, which in turn is linked to success in reading.

Many other authors have shown the important role of phonological awareness in reading ability and distinguishing struggling readers and good readers in transparent languages such as German, Dutch, Spanish, Italian, Portuguese,

Greek, Turkish, Finnish, Indonesian, etc. (Aro, 2004; Carrillo, 1994; Denton, Hasbrouck, Weaver & Riccio, 2000; Cossu, Shankweiler, Liberman, Katz & Tola,1998; Durgunoglu, Nagy & Hancin-Bhatt, 1993; Gonzalez & Garcia, 1995; Hollopainen, Ahonen & Lyytinen, 2001; Winskel & Widjaja, 2007). Some authors (e.g., McGuiness, 2004; Mayringer & Wimmer, 2000) have, however, shown how reading in transparent orthographies, as compared to English, is easier using a phonological strategy. It is contended that transparent orthographies may not demand the same level of phonological awareness competence in the early stages of reading and spelling that English demands.

Rapid naming speed is an additional process that compounds phonological awareness deficits, making learning to read more complicated (Kerins, 2006; Wolf & Bowers, 2000). This was identified a quarter of a century ago by Denkla and her colleagues (see Denkla & Rudel, 1972). Naming speed, according to the authors, refers to the speed at which an individual can name objects, letters, or colors either as a discrete process or in a serial naming task. Naming speed is associated with an individual's acquisition of reading, specifically the ability to rapidly decode and read words (Blackman, Bruck, Herbert & Seiderberg, 1984; Bowers, 1995; Cutting & Denkla, 2001; Manis, Doi & Bhadha, 2000; McBride-Chang & Manis, 1996; Wolf & Bowers, 2000). Individuals with naming speed deficits coupled with phonological awareness deficits are found to be more resistant to intervention compared to individuals without these kinds of deficits (Kerins, 2006).

Phonological memory is the third process implicated in the acquisition of language and reading skills. This refers to the coding or imprinting of phonological information for temporary storage in the working memory prior to storage in the long term memory (Kerins, 2006; Wagner, Torgesen & Rashotte, 1999). The phonological loop associated with working memory provides a brief, verbatim storage of auditory information (Baddeley, 1986; Baddeley, 1992; Gathercole, Willis& Baddeley, 1991). The loop stores recent auditory information, which can be accessed by repeating it out loud. When the phonological loop is deficient, phonological memory is compromised, and thought to interfere with the acquisition of written and spoken words.

Some researchers have found letter knowledge, phoneme identification, and pseudo-word reading to be the most sensitive predictors and enabling skills of reading development and spelling, especially in transparent orthographies (e.g. Good & Kaminski, 2002; Holopainen et al., 2001; Jimenez Gonzalez & Hernandez-Valle, 2000; Leppanen, Aunola, & Niemi, 2008; Lundberg & Høien, 1989; Lyytinen et al., 2004; Molfese, Beswick, Molnar & Jacobi-Vessels, 2006; Stanovich, 2000; Winskel & Widjaja, 2007).

2.6 Children at risk of developing reading and writing difficulties

Which children are at risk of reading and writing difficulties? What do we know about them and their background? This section tries to elucidate this question based on research and experience.

There are variations of terms regarding students at risk of developing reading and writing difficulties. Such terms include: *disabled reader* (Downing, 1973), *backward student* (Bryant & Bradley, 1985), *poor readers* (Zabrucky & Ratner 1992), *remedial readers* (Duffy-Hester, 1999), *struggling student* (Hall, 2007), and *slow learner* (Rathvon, 2006).These variations are related to variations in the definitions of reading and writing and subsequent difficulties.

In a school perspective, at risk students are those in danger of failing to obtain an adequate level of educational skills. Many of these students have problems, which can be identified very early. The basic problems can be traced to physical and medical conditions before or soon after birth, or to environmental factors in early childhood (see e.g., Delgado & Scott, 2006). Regarding disabilities there is plenty of data indicating higher prevalence for boys that for girls (see also e.g. Delgado & Scott, 2006).

Relating to this study, an at-risk student is one who has fallen behind in academic performance in reading and language (Slavin & Madden, 1989). Students at risk begin school showing inadequacy in prior knowledge, in verbal abilities, in the ability of attending to the sounds of language as distinct from its meaning, and in understanding the basic purposes and mechanisms of reading and letter knowledge (Sandberg & Norling, 2006). Many authors (e.g., Elbro, Petersen & Boström, 1998; Høien & Lundberg, 2000; Locke, 1997; Scarborough, 1990) have noticed that children from poor home backgrounds, with little language proficiency, with pre-school language impairment, with impairments, and those whose parents cannot read, are at risk. These may come to school already facing problems in learning.

Children at risk of reading failure tend to be less motivated to engage in reading tasks compared to other children. They have more negative self-concept, feel more helplessness and they avoid reading activities more frequently than their typical peers. It is suggested that the reasons could be related to consequences of the children's repeated failure to acquire reading skills (Morgan, Fuchs, Compton, Cordray & Fuchs, 2008).In a study in Morogoro in Tanzania (Kalanje, 2002), regular teachers and parents were asked about their opinions regarding reasons for learning difficulties and reading and writing problems. They pointed out home and school environmental factors, the child's IQ and the child's and parents' personalities. Other respondents, however, strongly related risk factors to the students' mental and academic abilities. Children struggling to read and write are believed to be a result of their low IQ, or that of their parents. It is thought that this is the source of the children's low performance in other subjects too. Some respondents attributed children's difficulties in learning to laziness, poor school attendance, absconding from classes, truancy and running errands.

On the home environment, particularly, the parents and teachers pointed out the following factors: lack of support from the parents or guardians; lack of age-related books for the student to read at home; the child being pressurized with house-chores (especially the girls); the child being engaged in petty business, thus leaving the child with little or no time to do school or homework. Other factors in the home environment suggested include the child's house being positioned close to public places such as beer bars or market places, where the activities produce noise or undesirable behavior, thus making it difficult for the child to concentrate with school or home work. On the school environment, a common view was that poorly equipped classrooms, lack of proper teaching and learning materials, lack of class and school libraries and having incompetent and committed teachers influenced children's learning.

The described reviewed factors certainly have an impact and influence on children's development. Although the factors mostly refer to learning difficulties in general, they may apply to the risk perspective of learning to read and write. More specific studies that target at-risk children in early literacy skills are needed in Tanzania. Indeed, home and school environments have been reported to influence children's literacy achievement (see, e.g., Coker, 2006). However, there is insufficient cohort research in this field in Tanzania to justify all the factors suggested above. There is ample evidence from studies in other countries, for example, that contradicts some views above. Some studies have shown that average or even the most intelligent pupil experience difficulty in learning to read and write (e.g., Lewis & Doorlag, 1995, p. 66; Hamilton & Glascoe, 2006).

2.7 Summary

Chapter two has tried to be restricted to discussing reading in the form of written text as distinguished from other kinds of reading, which are very inclusive. It has attempted to show how reading in the form of written text has also been understood differently over time. However, to many authors reading is fairly complete if it involves decoding and comprehension as the basic components. Nevertheless, reading is a much more complex process, which involves other kinds of strands intertwined with decoding and comprehension. These are such developmental milestones as attention, memory, language and motivation. It is now also important to regard reading and writing as associated skills.

In this chapter, it has also been tried to show how the written text form of reading is acquired. The models reviewed in this chapter are as follows: 1.The dual process model, which involves two mechanisms, namely the phonological indirect or non-lexical route and the visual or direct lexical route. 2. The stage model of literacy development, which involves acquisition of reading through the initial logographic stage, the phonological alphabetic stage, with fluent orthographic reading as the final stage. 3. The model which involves the use of analogical processes. Although these models are widely used, it is questioned as to whether they can be generalized for application to languages beyond English.

Learning to read and write also requires understanding of the orthography of that particular language. In this chapter, it has been elaborated what a transparent and an opaque orthography entails and highlighted the challenges faced in learning to read in these orthographies.

Since the aim of the study is to create a screening instrument for identifying beginning readers at risk of developing reading and writing difficulties, it was also necessary in this chapter to discuss what is implied by reading and writing difficulties. Another task here was to distinguish the terms reading difficulties, reading disabilities and dyslexia. These terms are often misunderstood. Often they are used interchangeably. In this chapter they have been given distinct meaning and an argument for using the general term 'reading and writing difficulties' in this study has been given.

After an understanding of how reading is acquired, the effect and challenges of orthography and the understanding of what implies reading and writing difficulties, then the necessary skills enabling children to be able to read with ease were presented. Most authors have agreed that phonological awareness, rapid naming and phonological memory are strong predictors. Finally, also, children at risk of developing reading and writing difficulties were defined and discussed. Definition of reading and writing and what we define as reading and writing difficulties are determinants of which children we are going to subscribe as being at risk. However, most authors indicate those children who have fallen behind in academic performance in reading and languages are at risk. It is important to note, however, that poor home and school environments and lack of motivation also put the child at risk. In the context of this study a child at risk was one who scored significantly low in the measures of phonological awareness, word identification, letter identification and spelling halfway in grade one.

3 Assessment of reading ability and screening instruments

The following chapter involves a discussion on the current understanding and practice of assessment of reading and writing ability and what is understood about screening instruments in the process of identifying at risk children. The discussion focuses on the essence and nature of assessing reading ability. It also touches on the merits and demerits observed in the available screens. In this chapter also a topic about the creation and validation of screens is presented and some challenges encountered in the process revealed. The components comprising such instruments are also presented and it is shown how they should relate to the operational definition of reading and writing in relation to reading difficulties. The last topic in this chapter is a review of some already available screens, set as examples or models. The perspectives ventured in this chapter are deemed necessary as a theoretical basis for the practical work of constructing a valid and reliable screen.

3.1 Assessment of reading ability

The focus in this section is on the discussion on assessment of reading ability and the presentation of the general types of assessment processes. In this way, the position of screening in the assessment process is clarified.

Assessment is an important ongoing process of data collection for evaluation purposes. Strategies used in the assessment process include direct observation, testing with instruments such as criterion referenced tests, standardized tests or information gathering through interviews, surveys, checklists and questionnaires (Peterson, 1987; Rhodes & Shanklin, 1993; Wren, 2004).

On reading in particular, Wren (2004) contends that the first step in implementing good reading instruction is to determine student baseline performances. He argues that the diversity of students in a classroom in terms of background and skills in literacy may be learning at different levels. It is therefore necessary to design literacy instruction to meet individual needs that require review of basic skills in reading. Individual needs can be determined by initial and ongoing assessments.

Types of assessment processes (Peterson, 1987; Salvia & Ysseldyke, 2001; Wren, 2004) can be distinguished in a number of stages. The expanded six-stage model includes case finding or referral, screening, diagnosing or classification, educational assessment or instructional planning, performance monitoring or pupil progress evaluation, and program evaluation. Figure 1 elaborates the type of assessment processes involved.

				Program Evaluation	Documenting program effectiveness
			Performance Monitoring		
		Educational Assessment	Pinpointing a child's skill and learning needs		
	Diagnosis	Analyzing the nature and degree of the procedure			
Screening	Verifying if further evaluation is needed				
Case finding	Assessing the population of 0-5 year olds, and soliciting referrals				

Figure 1: An evaluation model for early childhood-Special Education (Source: Peterson, 1987, p. 283).

3.2 Screening and screening instruments

In this section a discussion on the process of screening as a second stage in the general assessment process is made in detail (cf. Figure 1). The concentration is on screening as related to the purpose of this study. A distinction between screening tests and readiness tests, diagnostic tests and IQ tests is made in order to avoid confusion.

Screening has its origin in the medical and health professions. But it also has a long tradition in education (Catts, 2006). In this study the term is used in relation to educational practice. In screening, tests may be administered to identify students who differ significantly from their peers (in either a positive or negative sense) and who therefore require special education services, or more specifically, targeted evidence-based interventions (Catts, 2006; Jenkins & Johnson, 2007; Scarborough, 1998; Tzivinikou, 2004).

Screening tests and readiness tests have been clearly distinguished by Meisels, Steele, and Quinn-Leering (1993). The authors contrasted the two in that screens focus on the child's potential to acquire skills important to school success, whereas readiness tests have the purpose of assessing whether or not the child has developed the skills needed to function effectively given the demands of a specific curriculum. Readiness tests are curriculum-loaded, while screening measures are intended to be curriculum-free and to predict accurately school-related difficulties.

Screening, according to the authors, is an initial or preliminary stage, during which those who may evidence a particular problem, disorder, disability or disease are sorted out from among the general population in order to pursue more extensive assessment (Meisels, 1994). Screens have been further seen as having the purpose of predicting an outcome months or years in advance (Scarborough, 1998; Tzivinikou, 2004). Instead of waiting for students to fail, as argued earlier in the introduction, screens identify those who are on the path to failure so that they can obtain immediate help by being given further individual diagnostic testing (Lindsay& Desforges, 1998; Tzivinikou, 2004).Screens for

reading ability can measure both oral and silent reading. They can be administered either on an individual or group basis.

Diagnostic tests are also contrasted with screening tests. Although they may be used as screens, they are, however, mostly designed to pinpoint with greater precision the nature of the problems that are contributing to the overall low score of pupils scoring low on the survey test or screen test. They are used in placement decisions (Mantzicopoulus, 1999; Meisels, 1994). Then remedial instruction can be appropriately planned and carried out. Hence, diagnostic reading tests are designed to fill this need, although they may be used as screens as well. Diagnostic tests can also test both oral and silent reading. The Woodcock Diagnostic Test Battery authored by Woodcock¸ Mather and Schrank (2004), for example, measures important dimensions of phonological awareness, phonics knowledge, reading achievement and related oral language abilities. It is an individually administered test for ages 2 to 90. However, these tests can only be meaningful and useful if they are carefully constructed with maximum reliability and validity, well-administered and used with great caution. According to Dallman et al., (1974), standardized tests must be interpreted with a clear recognition of their shortcomings.

Many screens for children at risk of developing reading and writing difficulties have partly similar content to traditional IQ tests. The content covered includes expressive and receptive language, quantitative concepts, and reasoning, comprehension, memory and perceptual-motor skills. Screens, additionally, also measure gross and fine motor skills, as well as social and communication skills. Also, both screens and IQ tests use correlations (Shepard, 1997). They are similar to the extent that both function as surveys. However, not all screens are necessarily IQ measures. Screens are rather short and can be administered to large numbers of children in a relatively brief period of time by examiners with little specialized training. Also screens are susceptible to numerous technical inadequacies as compared to (IQ) tests.

Various kinds of tests have been developed for screening purposes. Generally, there are tests such as vision and hearing tests for identifying pupils with vision or hearing problems; intelligence tests, which are administered to identify students who may need special attention, either because of limited intellectual capacity or because of highly superior ability; achievement tests, which are measures of what has been taught to and learned by students, and which are routinely given to identify students who are experiencing academic difficulty and for whom further assessment may be appropriate.

3.3 Challenges and problems with screens

The challenges and problems with screening tests and attempts to minimize them are discussed in this section. This is necessary in order to be aware of the possibilities and limitations in the creation of reliable and valid screens.

There are, indeed, challenges and problems associated with screening instruments, commonly used to identify young children at risk of reading difficulty. It has been observed (e.g., by Lindsay& Wendell, 1982; Raforth, 1988; Satz & Fletcher, 1988) that these measures are often characterized by questionable reliability and validity, lack of a theoretical base, and a tendency to predict global achievement rather than specific deficits in reading skills (Beers & Beers, 1980; Horn & Packard, 1985).

The controversy with the practice of using screens has also been observed by Gredler (1997), May and Kundert (1992), May and Kundert (1997) and Shepard (1997). These authors have argued that screening measures have been inappropriately used to assign young children to pre-kindergarten or pre-first-grade programs. The children have been held back from school or have been kept in the same grade. Consequently, this has led to unjustifiable labeling of the children.

One of the biggest challenges has been to determine who really has a learning difficulty (i.e. a reading difficulty). According to Snow, et al., (1998), the determination that a particular student has a learning difficulty is dependent upon the definition of reading difficulties. The author distinguishes two classes of reading difficulties within the traditional categorical approach: (1) students having a reading disability based on discrepancy between IQ test scores and reading performance, and (2) students having 'garden variety' reading problems or general reading backwardness. The latter are all other poor readers without the IQ / performance discrepancy, rather due to poor instruction and / or weak motivation. In this model reading difficulties are viewed as separate diagnostic categories.

The dimensional approach, on the other hand, assumes that human abilities, such as reading, are distributed in a statistically normal way along a continuous dimension (Snow et al., 1998). From this perspective, reading difficulties form the lower tail of a bell shaped distribution.

It should be understood that, traditionally, LD, including reading disabilities, have been defined on the basis of exclusion components and formula component using discrepancies relating to intelligence and achievement (Reid & Wearmouth, 2009). It has been implicitly assumed that measured intelligence has an intimate relationship with reading skill, thus children who are bright but could not read were considered to have 'unexpected reading failure'. On the other hand, children with below average intelligence were not expected to read well and therefore their reading difficulty was expected (Joshi & Aaron, 2009).

The IQ model for LD has been criticized by some researchers (e.g., Siegel & Ryan, 1989; Siegel & Lipka, 2008; Sparks& Lovett, 2009; Stanovich, & Siegel, 1994; Vellutino, Scanlon &Lyon, 2000). They argue that the model is problematic because some areas such as emotional factors cannot be objectively measured. The exclusion component was also unclear to indicate areas for assessment and what tools to use. Further, it is claimed that IQ test is irrelevant

in the analysis of LD due to there being too many variations in the type of tests used and also for the lack of agreement on the cut-off for average IQ.

The IQ scores have also been observed to be inadequate predictors of the cognitive processes involved in reading, spelling, language skills and memory tasks (Siegel, 1989; Siegel & Lipka, 2008; Stanovich & Siegel, 1994; Vellutino, Scanlon & Lyon, 2000). This is because reading skills and IQ are interdependent constructs in a way that language abilities make substantial contribution to both of them. And according to Scarborough (1998), reading difficulties may slow down the rate of intellectual growth and thereby obscure the discrepancy between reading skill and IQ.

Despite this critique, some scholars still support the use of ability-achievement score comparison when they are considered as part of a more comprehensive diagnostic model (Sparks & Lovett, 2009). However, although the discrepancy model may remain a very popular diagnostic method, in contrast, Siegel and Lipka (2008) have called for the establishment of a diagnosis for LD on the basis of achievement tests that indicate if the child has deficits in skills compared to his or her age group.

The componential model for diagnosing and instructing children with reading difficulties that focuses on the source of reading difficulty has been proposed by Aaron, Joshi, Gooden and Bentum (2008). (see Figure 2). In contrast to the discrepancy model reviewed earlier, it is then possible within the componential model to target remedial instruction at the source. The componential model of reading is an elaboration of the simple view of reading reviewed earlier, as proposed by Gough and Tunmer (1986), and validated by Hoover and Gough (1990).The model considers reading to be comprised of processes of word recognition, and comprehension aspects are themselves comprised of processes such as phonological awareness, decoding, listening comprehension and vocabulary. To be comprehensive, the model has an addition of psychological and ecological domains to the cognitive domain of the simple view of reading model of reading. Figure 2 is the graphic representation of the componential model of reading.

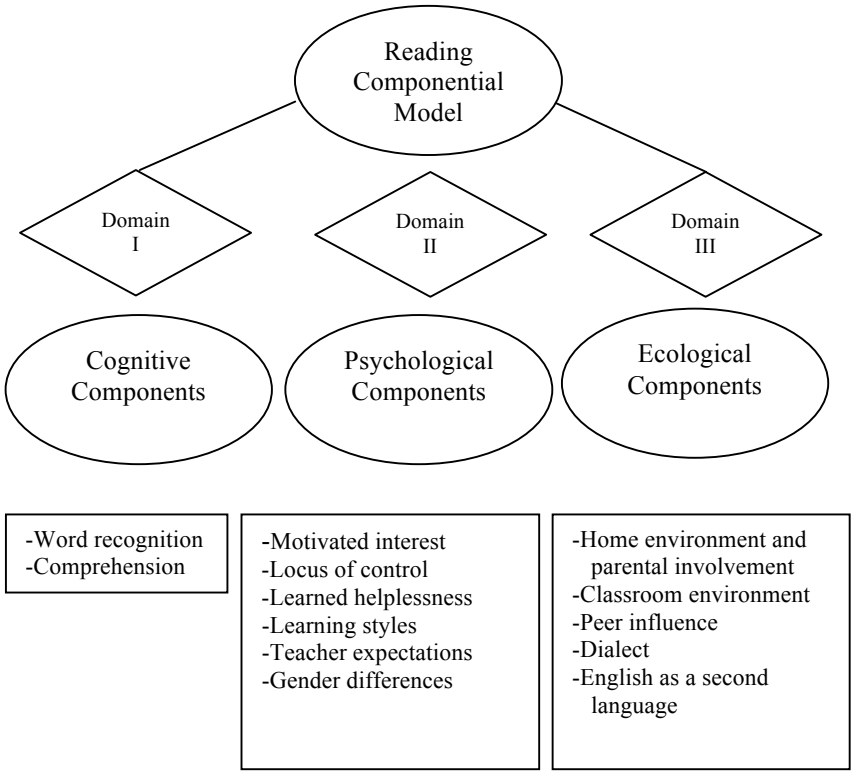


Figure 2: The componential model of reading (Source: Joshi & Aaron, 2009, p. 272)

There are other difficulties, specifically concerned with early screening measures for identifying LD children in general. One of the problems, according to Vervaeke, McNamara, and Scissons (2007), concerns the willingness among stakeholders to abandon formal classification of reading disability which focuses on identification of children at risk of becoming poor readers, rather than children who may have a learning disability per se. Within the early identification model, it should be necessary to distinguish children who are poor readers because of a learning disability, and those who are for a variety of reasons. It has been criticized that identification procedures that assess poor achievement such as reading failure with no regard for the reason for the problem cannot provide children with actual reading disabilities with the proper instruction needed. However, in response to this, based on research there is ample evidence to show that all children with difficulties in reading for whatever reason may benefit from early identification and intervention (Lyon & Fletcher, 2001).

Another difficulty, seen much earlier, is the discontinuity between the assessment tool and classroom instruction. According to Roth, McCaul and Barnes (1993), the tools designed have been administered outside the classroom

by non-classroom personnel. In effect, problems ranging from lengthy waiting lists to incompatibilities between assessment practices and classroom curriculum have arisen (Snow, et al., 1998).

The way of providing early and effective services have also been criticized by Lyon and Fletcher (2001). They oppose early identification tools and subsequent remedial programs as being simply ad hoc measures designed chiefly for ease of implementation, rather than a venture that aligns practice with research. According to Murray (2002), policy makers should call for the use of evidence-based practices.

3.4 Creation and validation of screening instruments

The concentration in this section is on highlighting some practical procedures of creating and validating screens and providing some examples.

The development of rapid assessment instruments (RAI) involves two primary components, each with its own series of steps (Springer, Abell & Hudson, 2002). First, conceptual design, which consists of (a) identifying and defining the construct of interest, (b) selecting the measurement tool format, (c) writing the items, (d) submitting the items for expert review on their appropriateness, both from content as well as a measurement point of view. Second, psychometric validation primarily consists of (a) determining the appropriate components for reliability and validity analyses, (b) designing the study, (c) administering the new tool along with the total measurement-packet to a sample, and (d) analyzing the data generated to assess the strengths and weaknesses of the RAI.

The creation of screens needs to go through three steps (Jenkins & Johnson 2007; Jenkins, Hudson & Johnson, 2007). The first step is to define the future outcome the screen seeks to predict (e.g. unsatisfactory reading ability, p. 2). The authors are keen to caution about the relativity of the future, which could range from several months to several years and is marked by specific points in the school curriculum (e.g., end of grades 1, 4, 8 and 12, p. 2). Reading screens attempt to predict which students will score poorly on a future reading test (i.e. the criterion measure). Norm-referenced test scores may be used for a criterion measure, defining poor reading by a score corresponding to a specific percentile (e.g., below the 10^{th}, 15^{th}, 25^{th}, or 40^{th} percentile, p. 2). But also poor reading may be defined according to a predetermined standard (e.g., scoring below "basic", p. 2) on the state's proficiency test.

The second step in creating a reading screen is the identification of early predictors of later reading outcomes. In order to for the screen to be effective, the authors emphasize that the screen needs to be sensitive to different levels of reading development. In the kindergarten the screen is expected to be sensitive to phonemic awareness, letter and sound knowledge and vocabulary, the linguistic aspects in which the children grow. In the 1^{st} and 2^{nd} grades the screen is expected to be sensitive to phonemic spelling, decoding, word identification,

and text reading, and in the higher grades on the aspect of comprehension of increasingly difficult texts (see also Invernizzi, Justice, Landrum & Booker, 2004/2005).

The authors present the third step as the determination of a cut-off point on the screening measure(s) that identifies children at risk of failing the future criterion test. In order to do so, working backwards is suggested by first selecting students who failed the criterion (later) reading measure, then identifying the score on the screening measure that best distinguishes those students from the students who passed the criterion measure (see also Invernizzi et al., 2004/2005).

Most authors (e.g., Invernizzi et al., 2004/2005; Jenkins & Johnson, 2007; Jenkins, Hudson & Johnson, 2007) also advocate that a screening device must be easily and efficiently administered, in a short time, with little formal training in its use, by a variety of persons, including special and regular educators and paraprofessionals. The authors also contend that the quantitative results of each screen should be easily understood and instructionally transparent.

Furthermore, screens need to be valid and reliable. A valid instrument measures what it purports to measure. This is demonstrated by comparing the results of an assessment with results obtained from other independent measures of the same constructs. Comparisons may be made with measures taken at the same point in time, i.e. concurrent validity or at future points in time, i.e. predictive validity (Invernizzi et al., 2004/2005).

3.5 Components for a screening instrument

In this section some important components for a screening instrument are highlighted. Basically, components for a screening instrument include a wide range of important different types of assessments that can be used for measuring development in reading skills. These can form components for screening instruments, depending on the predetermined objectives of a particular screen. Screens can accommodate tasks reflecting the following types of assessments: reading comprehension, language comprehension, decoding, background knowledge, linguistic knowledge, phonology, semantics, syntax, cipher knowledge, lexical knowledge, phoneme awareness, knowledge of the alphabetic principle, letter knowledge and the concept of print (see Social and Emotional Development and Learning, SEDL, 2010).

However, it is important to note that, firstly, components for a screen are essentially determined by the definition of reading and reading difficulties one subscribes to (see Dallman et al., 1974; Lee, 2008). Traditionally, screens have been developed based on IQ discrepancy and ability-achievement models, as reviewed earlier. In such cases, IQ scores and achievement scores have been important components for screens. This has not been necessary with screens developed within the simple view of reading, componential and RTI models (see Aaron, Joshi, Gooden & Bentum, 2008; Bradley, Danielson & Doolittle, 2005; Fuchs, Mock, Morgan & Young, 2003).

Secondly, the components for a screen are also determined according to their sensitivity to different levels of reading development. Most authors seem to be in agreement that battery screens for kindergarten, for example, should include measures assessing letter knowledge, phonemic awareness, phonological memory, rapid automatized naming, and expressive and receptive vocabulary. They contend that as children move into grade one, so screening batteries should include measures assessing phonemic awareness, decoding, word identification, phonemic spelling and text reading. By the second semester of grade one, it is recommended that the measures include speed as an outcome (Bishop, 2003; Compton, Fuchs, Fuchs, & Bryant, 2006; Foorman, Francis, Fletcher, Schatschneider & Mehta, 1998; Fuchs, Futchs & Compton, see 2004; Høien & Lundberg, 2000; Jenkins& O'Connor, 2002; Lipson & Wixson, 1997; McCardle, Cooper, Houle & Paul-Brown, 2001; O'Connor& Jenkins, 1999; Scarborough, 1998; Torgesen, 2002).

Batteries for grade two are recommended to include measures involving word reading, passage comprehension, fluency and phonemic decoding. The authors have also indicated the importance of vocabulary, language and comprehension in kindergarten through grade two. In the higher levels children gain the ability to comprehend increasingly difficult texts. Hence, the screen components at this level need to take account of this (again, see Bishop, 2003; Compton et al., 2006; Foorman et al., 1998; Fuchs et al., 2004; Høien & Lundberg, 2000, Jenkins & O'Connor, 2002; Lipson & Wixson, 1997; McCardle et al., 2001; O'Connor & Jenkins, 1999; Scarborough, 1998; Torgesen, 2002).

Thirdly, aspects of context are also important components to take into consideration when developing a screening instrument. Of these, the question of orthography seems to be important. As reviewed earlier, transparent and opaque languages may pose different but also similar challenges as one sets out to learn a particular language. It is therefore necessary to come up with relevant sensitive predictors in a screen that can work for that particular language.

Fourthly, other contextual aspects in the creation of the screen include cultural aspects. Culture may be expressed as attitudes, values, customs and language that family and friends transmit to children. These attitudes, values, customs and language have been passed down from generations of ancestors and have formed an identifiable pattern or heritage (Kirk et al., 2000).

Hence, language is a cultural aspect. It is a system of arbitrary verbal or vocal sounds used by people in a certain community with a common culture for the purpose of communicating. The ability to read and write effectively enhances our ability to communicate easily. However, the ability to read and write a language reflecting a foreign culture becomes limited. Comprehension in reading is enhanced as we read those aspects that are compatible with our experiences and frames of reference. On assessment measures, Kirk et al., (2000) contend that there is a concern that measuring instruments that may serve well for many children from the mainstream culture do not communicate

accurate information on students who come from different cultural backgrounds, and consequently this may result in inappropriate placement. The authors acknowledge the use of tests that provide for a past experience base for students.

From this context, the testing materials for the screening instrument in this study, firstly was needed to adhere to the Tanzanian traditional modes of testing that are familiar to young students in Tanzania. Hence, the reading and writing syllabus for standards one and two were referred to. Secondly, the concepts, objects, events or activities expressed by the words, sentences or paragraphs in the screening instrument needed to be culturally bonded by taking into consideration that a curriculum may be defined as selection from culture (Urevbu, 1991). The content selected for the measures in this study was culturally sensitive, relevant and acceptable to the children's age, culture and experience. Hence, the words and pictures used were understandable and culturally appealing to the children (see Arnold, Davis, Frempong, Humlston, Bocchini, Kennen & Lioyd-Puryear, 2006).

Fifthly, screens are more effective if a multivariate screening model is adopted, in contrast to the univariate model. Bishop and League (2006) found that a screen incorporating letter identification, phonological awareness, rapid automatized naming and phonological memory yields the highest correlations with oral reading fluency as the outcome measure. Other authors have supported the use of multivariate screening (e.g., Bishop & League, 2006; Compton, Fuchs & Fuchs 2007; Foorman et al., 1998; Jenkins & O'Connor, 2002). It is expected that the use of a combination of *word identification, letter identification, phonological awareness* and *spelling* in this study will boost the reliability of the group test.

Recommended target areas for early screening and progress monitoring as proposed by a number of authors is presented in table 4. This recommendation serves to illustrate and exemplify components for a screen at different levels of reading development.

Table 4: Recommended target areas for early screening and progress monitoring.

Measures	Recommended grade level	Proficiencies assessed	Purpose	Limitations
Letter naming fluency	K-1	Letter name identification and the ability to rapidly retrieve abstract information	Screening	This measure is poor for progress monitoring, since students begin to learn to associate letters with sounds. It is not valid for English learners in kindergarten, but seems valid for grade one.
Phoneme segments	K-1	Phonemic awareness	Screening and progress monitoring	This measure is problematic for measuring progress in the second semester of grade one. As students learn to read, they seem to focus less on phonemic skills and more on decoding strategies.
Nonsense word fluency	1	Proficiency and automaticity with basic phonics rule	Screening and progress monitoring	This measure is limited to only very simple words and does not tap the ability to read irregular words or multisyllabic words.
Word identification	1-2	Word reading	Screening and progress monitoring	This measure addresses many of the limitations of nonsense word fluency by including multisyllabic and irregular words.
Oral reading fluency (also called passage reading fluency)	1-2	Reading connected text accurately and fluently	Screening and progress monitoring	Although the measure has moderately strong criticism related validity, it cannot give a full picture of students' reading proficiency. Many students will score close to zero at the beginning of grade one. The measure still is reasonable predictor of end of year reading performance.

Note: K-1 = Kindergarten to grade one.

Source: Bishop, (2003).
A compilation based on Baker & Baker, 2008; Baker, Gersten, Haager & Dingle & 2006; Compton, Fuchs & Bryant 2006; Fuchs, Fuchs, Hosp & Jenkins 2001; Fuchs, Fuchs & Maxwell, 1988; Good, Simmons & Kame' enui, 2001; O'Connor & Jenkins, 1999; Schatschneider, 2006; Speece & Case, 2001; Speece, Mills, Ritchey & Hillman (2003)

In this study the components for the screen were determined based on ideas from the simple view of reading, the componential model and the response to

intervention model. However, expert advice and judgment were employed in order to come up with components that were perceived to be most sensitive to the context of the Kiswahili language, which has a transparent orthography, and for grade one level.

Figure 3 is a graphic representation of the conceptual model conceived from the literature reviewed in this study. It summarizes the reading and writing components as presented in the literature review in the context of this study. The model also indicates the proposed linguistic indicators and the tests as their measures.

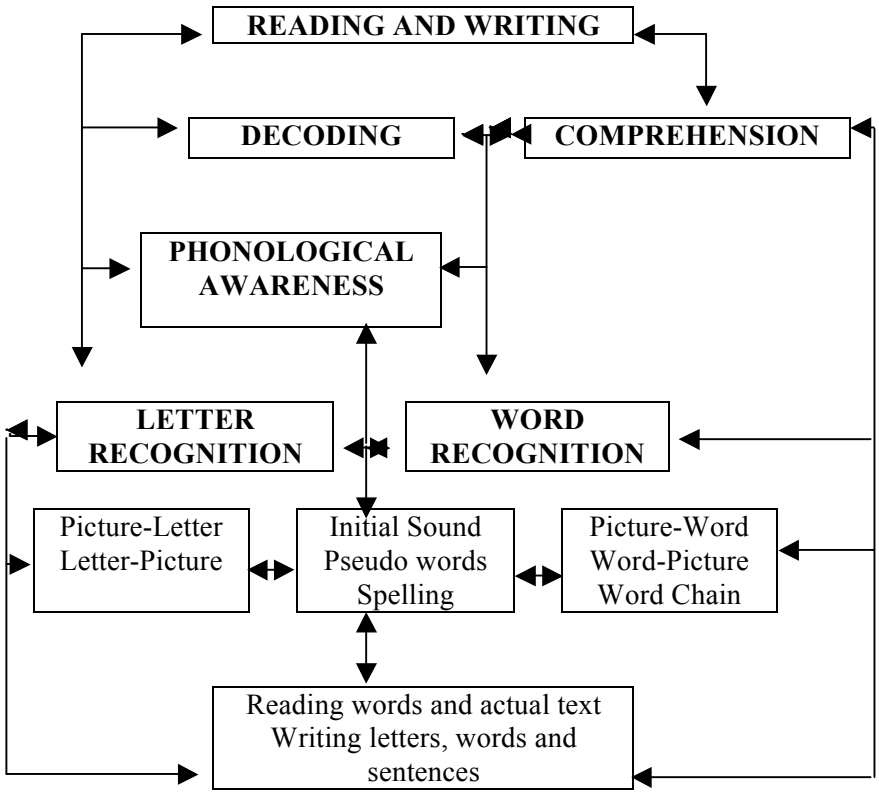


Figure 3: The conceptual model for test construction, in graphic form

The model above attempts to reflect and demonstrate the functioning of current theories in the acquisition reading and writing. If decoding and comprehension is the combined body that constitutes reading and writing, then phonological awareness is the heart that ignites letter recognition and word recognition, thus facilitating the whole process of reading and writing. According to Sousa (2005), successful reading is the result of the interaction between the decoding and

comprehension processes. Decoding includes phonemic awareness, phonics, and fluency. Comprehension requires adequate vocabulary and linguistic knowledge and interaction with the text to capture meaning. The model also indicates the positions of the selected linguistic predictors, which in this study form the test items or measures for identifying children at risk of developing reading and writing difficulties.

However, reading and writing involves much more complex intertwined aspects than depicted in this model. This model is sufficient for guidance on the construction of the desired screening instrument in this study. Home and school environmental factors are considered to have an influence in the aspects depicted. The conceptual model presented here is a product of a combination of the reviewed literature from which current theories of reading and writing have been extracted. The model is mainly based on ideas from the simple view of reading, the componential model and the response to intervention model. Expert advice from noted researchers in the field of reading and writing difficulties was also considered.

3.6 Review of screening instruments

The current study focused on creating and validating a survey test that would be administered as a group test, in Kiswahili, for the purpose of screening to identify first graders at risk of developing reading and writing difficulties. The children identified could then be recommended to undergo a diagnostic test. In this section some examples of screening instruments which share some elements with the current study are presented.

In Ontario, Canada, Harrison and Nichols (2005) validated the Dyslexia Adult Screening Test (DAST), which was created by Nicolson and Fawcett (1998). The screen was created following a demand for psychometrically robust screening tools capable of efficiently identifying students with specific learning disabilities (SLD), such as dyslexia. The researchers Harrison and Nichols (2005) investigated the ability of the screen to discriminate between 117 post-secondary students with carefully diagnosed SLDs and 121 comparison students. The117 (58 men and 59 women) participants with diagnosed SLD, were recruited from six post-secondary institutions. The participants ranged in age from 18 to 41 years (M=22.75, SD=4.77). These participants completed a demographic questionnaire and the DAST during the data collection period. The screen has 11 subtests, 9 of which tap skills in domains relevant to dyslexia such as *phonological awareness*, *auditory memory* and *motor coordination*. The tests include *rapid naming*, *one minute reading*, *postural stability*, *phonemic segmentation*, *two minute reading*, *backwards digit span*, *nonsense passage reading*, *one minute writing* and *verbal fluency*. In addition, two subsets, *nonverbal reasoning* and *semantic fluency*, are designed to probe areas of relative strength among individuals with dyslexia. The screen is administered individually and requires approximately 30 minutes completing.

The DAST as a screening instrument, however, differs from the instrument needed in this study, firstly in terms of the construct measured. Apart from *phonological awareness*, the instrument was specific to screen dyslexia. Secondly, the sample used was older (Mean =23) compared to the present study (Mean=7).

The general results of the study indicated that the DAST correctly identified only 74% of the students with SLD as 'highly at risk' for dyslexia. However, employing the cutoff for 'mildly at risk', the screen correctly identified 85% of the students with no major history of learning problems identified as 'at risk' for dyslexia.

In Malaysia, Lee (2008) carried out a study in an effort to develop and validate a reading-related assessment battery for the purpose of dyslexia assessment in alphabetic Malay language, and with a transparent orthography. The battery consisted of ten tests: *letter naming*, *word reading*, *non-word reading*, *spelling*, *passage reading*, *reading comprehension*, *listening comprehension*, *elision*, *rapid letter naming* and *digit span*. Content validity was established by expert judgment, while concurrent validity was obtained through regression analyses and factor analyses. The study reported that *phonological awareness* was the most significant predictor of word-level literacy skills in Malay, with *rapid naming* making independent secondary contributions. *Decoding* and *listening comprehension* made separate contributions to *reading comprehension*, with *decoding* as the more prominent predictor. Factor analysis revealed four factors: *phonological decoding*, *phonological naming*, comprehension and verbal short-term memory.

There are similarities and differences when comparing Lee's (2008) study and the current one. The former study is similar to this one as they both share the purpose of creating and validating a screening instrument in an alphabetic language (Malay for the former study and Kiswahili for the current one) with a transparent orthography based on a sound theoretical foundation. The two studies are also similar in some of the measures used for the identification of at-risk children. However, while the Malay study has dyslexia as the construct in focus, the current study focuses on general reading and writing difficulties.

Both the Malaysian study and the current study have included *letter naming*, *word reading*, *non-word reading*, *spelling*, *passage reading*, with the exception of *reading comprehension*. The former study has not used *initial sounds*, *chain word* and *one minute reading*, while the current study has not used *listening comprehension*, *elision*, *rapid letter naming* and *digit span.* There are similarities and differences, too, in the way of validating the developed instrument. The Malaysian study used school tests as criterion measures. The current study, in contrast, used an individual test, school marks and a Follow-up test as criterion measures for validation. There is no knowledge about how well teacher-made school tests in grade one in Tanzania measure emergent literacy skills.

Another screening tool that could be administered to children in the pre-kindergarten year by non-professionals who may know very little about emergent literacy and reading readiness, has been provided by Whitehurst (2003). The tool was created and expected also to produce a score for individual children that could be linked empirically to the potential for reading success. The researcher used the procedure of secondary validation of the screening tool. In the procedure the screening instrument was linked statistically to results from another professionally administered assessment, the gold standard that had already been validated as a predictor of later reading problems (Springer et al., 2002).

A strong correlation between the results from the screening tool and results from the same children on the gold standard would indicate that the screening instrument is also valid. According to the researcher, the procedure of secondary validation can be contrasted with primary validation, which would require that children assessed on the screening instrument be followed into school and tested on their reading skills two or three years later. While Whitehurst's study contrasts the current one in terms of the target group and the purpose of the screening tool created, nonetheless, there are similarities in terms of the procedure for validation.

The Makwami study, in Bagamoyo, on the acquisition of literacy in Kiswahili, is a large study in Tanzania performed by Alcock et al., (2000). In so doing they have gone beyond most previous educational research, which has focused more on the debate surrounding the use of Kiswahili versus English or other local languages as the medium of instruction in primary and secondary school. In their study the authors have been motivated by the need to develop a reading test suitable for administration by teachers with the equivalent of secondary school education, and for use with large numbers of primary school children as part of a study examining the impact of health on cognitive development.

In their first study, they first examined data obtained from existing reading tests: both standardized and teacher created reading tests. The 67 children's scores on both the tests fell into a bimodal distribution with all children's scores at the extremes and none in the center of the distribution. The authors then developed new tests and tested 314 children in order to avoid problems such as children reading aloud words that they cannot understand, children with non-zero knowledge of letters scoring at floor level and all children with any-word reading knowledge scoring at ceiling level. In the new test the authors used tests which consisted of three levels: a *letter*/*pseudo-letter* discrimination task, a *word*/*pseudo-word* discrimination task and a *silly*/*sensible* sentence discrimination task. The task demands and the distribution of the scores indicated that the tests are measures of both comprehension and decoding.

The current study has some similarities with the Makwami study, in the sense that both studies deal with the acquisition of literacy in Kiswahili and the use of tests with linguistic aspects. The other similarity is the administering of the tests

to a large number of children. The Makwami study used a sample of 67 children in their first study and 314 in their second one. The current study has used 337 children for the group tests in grade one, and a sample of 64 children from the main sample for the individual tests a month later; and 330 children for the follow-up, a year later.

A compiled list of more screening and monitoring measures is presented in table 5. The measures indicate their psychometric properties. They have been selected for use in this study to serve as close models, indicating similarities and differences with this study. It can be observed that some of the measures presented in the figure have been utilized in the present study, although at a different level of reading development of the children.

Table 5: Technical adequacy of screening and monitoring measures

Measure	Author	Validity	Reliability
Test of phonological awareness-Kindergarten	Torgesen& Bryant, (1993).	Concurrent validity with segmenting and sound isolation (.50-.55); Concurrent validity with word identification and word analysis of Woodcock Reading Mastery Test-Revised (60-.66); Predictive validity (.59-.75)	Internal consistency (.90-.91); Total score reliability (Cronbach's Alpha= .91)
Non-word spelling	Torgesen& Davis, (1996).		Internal consistency (.88)
Digit naming rate	Torgesen& Davis, (1996).		Split-half reliability (.91)
Bruce phoneme deletion test	Bruce, (1964).	Predictive validity to learning to read novel words (.67)	Cronbach's Alpha (.92)
Auditory analysis test	Rosner& Simon, 1971, cited in Macdonald & Cornwall, (1995). Yopp, (1988).	Predictive validity (accounted for 25% of the variance in word identification and spelling skills at age 17); Construct validity for compound phoneme awareness	Cronbach's Alpha (.78)
Rapid letter naming	Good& Kaminski (2000-2001).	Concurrent criterion-related with the Standard Diagnostic Test (.50) and oral reading fluency (.45)	Spearman-Brown Prophecy formula (.83 for first grade)
Segmenting fluency	Good& Kaminski (2000-2001).		Alternate form reliability (.60, Spearman Prophecy formula)
Oral reading fluency	Children's Educational Services, (1987).	Coefficient with Stanford Diagnostic Reading Test, Woodcock Reading Mastery Test-Revised, and Peabody Individualized Achievement Test (.52-.91)	Alternate form reliability (.97)
Nonsense word fluency	Good, (1998).	Criterion reliability with curriculum-based reading measures (.80)	Alternate form reliability (high .80s)

Source: Bruce, (1964).
(Compiled from: Children's Educational Services, 1987; Good& Kaminski, 2002; Rosner& Simon, 1971; Yopp, 1988)

Other examples of measures that are specifically for grades one through three include the following:

1. Test of Word Reading Efficiency (TOWRE): word identification fluency with reliability coefficients exceeding .90; alternate forms .95; test-retest .83-.96; inter-rater reliability .99 and validity ranging from .75-.90 (Torgesen,Wagner & Rashotte, 1999).
2. Woodcock-Johnson Diagnostic Reading Battery (W-JDR): word identification fluency with reliability coefficients exceeding .90; norms established with a population more than 8,800 (Woodcock, Mather & Schrank, 2004).
3. Dynamic Indicators of Basic Early Literacy Skills (DIBELS): oral reading fluency with reliability alternate form .92; test-retest .92-.97; concurrent validity with WJR .80 (Good& Kaminski, 2002).

3.7 Summary

Chapter three concentrated on discussing the current understanding and practice of assessment of reading and writing ability. The importance and the types of assessment in the continuum of learning have been highlighted. The discussion was necessary for establishing the position of screening in the whole process of assessment. In this chapter also an attempt has been made to distinguish screening from other kinds of assessments in the overall process of assessment. This distinction is important since this study is not focused on identifying the cause of difficulties or diagnosing specific types of reading and writing difficulties, but rather on screening for identification only.

The challenges that are associated with screening instruments for identifying children at risk of reading and writing difficulties have also been highlighted in this chapter. Understanding and being aware of problems is a crucial step towards constructing effective screening instruments. This chapter has also tried to show the practical procedures of creating a screen through selected study examples. Through the examples it can be learned how and why it is important to contextualize our own screens as we take notice of differing purposes for the screens created. It is important to understand the background contexts in which the instrument was constructed. Other contextual aspects discussed included cultural considerations in the form of the traditions and norms and customs maintained in society. These have an influence in the way classrooms are organized and the way the lessons are planned and executed. It has been shown why it is important to understand language as a cultural aspect used for communication among a people with common traditions, experiences and frames of references. This understanding has been posited as being essential to the planning and creation of a screening instrument.

Further in this chapter, it has been indicated how the components for the screen created in the current study have been determined. Based upon the definition of reading and writing difficulties based on the simple view of reading, the componential model and the response to instruction model, from whose concepts

the test measures and items were conceptualized and operationalized. (see Figure 3, The conceptual model for test construction, in graphic form).

Finally, the chapter has also reviewed a selection of model screening instruments. These were chosen based on their similarities to this study. The interest has been to study their psychometric properties and compare them with this study. This is important for future improvement of this undertaking.

Constructing a screening instrument for identifying children at risk of developing reading and writing difficulties is a process requiring careful and diligent planning. It involves various aspects that must be taken into consideration. It needs to be clear in terms of the objective and in the whole concept of reading and writing and the difficulties involved. The components for the instrument have to be related to the relevant language and its linguistic nature. Also the testing materials ought to have a cultural orientation. Not all developed instruments are successful in attaining their objectives. Administering the test and interpreting the scores is equally important. Haphazard planning may lead to unfair labeling, placement and segregation of the students. Although perfect screens do not exist (Jenkins & Johnson, 2007), adherence to the important issues discussed in this chapter relating to creation, relevant components and validation may minimize the possibility of constructing screens with questionable reliability and validity.

4 Kiswahili language and reading in the Tanzanian school

The following chapter presents some important themes on the Kiswahili language and literacy acquisition in Tanzanian schools. As the study aims to construct a screening tool in Kiswahili, it is imperative to understand the historical background of the language, its spread, and its role in Tanzanian society. Secondly, this chapter tries to explain elements of Kiswahili language that qualify its status as a transparent orthography. Thirdly, in this chapter, early literacy acquisition in Tanzanian schools is described. Some approaches and methods used in teaching reading and writing in grade one in Tanzanian school, the teaching of Kiswahili and also the assessment of reading and writing in the lower grades is highlighted.

4.1 Kiswahili language

It is important in this section to review the Kiswahili language. This is the language used in primary schools in Tanzania as a medium of instruction and as a subject in secondary schools and in higher institutes of learning. It is also a national and official language. The screening instrument in this study is created in the Kiswahili language. Hence, understanding the language's historical background, its spread, its role in society, its current status and its orthography provides a useful context in which the screen is created.

4.1.1 Historical background of Kiswahili

Many authors, local and foreign, have written about Kiswahili. Most of them have a common stand about the language´s historical background, origin, spread, and its role in society and its current status and orthography (see Hinnebusch, 1979; Hinnebusch & Mirza, 1979; Myachina, 1981; Nurse & Spear, 1985; Nurse & Hinnebusch, 1993; Shihabuddin & Mnyampala, 1977; Wald, 1987; Whiteley, 1969). Kiswahili is the Swahili word for the Swahili language, and this is also sometimes used in English. *Ki* is a prefix attached to nouns of the noun class that includes languages, e.g., *Ki*chagga, *Ki*nyasa, *Ki*masai, *Ki*ngoni, *Ki*zulu, etc. Kiswahili refers to the Swahili Language, Waswahili, or Swahili in English, refers to the people of the Swahili coast or also designates the culture of the inhabitants of the East African Coast, which is the coastal strip stretching from southern Somalia and extending to northern Mozambique. Uswahili refers also refers to the culture of the Swahili people. (A common colloquialism, *uswahili*, has been used for years in Tanzania as a derogatory term for base behavior or attitude. Its relationship to actual Swahili culture is unclear and somewhat controversial.)

The term Swahili, originated from the Arabic *Sahel,* meaning coast. However, Swahili language had its origin on the East African Coast, where it was used as a lingua franca and trade language starting from 7th and 8th Century A.D. The typical Bantu structure (the lexis and grammar) of the language, plus its

extensive stock of vocabulary that is Bantu origin, demonstrate that although Swahili has borrowed a large number of Arabic words, it remains a distinctly African language. In the modern era it has borrowed extensively from English. There is also a loan set from Portuguese, German, Persian, Hindi, etc. Such borrowing is comparable to the proportion of French, Latin and Greek loan words in English. Although this proportion of Arabic loans may be as high as fifty percent in classical Swahili poetry (traditionally written in Arabic script), it amounts to less than twenty percent of the lexicon of the spoken language.

Bantu languages cover East and Central Africa. Central Africa as far as Kiswahili is concerned includes such countries as Rwanda, Burundi, and the Democratic Republic of Congo. Politically, Central Africa also includes the Central African Republic, Gabon, and Congo Brazzaville, as well as Southern Africa (except Namibia).

4.1.2 The spread of Kiswahili, its role and status

Swahili spread through eastern Africa beginning in the 19th century, when Arab/Swahili trade expanded along the East African Coast, to Zanzibar and in trading centers in the interior. Long before the arrival of European colonizers, it was the Swahili dialect of Zanzibar Town (Kiunguja) that spread inland and eventually became the basis for standard Swahili in colonial and post-independence East Africa.

As is the case with all Bantu languages, a lot of specific vocabulary in the fields of science, technology, politics, psychology, etc. does not have vernacular equivalents. But because Kiswahili is also used as a medium of instruction in many Tanzanian schools, especially in state schools, there was a need to for such vocabulary to be formulated. Thus, the Institute of Swahili Research (known in Swahili as Taasisi ya Utafiti wa Kiswahili, TUKI) at the University of Dar es Salaam, and the National Swahili Council (known in Swahili as Baraza la Kiswahili la Taifa, BAKITA) were launched in Tanzania. The former is involved in neologism (i.e. coining new vocabulary) and the later in standardizing it for the general public.

Although English is still an important language in post-independence East Africa, Swahili plays an increasingly vital role in the daily commercial, political cultural and social life of the region at every level of society. This is especially true in Tanzania, where the language is used throughout the country in government offices, courts, schools and mass media. It has, in fact, become a more important language than English, and in some cases is replacing English as the language of choice among the educated. According to Rubagumya (2000, pp.112-120), concerning the ongoing debate about the use of Kiswahili or English in schools, both languages are seen as resources that teachers and learners can draw upon. Both languages should be encouraged and used systematically and strategically so that students can develop additive bilingualism.

Swahili, in fact, is spoken as a mother tongue by various groups traditionally inhabiting about 1500 miles of the East African coastline, and has become a second language spoken by tens of millions in three countries: Tanzania, Kenya, and Congo (DRC), where it is an official or national language. It is estimated that some 90% of approximately 39 million Tanzanians speak Kiswahili (Brock-Utne, 2001). Indeed, Kiswahili is widely spoken in Tanzania and the people are unified by it. The neighboring nation of Uganda made Swahili a required subject in primary schools in 1992.This mandate has not been well implemented. However, Swahili was declared an official language in 2005. Swahili is also used by relatively small numbers of people in Burundi, Rwanda, Mozambique, Somalia, Zambia, and nearly the entire population of the Comoros. The language is now the only African language among the official working languages of the African Union. Swahili is also taught in the major universities in the world, and several international media outlets such as the BBC, Voice of America, German DW and Xinhua have Swahili programs.

4.1.3 Characteristics

Languages are characterized by a particular orthography. Understanding the orthography of a language is crucial to the acquisition of that language. In this section the presentation focuses on the attempt to justify Kiswahili as a transparent (shallow) orthography.

Apparently, Kiswahili is an alphabetical language, which qualifies it to be placed in the realms of regular (transparent or shallow) orthographies (based on Aro's (2004) orthographic theory, and also Alcock et al., (2000). Kiswahili has a consistent G-P correspondence. The written code fully represents the phonetic structure of the spoken language. This contrasts with English orthography, from which the components for the screening instrument for identifying students at risk of developing reading and writing difficulties are reviewed.

Kiswahili has been supported to possess a transparent orthography (Alcock et al., 2000; Mohammed, 2001; Iraki, 2002; Kahigi, 2003; Gakuru, Kang'ethe & Ngugi, 2004). The authors have written about Kiswahili orthography and phonology. They elaborate by implying that, generally, the Kiswahili language uses the alphabet based on the Roman alphabet, which is also used in English: *a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z.* However, the Kiswahili language differentiates itself from the Roman alphabet in that Kiswahili language has *ch* instead of the roman *c*. Also, the Kiswahili alphabet does not include or use *q* and *x.* Generally, in Kiswahili, one grapheme represents one phoneme. For example, b represents */b/*, and d represents */d/*. However, the number of Kiswahili phonemes does not necessarily equal the number of sound letters of the alphabet. Appendix 1 is an attachment of the Kiswahili phoniset compiled from Alcock et al., (2000) and Gakuru et al., (2004).

In addition to the above mentioned phonemes, the Kiswahili language has the following graphemes: *sh, gh, ng, th, and dh*. Another grapheme *kh,* represents the *x* phoneme in certain Kiswahili dialects, which are more affected by Arabic. In standardized Kiswahili, *x* is manifested as *h*. Letter names in Kiswahili are syllables. According to Armbruster et al., (2001), a syllable is a word part that contains a vowel sound. In English orthography examples are *e – vent, newspaper*, *ver- y*. The letter names *a, b, c, d, f,* in English would be *a, be, che, de, fe,* etc. in Kiswahili. In Kiswahili orthography, all the vowels in Kiswahili can be syllables, as in ua (flower or kill)*, ita (call), elewa (understand)*, etc. All consonants are pronounced by placing a vowel after them, as follows: a, e, i, o, u; ba, be, bi, bo, bu; da, de, di, do, du; etc.

4.2 Early literacy acquisition in the Tanzanian school

In this section a brief highlight on the situation of acquisition of literacy, approaches and methods of teaching of reading and writing in the lower grades in Tanzania is made.. These are some of the factors that provide an important environment in which school beginners are set for learning. It is also in this context that the screening instrument in this study was conceived.

It is of interest here to understand that in Tanzania children are enrolled to begin primary school education at the age of seven. This is in response to the Tanzanian education and training policy, which states that primary education, is universal and compulsory from the age of seven years, until the end of the cycle of seven years of primary education (MOEC, 1995). However, for various reasons and problems encountered in a developing country like Tanzania, some pupils older than the official age may be found in many classrooms, especially now with the policy and call to enroll all children of school age (MOEC, 1995). Older students in grade one include repeaters or children who were unable to enroll at seven because of obstacles such as coming from families with low socio-economic background. Children younger than seven may also be found in the classes. These children may be enrolled if they have demonstrated the ability to follow the grade one curriculum. This ability may have been observed at home or at nursery school.

In Tanzania not all grade one students have gone through nursery school, although the policy requires this. The policy directs all public primary schools to have nursery schools (MOEC, 1995), but this has yet to be realized. Most nursery primary schools are privately owned and some parents cannot afford the fee demanded. Some schools may also be located at a distance that requires the parent to pay for transport. Some parents cannot afford this, too. Hence, for most children in Tanzania literacy development begins in grade one and extends to grade three. It is this period that reading, writing and arithmetic (3Rs) skills are oriented and consolidated. Hence, it is in this context that it was necessary to conduct this study with a sample from grade one instead of the kindergarten.

Kiswahili language, apart from being the national language, is currently also the medium of instruction both in pre-school and primary school (MOEC, 1995). In grade one the pupils are taught the skills of reading and writing in Kiwahili through class periods known as *kusoma* (*reading)* and *kuandika* (*writing)*. There is also a period allocated for *lugha* (*language*), which focuses on language usage (lexis). It is worth knowing also that English is taught beginning in pre-school and in grade one as a compulsory subject (MOEC, 1995). In the cities and towns in Tanzania, most children use Kiswahili in their homes and at school for communication. In the villages many children may also have the influence of their tribal vernacular as they learn Kiswahili and English.

4.2.1 Approaches to teaching reading and writing in lower grades

What types of approaches are used by teachers in the lower grades when instructing reading and writing? In this section the methods are presented. The approaches in use make a difference in terms of the effectiveness of how the children are going to learn to read and write. Teachers in the lower grades, especially in grade one, are expected to have undergone a specialized program or attended a seminar dealing with the three Rs (popularly known as KKK in Tanzania).The teachers are oriented also to various teaching methods for reading and writing. The approaches presented below are commonly used in Tanzania.

The whole word approach: The whole word approach, which is also popularly known as *look-say* instruction, involves the child saying the words that are pronounced by the teacher. The process requires the teacher to present a flash card with a word on it or to write the word on the chalkboard. The teacher often uses the demonstration method by pronouncing the word and then asking the child to say it as well. The process is usually started with a small set of words and gradually expanded.

The general rationale behind the whole word method of reading instruction is that the child does not recognize that the letters represent sound units, so the entire pattern of letters is taught holistically as representing a particular word (Rayner & Pollatsek, 1989, p.348). In Tanzania, common and frequent words such as *mama (mother), dada (sister), kitabu (book)*, and *kalamu (pen or pencil)* are presented on flash cards or written on the chalkboard with the accompaniment of pictures or real items. Reading for meaning in such a case is promoted at a very early stage of reading development. This approach is also particularly useful when teaching English, where many words are irregular in spelling**.**

The phonics approach: The phonics approach requires starting with a limited set of letters, which can then be built into many different kinds of words. Gradually, more letters are added and then the children are given consonant blends. Through this method the child also develops a sight vocabulary during the early stages, as some words keep recurring. The individual letters are taught by the sounds they make, and then children are induced to blend the sounds of more

letter combinations. According to Rayner and Pollatsek (1989), this is an analytic approach to learning words and is designed to exploit the alphabetic principle. Through this approach, Tanzanian children start to learn the vowels *a, e, i, o, u* and combinations to form words such as *au, oa* and *ua.* The children then are gradually led to learn the syllables, such as *ba, be, bi, bo, bu* and combinations to form such words like *baba, bibi, bubu,* and *beba,* as explained earlier. This approach has been criticized as boring to the child as it obscures the function of reading, i.e. extracting meaning from the print. Reading and Deuren (2007), however, have confirmed in their research that learning phonic awareness skills during the first grade supports grade level reading and that the learning can occur within a short time period, although learning these skills beyond a sufficient level does not necessarily result in improved oral reading. However, it is a very useful approach for teaching Kiswahili, where most words are regular in spelling.

The linguistic approach: This approach is halfway between phonics and the whole word approach. The method was advocated by Bloomfield in the early 1960s (Bloomfield & Barnhart, 1961). Teachers in Tanzania have used this method, especially in the 1970s.The method involves the child encountering a limited set of words similar in spelling construction that are combined to form sentences (e.g., *Katibukata amekataa kata kata*), or an English example, (*A fat cat ran after a bad rat*). The approach is similar to the whole word approach, except that a phonics type of decoding is also encouraged by the teacher. One disadvantage is that some pupils can simply read some text by simply looking at the picture illustrating the text by way of cramming.

Of the approaches presented above, the phonics method seems to be preferred the most considering the status of Kiswahili as a regular spelled alphabetic language. However, teachers are no longer limited to the use of the phonics method, as they also see some advantages in using the other two approaches. Hence, the common practice nowadays is an eclectic one.

4.2.2 Teaching to read Kiswahili among beginners

How are beginning readers taught to read in Kiswahili? According to Kahigi (2003), in the CV syllable cluster type, beginning readers are taught practice reading *ba,be, bi, bo,bu; da, de, di, do, du; ...*up to *za, ze, zi, zo, zu*. By connecting the syllables, words are formulated. For example:

ba-ba= *baba (father)*, **bi-bi** = *bibi (grandmother)*
a-na-ku-la = *anakula (a = third person + is eating i.e. John is eating).*

There is also the CCV syllable cluster. Teachers are expected to start with *ch, sh, ny,* by asking the pupils to pronounce various examples such as *chai (tea), chache (few), chuchu (nipple), shaba (copper), shaka (uncertainty), shati (shirt), shuka (bedsheet), nyanya (tomato), nyinyi (you, in plural form), nyigu (wasp),*

nyika (plain), and nyuki (bee). After that the teacher is directed to teach *ng, th, dh, and gh* as in *ng'ambo (across), ng'ombe (cow), ng'ata (bite), thelathini (thirty), themanini (eighty), dhani (conceptualize), dhuru (affect), ghala (storehouse), ghali (expensive), etc.* Beginning readers also are expected to be conversant with the CCCV syllable cluster. Examples are *shw* as in *angushwa (brought down); nyw* as in *nywea (shrink),* etc. The children also have to be conversant with various combinations of consonants to get varied sounds, as in the following:

(a) **Consonant clusters** (i.e. **consonant** + ***w*** or ***y):*** children are expected to be conversant with the following combinations:

bw	***fy***	***kw***	***mw***	***pw***
Bwana (mister)	*Fyeka (slash)*	*Kwapa (armpit)*	*Mwaga (pour)*	*Pwani (coast)*
Bweni(dormitory)	*Fyonza (suck)*	*Kwato (hoof)*	*Mwali (virgin)*	*Pweza (octopus)*

The above combination examples are best mastered if the children are given enough pronunciation exercises, e.*g. bwa, bwe, bwi, bwo, bwu; fya fye, fyi, fyo, fyu, etc.*

(b) **King'ong'o + consonant**: These are nasal sounds that are represented by the letter ***m, n*** and clusters ***ny, ng,***and ***ng'***. Examples are as follows:

mb	**nd**	**nj**	**ng**	**ng'**
Mbali (far)	Ndama (calf)	Njano (yellow)	Nguruwe (pig)	Ng'ombe (cow)

(c) **Other words differing in pronunciation but without differing in the clustering,** such as**:**

mb	***nt***	***nch***	***nj***	***ng***
Mbu (mosquito)	*Nta (wax)*	*Nchi (land)*	*Nje(out)*	*Nge (scorpion)*

The pronouncing of such words starts with n, which is syllabic and proceeded by pronouncing of *n + b; n + t,* etc.

(d) **Consonant + Consonant clusters**: Example of such clusters are as in:

bl (bluu=blue, blanketi=blanket)
sk (sketi=skirt, skonzi=scones)
ks (soksi=socks)
fr (friji=fridge)

(e) **When a vowel appears twice in succession, it is pronounced as a long vowel**. For example:

baa,	*tee,*	*hii,*	*choo,*	*suu,*
kaa,	*fee,*	*fii,*	*koo,*	*fuu,*
vaa,	*pekee,*	*halii,*	*popoo,*	*mkuu.*

(f) **Words having differing vowels in succession**. As in:

bai (bye), bia (beer), cheo (office position), doa (dirt-spot), chuo (college), hoi (exhausted), gae (tile), mkia (tail), nywea (drink from or shrink), poa (cool down), pua (nose), ungua (burn), hatoi (will not release), au (or), lia (cry or eat from), oa (marry), ua (flower), huo (that is), sinoi (I am not sharpening).

(g) Structural analysis skills

Structural analysis is concerned with identifying meaningful indicators in a word. As in many other languages, Kiswahili has a number of structures, e.g. *affix + root, compound words*, *shortened words*, etc. An affix is a structure that is attached to the root word. The root word cannot be broken further. The root word can, however, have a prefix before it (*Prefix-root*), as in ***m***-tu (person), ***m***-toto (child), and ***ji***-bwa (big and fierce dog). Also *(root-suffix*) as in shughuli-***ka*** *(actively busy)*, sali-***o*** *(credit),* starehe-***ka*** *(entertained or relaxed)*, burudi-***sha*** *(entertaining),* etc. Also (*prefix-root-suffix)* as in ***u***-karibish-***o*** *(invitation),* ***m***-sem-***aji****(spokesperson),****m***-som-***i*** *(an intellectual person or academician).*

Compound words are structured by two words. In Kiswahili, examples are as follows:

mwana*nchi (citizen),* ***mwenye****kiti (chairperson),* ***mbwa****mwitu (wild dog, jackal, fox),* ***mwana****mme (man),* ***mwana****mke (woman),* ***mwana****chama (member),* ***mwana****kijiji (village member),* ***mwana****jeshi (soldier),* ***mwana****michezo (sportsperson),* ***katibu****kata (ward secretary),* ***mchapa****kazi (hard worker).*

Shortened words results from shortening the second word, as in:
*mama yangu – mama**ngu** (my mother)*
*kaka yako – kaka**ko** (your brother in singular possessive form)*
*baba yetu – bab**etu** (our father)*
*mama yetu – mam**etu** (our mother)*
*mjomba wetu – mjomb**etu** (our uncle)*
*dada yenu – dad**enu** (your sister in plural possessive form)*

Each language follows some principles in structuring syllables that can be used. When words are broken into their syllables, the child is assisted in identifying words. When a child breaks a word into its syllables, in Kiswahili orthography

he/she can pronounce the syllables and join them in varied ways and is able to identify words which might be in his/her vocabulary, which he/she is used to hearing. For example, words such as *ua (flower or kill), mama (mother), kazi (work), panga (machete), tema (spit),kitu (something)*, futa (erase), when broken into syllables, that is in ***u****-a,* ***ma****-ma,* ***ka****-zi,* ***pa****-nga,****te****-na*, ***ki****-tu*, can bring about other words such as ***zi****ma (extinguish, i.e. extinguishing fire or whole i.e. one whole),* ***ka****ma (as),* ***ma****te (saliva),* etc.

In stage two beginning readers in Tanzania are expected to continue learning other clusters of **consonant** + **consonant**. Some of these are:

br*: breki, brashi (brake, brush****);***
ft*: daftari (exercise book);*
gr*: grisi (grease);*
kr*: shukrani (acknowledgement);*
lm*: almasi (diamond);*
tr*: treni (train)*
fr*: friji, friza (fridge, freezer)*
gl*: glopu, gluu (glope,glue)*
kl*: klabu, kliniki (club, clinic)*
lf*: elfu, alfajiri(a thousand, daybreak)*
st*: bustani (garden);*

Structural analysis skills are contained on the basis of singular and plural and word building. Many nouns in Kiswahili orthography are divisible into groups of singular and plural, such as:

Singular	**Plural**
Mtu (person)	Watu (persons)
Jiwe (stone)	Mawe (stones)
Kitu (something)	Vitu (things)
Meza (table)	Meza (tables)
Ukuta (wall)	Kuta (walls)
Ndizi (banana)	Ndizi (bananas)

In this case, by considering the reviewed components of reading and Kiswahili phonology and orthography, the test items that form the battery for identifying beginning readers at risk of developing reading and writing difficulties were constructed. In the context of this study the conceptualized and proposed model battery is graphically illustrated and presented in Figure 3.

4.3.3 Assessment of reading and writing in the lower grades

How is reading and writing assessed in the lower grades, in Tanzania? The following part concentrates on answering this question.

Reading and writing skills in the lower grades are assessed, like any other subjects, following set long-term curricula objectives and teachers planned short term and more specific and elaborated objectives. Children sit for examinations a monthly, end of term and end of year basis. The tests are teacher-made and

prepared solely for the purpose of determining the extent to which the syllabus has been covered. In most cases the tests are criterion-referenced. In this way, it is easy to know who is reading and who is not, but without informed knowledge as to who is in real difficulty and the nature of the difficulties and who is requiring specialized assistance. For the context of this study, however, the average score from three teacher-made test scores, i.e. *Kusoma* (*Reading*), *Kuandika* (*Writing*), and *Lugha* (*Language*) done in the middle of grade one in July, 2007 were used for sampling and validation purposes, as explained in earlier chapters.

4.3 Summary

In Chapter four, a presentation has been made on the source and nature of the Kiswahili language in terms of its historical background, its spread in and outside the country, its role in society and its current status.

It has been highlighted that the language originates from Bantu languages but is also influenced by some foreign languages. The language is widely spoken in Tanzania. It has been an important language during the colonial rule and during the struggle for independence. It has remained an important language for instruction in schools, for communication in everyday life, for business and official use. It has also been elaborated why the language is regarded as a transparent orthography and the challenges expected in learning the language have been indicated.

Other themes in this chapter relate to how and when children are engaged in early literacy in Tanzania, the methods of teaching in the lower grades, major aspects that are dealt with in teaching Kiswahili in lower grades and how reading and writing is evaluated in the lower grades. Awareness of all this is important for the choices made in constructing the screening instrument.

5 Method

Some conceptions of reading and writing, reading and writing difficulties, children at risk and screening instruments have been reviewed in the theoretical background. An overview of the components to be included in the screening instrument have also been presented and related to the context of the Kiswahili language and some other aspects related to the Tanzanian context in general. The next chapter is a presentation of the method used in realizing the general study design. The initial phase of creating the screening instrument started with listing some matters which should be taken into consideration in the construction and validation of the instrument. The following questions guided the process:

1. With what cognitive linguistic predictors can the screening instrument be constructed so that it can reliably and validly be used to predict reading and writing difficulties among first graders?
2. What test items should be included in the group test battery?
3. What test items should be included in the individual test battery?
4. What test items should be included in the follow-up test battery?
5. What personal, home and school environmental information of the children related to their reading and writing skills development can be incorporated as background factors?
6. How well can the group test be validated?

The presentation in this chapter is organized into eight sections (5.1 through 5.8). The first section (5.1) deals with the nature and design of the study. In the second section (5.2) the process of constructing the screen and validation is described. Other related tasks are also presented. The third section (5.3) involves the sampling of the area of study, the schools, and the participants. The pilot study comes in the fourth section (5.4). The actual study involving the group test and the individual test is described in the fifth section (5.5). The sixth (5.6) involves the description of the follow-up study. Data collection and procedure is described in the seventh section (5.7). Here, the subscales and their related tasks for the group test, individual test, and follow-up test and what they measure are described. The administration procedures marking and scoring are also presented. Data analysis procedures and the questions of reliability and validity are presented in the eighth (5.8) section. Ethical issues are also highlighted in this section.

5.1 Research nature and design

This section deals with the nature and the design of the study. This is followed by a graphic illustration of the design.

The design was longitudinal. It included, firstly, a group test and a questionnaire, and secondly, an individual test and a follow-up study for comparison and validation. The results of an established teacher-made test were also used for validation. For the Group test, 337 first graders from four purposively selected

schools in the municipality of Morogoro were administered a battery of seven different subscales. The results of these tests were used for comparison with the results of six subscales given on an individual basis and those of established teacher-made (school) tests. For the Individual test, 64 of the 337 first graders were chosen by stratified sampling, followed by a simple random sampling based on the results of the established teacher-made test. These were administered a battery of seven subscales. A year later when the children were in grade two, a Follow-up study was conducted. A battery of four subscales was administered to 330 children, since seven had dropped out of school. Figure 4 describes the basic research design for empirical data collection:

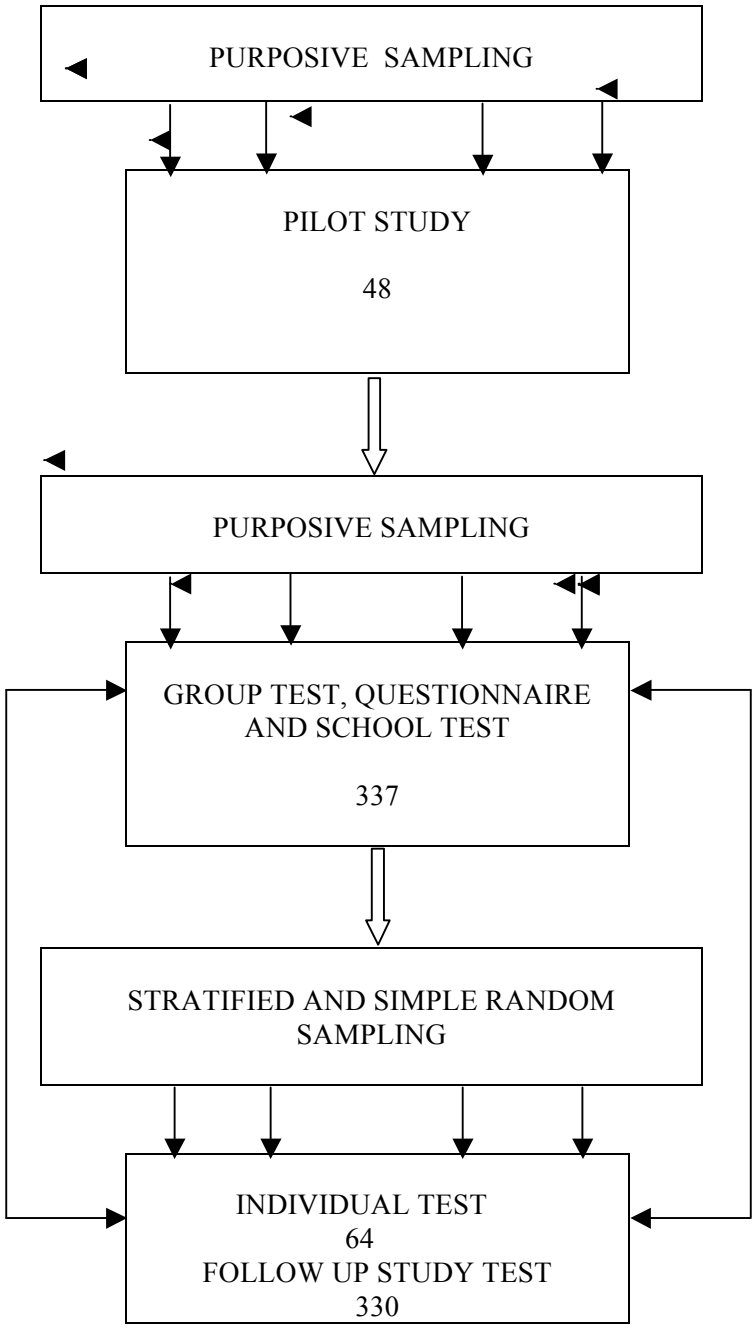


Figure 4: General study design and participant populations

5.2 Instrument construction process, validation, and related tasks

In this section a detailed process and steps that led to the creation and validation of the screening instrument in Kiswahili in the Tanzanian context for identifying first graders at risk of developing reading and writing difficulties is presented. The process and the tasks involved are related to both the theoretical and practical basis.

Based on my thirty years of teaching experience in primary schools and teacher training colleges, I noted that there were no theoretically-based screening tools in Tanzania, especially in Kiswahili, for identifying children at risk. Many children with reading and writing difficulties had not been identified and aided. This made people critical of the quality of education in Tanzania. "*How could some children complete school without being able to read and write?"* They questioned. The discussions I had with teachers and experts in the field in 2005 emphasized the need to develop measurement tools for early screening and identifying children in need of special attention and early intervention. Up to this point the teachers used only their experience and tests which lacked the cognitive linguistic indicators sensitive to reading and writing difficulties.

Clearly, there was a need to assist those children who lagged behind in literacy skills and minimize the risk of them falling into the 'Mathew effect' (see Stanovich, 1986). The 'Mathew effect' refers to a pattern of increasing advantage or disadvantage following initial advantage or disadvantage. The term comes from the Gospel according to St. Mathew: "For unto one that hath shall be given and he shall have abundance: but from him that hath not shall be taken away even that which he hath.'' In reading, the Mathew effect refers to the notion that 'over time better readers get even better and poor readers become relatively poorer' (Bast & Reitsna, 1998, p. 1373; Walberg and Tsai, 1983).

At this stage, I wrote a proposal for my intention of creating an instrument for identifying children who were beginning readers and at risk of reading and writing difficulties. I consulted experts in the Kiswahili and special education field at the University of Dar es Salaam, The Tanzania Institute of Education (TIE), The National Examination Council of Tanzania (NECTA) and Morogoro Teachers' College (MOTCO). I also consulted a number of Kiswahili teachers from primary schools in Morogoro municipality, especially those who teach in the lower grades, and grade one in particular. All of them had the opportunity to read the proposal. They all expressed the importance and need of having a screening instrument. They also provided invaluable advice on how to construct it.

The decision to create and validate the instrument was followed by a diligent reading of related literature starting from 2006 in an effort to identify the rudiments for such an undertaking and to identify the fundamental components for a screening instrument. The development of rapid assessment instruments (RAI) according to Springer and his colleagues (2002) involves two primary components, each with its own series of steps. First, conceptual design, which

consists of (a) identifying and defining the construct of interest, (b) selecting the measurement tool format, (c) writing the items, and (d) submitting the items for expert review on their appropriateness, both from a content and a measurement point of view. Second, psychometric validation primarily consists of (a) determining the appropriate components for reliability and validity analyses, (b) designing the study, (c) administering the new tool along with the total measurement package, to a sample, and (d) analyzing the data generated to assess the strengths and weaknesses of the instrument.

For the conceptual design my knowledge and experience were clarified by the literature reviewed on the definition and acquisition of reading and writing and the difficulties involved. At this juncture, more expert advice was also consulted, including professors in the field at Umeå University in Sweden and at Åbo Akademi University in Finland. The consultation confirmed the kind of tests to give and the mode and design for validation of the tool. The panel recommended having a word count of some Kiswahili books commonly used in grade one so as to distinguish high frequency (HF) words and low frequency (LF) words and use these words for constructing the relevant tests.

The words for the test construction were purposively selected. First, they had to be those found in the grade one curriculum and in the most common and popular grade one class readers[3]. The words had to meet the criteria as high frequency words. In order to distinguish high and low frequency words a manual word count was done on three class readers (Kihampa, 1997; Mkinga, 2000; Tanzania Institute of Education, TIE, 2000). The process involved finding out how many times each word used in a particular book appeared in the whole book by five stroke tallying. A word with three and more tallies was considered as a high frequency one, while one with less than three as low frequency. High frequency words from all the books made one list, and likewise with low frequency words. Both the lists are presented in Appendix 2.

The tests used high frequency words and were constructed by starting with easier words and then moving to more difficult ones except for the picture-letter, letter picture, initial sound and the writing of the letters of the alphabet in the writing test. In these scales the items were mixed up so that the children would not easily guess what letter would follow in the arrangement. Short high frequency words without clusters or with a single cluster and no succession of clusters were considered easier to read. Longer high frequency words without clusters or with single or a succession of single consonant clusters and with a nasal sound were considered more difficult to read. Low frequency words were very limitedly and sparingly used. These were considered to be more difficult to read. They were used when it was felt or observed that the high frequency words were exhausted. Table 6 presents an example of the selection of the words for the tests:

[3] Class readers refer to Kiswahili textbooks authorized for use in grade one for reading and instruction.

Table 6: Selection of words for the test construction

Word status	Without clusters	With clusters/nasal
HF short	wa, na, kaka, ana, mama	mkate, mboga, tembo
HF long	anakula, watoto, sufuria	mwanafunzi, gwaride thelathini, mchwa
LF short	kasirika,	kunguru, bunduki
LF long	harakaharaka	kimeanguka, ng'ang'ania

Note: HF=high frequency; LF=low frequency

The definition of reading and writing, reading and writing difficulties and children at risk in the context of this study were determinants for the selection of the cognitive linguistic predictors, which are *letter recognition*, *word recognition, phonological awareness* and *spelling*, as latent constructs. The conceptual model for the determination of the linguistic factors and the test construction is shown in Figure 3. The expert panel also suggested and verified the most effective linguistic predictors regarding Kiswahili with its highly transparent orthography. Based on convenience and appropriateness, I created a group test with seven subscales manageable to test 337 first graders, and an individual test with six subscales manageable for testing 64 pupils. For the follow-up I created a battery of four subscales. A questionnaire was also included for probing school and home background factors. Before the pilot study, the tests were tried out in small groups of grade one children and individuals around the Morogoro Teacher Training College neighborhood. They were also analyzed by teachers. The tests were also analyzed by Bachelor of Education students (Special Education) at the Open University of Dar-es-Salaam, who were attending a face-to-face seminar. Some necessary changes were made.

Administration of the tests was done in mid-July, 2007, for the group test, with a sample of 337 children. At this time the children and their parents also responded to a questionnaire on home and school background factors. A month later in August, 2007, the individual test battery was administered to 64 children purposively and randomly sampled from the main sample. A year later, half-way through grade two, a follow-up test was administered to 330 children. Seven children had dropped out. The scripts for all the tests were marked as scored by the grade teachers from the sampled schools, under the researcher's supervision. Ten percent (a computer based random sample) of the group test scripts were sent to Vaasa, in Finland for cross-marking by an independent marker. No

systematic errors were found. The new scores on some scripts did not alter the general performance of children in the group test.

5.3 Sample

The following section involves a description of the criteria used in selecting the area for the study, the schools and the participants. The basis of selection is argued for, and Morogoro municipality is described in terms of its geographical position, socio-cultural environment, socio-economic environment, educational status and physical environment, factors that apply to all the schools. In the context of this study, it is anticipated that the differences or similarities of the position and status of the sampled schools will provide useful information when interpreting the results.

Area: Morogoro municipality

Geographic environment: Morogoro municipality, from which the schools for this study have been selected, is situated on the foot of the Uluguru range of mountains. It is positioned in the eastern part of Tanzania, with some 600,000 inhabitants. Morogoro municipality has a metropolitan characteristic. For this reason, it attracts many people looking for employment in factories, sisal estates, as well as people seeking jobs in public services, hotels and offices. Hence, its population has been continually increasing. The municipality, also known as Morogoro Urban District, is the regional headquarters and the main administrative town for Morogoro region. It is one of the six districts that comprise the region. The other districts are Ulanga, Kilombero, Kilosa, Morogoro rural and Mvomero. It is positioned west of Pwani (coast region) and some 200 km. from Dar-es-Salaam, the current capital city of Tanzania.

As a region, to its west Morogoro is bordered by the interior and mountainous Iringa region. To the north is Tanga. To the northwest and near the central part of the country is Dodoma region. Dodoma city, which is designated to be the future capital city, is only some 400 km. away. To the south Morogoro is bordered by Ruvuma region. This region also has a variation of physical features. The profile is such that the land rises from the coastal plains to the mountainous Uluguru and Udzungwa ranges. The vegetation also varies from scrub to thick forests. The temperatures also vary from warm in the lower areas to cool in the mountains. The savanna in Mikumi area harbors a variety of wildlife.

Socio-cultural environment: The indigenous people are the Luguru, Kutu, Pogoro, Sagara and Ndamba, all of which are Bantus. However, due to its position and reachability, Morogoro has neighbors such as the the Zaramo, Zigua, Kwere, Gogo, Ngoni and Hehe, and also other tribes such as the Chagga, Maasai, Meru, Haya, Nyamwezi, Nyakyusa, and many others in the town centre especially. There is a good representation of both coastal and hinterland indigenous tribes; agriculturalists and pastoralists; low and higher land dwellers;

Swahili speakers only or as first language, and Swahili speakers as second language and urban and rural dwellers. Such a variation of people and culture favors Morogoro's representation in the study.

Socio-economic environment: People's social relationships are influenced by their economic activities such as farming, business, herding and working in business, industrial companies, public or private offices and estates. The people's socio-economic status ranges from very low to very high profile.

Educational status and environment: Morogoro has a variation of educational status and environment. The range is from places with no school at all or far away to places with an abundance of schools. School variety also ranges from pre-schools to higher institutes of learning. Hence, some children will come from very well educated parents and also from parents with no formal education at all. The municipality has high and low capacity schools, and high and low performance schools. Children in schools mostly come from urban locations. However, some schools in urban settings also have children from the rural suburbs.

The schools: The area for the study was Morogoro municipality in Morogoro region. Four schools in the municipality were involved in order to validate the instrument. The schools were purposively selected because of their different performance status among the 37 primary schools in the municipality.

The performance status of the schools is defined and described according to information from the District Education Office (DEO) and the School Inspectors' Report (SIP) at the time of the study. Schools in the municipality may be inspected at any time of the year as needed, but all of them must receive a major inspection yearly based on the rubrics or criteria set by the inspectorate. Certificates of merit are issued for the major inspection, which indicate how the schools have fared in the inspection. Many aspects are considered before a school can be judged as high, good, average or low. These include teachers planning their lessons from syllabi and schemes of work and using the plans; teaching in the classes; preparation of teaching and learning aids and using them; children getting food at school; the availability of adequate school buildings, furniture, library and other facilities. The inspection also covers the effectiveness of the school administration and the managing of the school finances.

It is assumed that if a school is faring well in many aspects, this should be reflected in the various examinations that children take, including the Grade four National Examination and the final Primary School Leaving Examination (PSLE) taken at the end of each year by grade seven children. The following is the list of schools selected for the study and their descriptions.

SUA Primary School: This school was selected and coded as a *high performance* (HP) school. The school is situated about five kilometers from the town center. The school is in the Sokoine University of Agriculture (SUA) campus. Most of the university staff and lecturers send their children to this school. Some children

come from the neighborhoods. The school enjoys good support from the parents through a strong parents' committee at the school. It extends up to grade seven. Up to the time of the study it was rated as an excellent school by the inspectors' report and it was the best performing school in the Primary School Leaving Examinations (PSLE).

Kigurunyembe Primary School: This school was selected and coded as a *good performance* (GP) school. The school is situated five kilometers from the town center. The school is in the Morogoro Teacher Training College Campus. Most of the college staff and tutors send their children to this school. The school also absorbs children of staff and teachers from a nearby Lutheran Junior Seminary. However, the school also absorbs many children from the nearby Kigurunyembe suburb. It extends up to grade seven. Up to the time of the study it had been rated as a good performing school by the inspectors' report and it was performing well in the PSLE.

Mwande Primary School: This school was selected and coded as a *peripheral average performance* (PA) school. The school is situated about five kilometers from the town center and about a kilometer away from Kigurunyembe Primary School. It is outside the Morogoro Teacher Training College campus. Most of the children enrolled in this school are from the outskirts of Kigurunyembe and some from further away. The school is relatively new compared to the other three schools, as it was built to support Kigurunyembe School. At the time of the study, it had only extended up to grade four. Up to the time of the study it was rated average by the inspectors' report and was performing fairly well in the grade four national examinations.

Kilakala Primary School: This school was selected and coded as a *low performance* (LP) school. The school is situated about three kilometers from the town center. It is close to the Field Force Unit (a police force) campus, from which some children come. But also the school is close to numerous mini-markets, local beer shops and bars, trading stores and centers of small businesses. The children come from around all these places. The school extends to grade seven. Up to the time of the study the school was rated low by the inspectors' report and it had not been performing well in the PSLE.

Participants: Grade one children

For the context of Tanzania and this study, it seemed reasonable to identify at risk children when they are in grade one. Grade two was excluded because in this grade the children are advancing and learning more skills in reading and writing. It would be late to identify at risk children in this class. Identification of at risk children done by the end of the year in grade one or the end of grade four examinations, as it has been done in Tanzania, may be too late.

Findings from research indicate that 88 percent of students who had difficulties in learning to read at the completion of the first grade continued to have difficulties at the end of the fourth grade (Juel, 1988). However, Deshler, Ellis

and Lenz (1996) noticed that children with reading and writing difficulties still can learn despite the enormous problems they face. They found that when the diagnosis is made in the first two grades, more than 80 percent of the children are brought up to grade level. In contrast, if it is not made until fifth grade, only 10 to 15 percent are helped (Kirk et al., 2000). Researchers at Yale University (USDE, 2001) found that the problem continued through high school. Rathvon (2006) confirms the above reviewed findings by reporting that 88 percent of children who are poor readers at the end of first grade are poor readers at the end of elementary school. 75 percent of poor readers who are unidentified until third grade never catch up; regardless of the length and type of remedial services they receive.

Pre-school would be the ideal place for identification, since some children begin reading very early (Haats, 2002). In Tanzania, formal education starts at age 5-6. However, although the policy emphasizes that each public primary school has to have its own nursery school, for this age group, for practical reasons, this has yet to be realized. For this reason, and probably many others, still not all children in grade one in Tanzania pass through nursery school. Hence, the sample for this study was based on grade one children.

5.4 The pilot study

This section involves a description of the pilot study. The participants for the pilot study are described (Table 7). The outcome and suggestions for improving and amending the subscales are presented.

The participants for the pilot study were, firstly, 48 children random sampled in grade one from four purposively selected primary schools in Morogoro municipality. These were similar in characteristics to those in the actual study. The schools were selected on the basis of performance status following the most recent inspectorate report. The composition of the sample for each school status was N=12. Gender balance was observed. Table 7 shows the sample composition for the first phase (pilot) of the study:

Table 7: Sample composition for the first phase (pilot) of the study

	HP	***GP***	***PA***	***LP***	***Total***
Boys	6	6	6	6	24
Girls	6	6	6	6	24
Total	**12**	**12**	**12**	**12**	**48**

Note: HP=high performance; GP=good performance; PA=periphery average performance; LP=low performance.

The pilot study, as *the first phase* of the study, was conducted for the purpose of further refinement of the test items. During the pilot study it was also necessary

to reduce the number of items in some tests, as it was observed that a test with a large number of items bored the pupils. The panel scrutiny involved experts, grade one teachers, some Kiswahili teachers and the researcher. The recommended number of test items by the Ministry of Education for grade one in Tanzania is between twenty and thirty. The number of picture-word and word-picture items was reduced from the originally intended twenty to ten for each subtest. This is because it was observed that the pupils slowly got tired and bored with matching pictures, as they were not used to performing a task with many items. The number of pseudo-words was also reduced from the originally intended forty to thirty. The pupils were not used to reading non-words. Some pupils said that it was '*English*' since they could not get the meaning of the words. They said that their teacher had not taught them such words. This implied that the examiners had to be very cautious and careful during the administration and when giving the instructions for the tests. It was essential that the pupils really understood what each test required them to do.

Another major change that was done during the pilot study was to type the tests and abandon hand-written tests, as some of the test items were confusing the pupils. For example, the hand written pseudo-word *maandagi* was read as *maandazi (buns)* which is a real word, and the pseudo-word *vangi* was red as *rangi (colour),* a real word. The Kiswahili lower case *z* is written to resemble *g* and also lower case *v* may be easily confused with lower case *r* when hand-written. The time spent for the children to complete the tests was also observed. This was important for planning the administration schedules. The pupils, too, were observed as to how they reacted to the tests. Generally, most of them showed a positive attitude. It seemed play-like to them. The teachers also liked the tests. Some gave comments like, *"They are very good!"* Others said, *"They are challenging!"* Others remarked, *"They provide a different way of testing pupils' reading and writing skills!"* And others judged, *"They are demanding, but good!"*

5.5 The main study

The main study involved, firstly, administering a Group test to 337 grade one children, from four selected schools based on performance status, the filling in of the questionnaire by parents and teachers to tap home and school background factors. This was done in mid-July, 2007, when the children were half-way through grade one. Secondly, it involved selecting 64 children representing the four schools from the main sample for the Individual test. This was done in August, 2007. The detailed process is described in this section.

The main study involved 337 first grade children ranging from age 6 to 12, with a mean age of 7.25, from the four purposively selected primary schools in Morogoro municipality. Gender-wise the sample consisted of boys (N=181) and girls (N=156). Since the study was focused on regular schools, which have yet to fully accommodate inclusion, the number of disabled children was very small

(N=5, 1.5%), mainly the physically impaired and with albinism. Also, the children in the sample generally enjoyed good health, with very few pupils suffering from chronic illness (N=5, 1.5%) mainly asthma, diabetes and sickle cell anemia. The schools were selected on the basis of performance status following the inspectorate report. Hence, the cohort was made up of children from *high performance* (N=107), *good performance* (N=80), *periphery average performance* (N=71) and *low performance* (N=79) schools. The children also responded to the questionnaire on home and school environmental factors. Table 8 shows the sample composition for the study.

Table 8: Sample composition for the second phase (actual study with group testing)

	HP	***GP***	***PA***	***LP***	***Total***
Boys	64	43	34	40	181
Girls	43	37	37	39	156
Total	**107**	**80**	**71**	**79**	**337**

Note: HP=high performance; GP=good performance; PA=periphery average performance; LP=low performance.

The second phase of the main study was conducted a month later, in August, 2007. It involved selecting 64 from the 337 first graders who were administered the group test a month earlier. The selection was done by stratified sampling, followed by a simple random sampling based on school marks. The school marks were obtained from a teacher-made test done in mid-July, 2007. The score for each child was the average obtained in *reading, writing* and *language* tests in Kiswahili. The children were selected based on the four quartiles representing their general positioning on the School marks. The quartiles were the '*strata'* interpreted as *Excellent, Good, Average, and Poor*. This is in accordance with Tanzanian primary school grading system of internal testing. Hence, this is how 'stratified sampling' was used in this phase of the study. The school marks were also used as criterion measure for validating the group test.

From each stratum, only four children were selected based on simple random sampling. This was done by picking out two boys' names and two girls' names with a '*YES'* marked piece of paper from a pool of children's names randomly coded either '*YES'* or '*NO'*. The first two boys and two girls to be picked were the representatives for that particular stratum. In this way, each type of school (HP, GP, PA and LP) was represented by (N=16) children, that is N=4 from each quartile, with a gender balance. The selected children were then administered the Individual test. Table 9 shows the sample composition for the third phase of the study.

Table 9: Sample composition for the third phase (actual study with individual testing)

	HP	*GP*	*PA*	*LP*	*Total*
Boys	8	8	8	8	32
Girls	8	8	8	8	32
Total	**16**	**16**	**16**	**16**	**64**

Note: HP=high performance; GP=good performance; PA=periphery average performance; LP=low performance.

5.6 The follow-up study

In this section, a description of the third phase of the main study, which involved a follow-up study, is presented. This was done in mid-July, 2008, when the children were half-way through grade two.

The third phase involved a follow-up study. The procedure was similar to that of the group test in grade one. Four new subscales were created through a similar procedure as the original tests administered as group test in the *first phase*. The new test for the Follow-up study included the following subscales: spelling, picture-word, initial sound and word-chain. The test was administered in mid-July, 2008, to 330 children when they were half-way through grade two. As expected, some children had repeated grade one (N=49), but all of them participated in the follow-up test. The number of repeaters from the different schools was as follows: high performance school (6), good performance school (10), periphery average performance school (16) and low performance school (17). Seven children were not available for testing due to drop-out. Four children had dropped-out from the Good performance school and three from the Periphery performance school. There were no drop-outs from the highest and lowest ranked schools. The final sample composition in the follow up study is presented in Table 10

Table 10: Sample composition for the fourth phase (actual study with follow up group testing)

	HP	*GP*	*PA*	*LP*	*Total*
Boys	64	41	31	40	176
Girls	43	35	37	39	154
Total	**107**	**76**	**68**	**79**	**330**

Note: HP=high performance; GP=good performance; PA=periphery average performance; LP=low performance.

5.7 Data collection and procedure

Measurement instruments: When the study was conducted, primary schools in Tanzania began the school year in January and ended in late November. During the school year there are three mid-term holidays, with the major one coming in

June. Hence, the group test battery was administered to 337 grade one pupils in July, 2007. At this time, the children were halfway through grade one. The battery included the following seven subscales: word-picture, picture-word, letter-picture, picture-letter, initial sound, word-chain and spelling. At the same time, the children also responded to a questionnaire relating to their home and school environment. The child's parent, guardian or the teachers assisted in giving information when the child could not provide it.

The individual test was administered to 64 grade one pupils in mid-August, 2007, a month after administering the group test, as the second measure for validation. The test battery included the following subscales: pseudo-word, reading words aloud, actual text reading, one-minute reading, and writing. Comprehension was administered but omitted from analysis because it was found that what was measured was listening comprehension instead of the intended reading comprehension. During administration of the test the administrators read the sentences for the children instead of letting the children read them and comprehend.

The follow-up study was conducted in mid-July, 2008, administered to 330 children, a year later when the children were half-way through grade two, as explained earlier. In this section, detailed descriptions of each of the three batteries (group test, individual test and follow-up test), the questionnaire, and their administration schedules and the requirements or demands for each subscale are presented. Reliability scores for all the scales included in the tests are indicated.

5.7.1 The group test battery and questionnaire

The group test battery (Appendix 3) included measures of word recognition, letter recognition and phonological awareness. A questionnaire was also given to parents of each participant in the study. Teachers also had their part to fill in the questionnaire. Measures of word recognition included three subscales, namely word-picture, picture-word and word-chain. Measures of letter recognition included two subscales, namely picture-letter and letter-picture. Measures of phonological awareness included two subscales, which are initial sound and spelling. These were chosen based on literature review and following expert advice on indicating the most effective predictors for reading acquisition in a transparent orthography. It was also judged that these subscales were manageable, and could easily and effectively be administered to a group of pupils, considering the large class sizes[4] in many Tanzanian primary schools. The total was 7 subscales. The schedule for administering the group test and a description of the group test subscales are presented in the following part. The questionnaire is also described.

[4] The standard class size in Tanzania is 45 pupils. Large class in this case means more than 45 pupils. At the time of this research class sizes could be as large as 80 and more.

The schedule and administration of the group test

The administration of the group test battery was done in four days beginning on 18th July 2007 and ending on 21st July 2007. Each day the first session began at 8.00 am and ended at 10.00 am. Another session started at 10.30 am and ended at 12.30 noon. It was necessary and ideal to administer the test in two sessions because each school had two grade one streams. The number of children in each stream was manageable for supervision and invigilation.The test administration was carried out in one of the four schools in the four sampled schools, i.e. Kigurunyembe (GP), Mwande (PA), Kilakala (LP) and SUA (HP). Table 11 presents the schedule for the administration of the group test.

Table 11: Schedule for administering the group test

Date	***School***	***Session1*** *08.00- 10.00*	***Session2*** *10.30- 12.30*	***Total***
18/07/2007	Kigurunyembe	39 tested	41 tested	**80**
19/07/2007	Kilakala	40 tested	39 tested	**79**
19/07/2007	SUA	53 tested	54 tested	**107**
20/07/2007	Mwande	35 tested	36 tested	**71**
Total		**167**	**170**	**337**

Before administering each subscale, it was first necessary to give the children relevant and clear instructions on how to do it. Three examples were given by using the chalkboard for every child to see, follow and practice. All the group test subscales were untimed, but each lasted for approximately twenty minutes. However, the children were instructed to work as fast as they could and to put a pencil mark at an appropriate space as demanded by the relevant scale, and not to leave any space unmarked. Two teachers and the researcher were the examiners in each session. They administered, invigilated, marked and scored the test. The examiners went around the classroom to offer assistance to any child in need. Also, they made sure that the children worked silently, independently, and that each space was marked on each child's script.

However, our control to make each child put a mark on each space was limited to the extent that a child showed the willingness to do so. Some children showed signs of becoming tired, bored or wanting to go outside despite having not finished their work. The test administrators refrained from forcing these children to write anymore. This resulted in some children having blank spaces towards the end of their scripts. It should be noted that a wrong response, a blank space or marks on every alternative given instead of choosing the one correct alternative, was awarded a zero. This may explain the many zeros on the last items of the subscales. After each task the children were allowed a five-minute break to relax, play around and be ready for the next task. The administrators tried to make the whole exercise enjoyable for the children as much as possible.

Measures of word recognition

Word-Picture: This subscale measures word recognition in terms of identifying one correct picture among the four pictures presented that corresponds to the given word. The pictures used were familiar to the children. The subscale was used as a variation of the word-picture interference task introduced by Rosinski, Golinkoff and Kukish (1975) and following expert advice. The child was required to identify ten pictures by putting an 'X' mark on the correct picture for each of the ten words given. The words were given in lower case. The minimum score was (0) and the maximum was (10). The scores were converted into percentages. The scale was not timed. However, it was observed that some children were able to complete the task in 5 to 10 minutes, while it took 20 minutes or a little more for others. The mean score of the word-picture subscale was 59 (SD = 35). The internal reliability was .93 (Cronbach's alpha).

Picture-Word: This subscale measures word recognition in terms of identifying one correct word among the four words presented that corresponds to the given picture. The pictures used were familiar to the children. The scale was used as a variation of the word-picture interference task introduced by Rosinski et al., (1975) and according to expert advice. The child was required to identify ten words by drawing an 'X' mark on the correct word for each of the ten pictures given. The words were given in lower case. The minimum score was (0) and the maximum score was (10).The scores were converted into percentages. The scale was not timed. However, it was observed that some children were able to complete the task in 5 to 10 minutes, while it took 20 minutes or a little more for others. The mean score of the picture-word subscale was 60 (SD = 34). The internal reliability was .92 (Cronbach's alpha).

Word-Chain: The word-chain subscale consisted of twenty-five word-chains, with a total of sixty words to recognize. The scale was intended to measure word recognition. All the words used for this scale were from the list of high frequency words obtained after the word count of three common Kiswahili books used in grade one and selected on the basis of an item analysis during the pilot study and according to expert judgment. The words were given in lower case. The chains were constructed and listed in order of increasing difficulty. Examples of the chain words are *au/oa (or/marry), mama/tembo (mother/elephant),* and *kichwa/mkate/nyoka (head/bread/snake).* The scale demands rapid reading of word chains of two to three words without space between them. The pupil's task was to recognize and separate the words by making dividing lines between them. The scale was scored basing on the number of correctly recognized words. The minimum score was (0) and the maximum was (60). The scores were converted into percentages. The mean score of the word-chain subscale was 50 (SD = 33). The internal reliability was .97 (Cronbach's alpha)

Measures of letter recognition

Letter-Picture: This subscale measures letter recognition in terms of identifying one correct picture among the four presented that corresponds to the given letter. The pictures used were familiar to the children. This scale was also adopted as a variation of the word-picture interference task introduced by Rosinski et al., (1975) and following expert advice. The child was required to identify twelve pictures by drawing an 'X' mark on the correct picture for each of the twelve letters given. The letters were given in lower case. The minimum score was (0) and the maximum was (12). The scores were converted into percentages. The scale was not timed, but it was observed that some children completed the task in 5 to 10 minutes, while it took about 20 or a little more for others. The mean score of the letter-picture subscale was 56 (SD = 34). The internal reliability was .94 (Cronbach's alpha).

Picture-Letter: This subscale measures letter recognition in terms of identifying one correct letter among the four letters presented that corresponds to the given picture. The pictures used were familiar to the children. This scale was also adopted as a variation of the word-picture interference task introduced by Rosinski et al., (1975) and following expert advice. The child was required to identify twelve letters by drawing an X mark on the correct letter for each of the twelve pictures given. The letters were given in lower case. The minimum score was (0) and the maximum was (12). The scores were converted into percentages. The scale was not timed, but it was observed that some children completed the task in 5 to 10 minutes, while it took about 20 or a little more for others. The mean score of the picture-letter subscale was 57 (SD = 34). The internal reliability was .94 (Cronbach's alpha).

It should be noted that the picture-letter and letter-picture tasks involved 12 letters of the Kiswahili alphabet for each of the task. This is because the Kiswahili alphabet uses 24 letters. The idea was to use all the letters of the alphabet for the tasks. In both the tasks, however, the used letters were not given in their chronological order they appear in the alphabet i.e. *a, b, c, d,…z*. Rather, they were presented in mixed form so as not to prompt the children to guess the upcoming letter. It was reasoned that the mixed form would provide a good enough distraction. For example, in the letter-picture task '*z*' was presented earlier than 'a'. In learning, the children first start with the vowels *a, e, i o, u* and later the consonants *b, c, d…z*. It should be expected that the vowels, especially those that come first: *a, e,* and *i,* would be easily remembered and recognized by the children than those vowels and consonants that come later in learning. However, proactive and retroactive interferences may influence children to forget something learned later or earlier.

Measures of phonological awareness

Initial Sound: This subscale measures phonological awareness. The Initial Sound test comprised ten picture items familiar to the children and used in their class readers. The pictures were also selected based on item analysis during the pilot test and following expert judgment. For each of the ten pictures, four other different pictures were given alongside each one of them. In only one of the four pictures given did its initial sound match the initial sound of the given target picture. The pupils' task was first to give a proper name to the target picture and after that to identify the one correct picture from the four, after giving them their proper names, whose initial sound matched the initial sound of the target picture (Schatskneider, Fletcher, Francis, Carlson & Foorman, 2004). The task involved the children working silently and finally to put the mark 'X' on the correct picture. For example, a list of four pictures of *mwezi (moon*), *nyoka (snake*), *ua (flower)* and *samaki (fish*) was presented, with a target picture of *nyota (star*). For Kiswahili, the correct response in this example is *nyoka,* which corresponds to the first sound in *nyota.* For English the correct response, of course, would be *snake,* which corresponds to the first sound in *star*. It should be noticed also that in this task the sounds were presented in mixed form, as explained in the picture-letter and letter-picture tasks for the purpose of creating distraction based on proactive and retroactive interference. This task had an added challenge in that it was demanding for the children, but the test administrators also used extra effort to make the children understand how to perform the task. The score was based on the total of correctly matched initial sounds. The minimum score was (0) and maximum score (10). The scores were converted into percentages. The scale was not timed. However, it was observed that some children completed the task in less than 12 minutes, while it took more than 25 minutes for others. The mean score of the initial sound subscale was 48 (SD = 33). The internal reliability was .91 (Cronbach's alpha).

Spelling: This subscale measures spelling and phonological awareness. The spelling consisted of twenty-five items, which varied in terms of phonological structure, word length and difficulty. All the words used had been drawn from the high frequency list, and were arrange in order of increasing difficulty. Examples of words include *au (or), maji (water), gwaride (parade), mwanafunzi (student)* and *ng'ombe (cow)*. The pupils were required to write down the spelling of each word as it was pronounced by the researcher (Invernizzi, Meier, & Juel 2005). The researcher pronounced a word three times before proceeding to another. The scale was scored on the basis of the total number of words spelled accurately. The minimum score was (0) and the maximum score (25). The scores were converted into percentages. The mean score of the spelling subscale was 47 (SD = 34). The internal reliability was .97 (Cronbach's alpha).

The Questionnaire

The questionnaire (Appendix 4) was constructed based on ideas from Basic Skills Assessment Tool (BASAT) developed by Ketonen & Mulenga (2003).

The questionnaire items were formulated to suit the context of this study. Parents and teachers completed the questionnaire on the pupils' home and school background information at school. Parents were asked to come to school to offer information on their child and to respond to the questionnaire. The first part of the questionnaire had ten questions that were intended to gather information about the child's status and demographic information. The first part of the questionnaire was mostly filled in by teachers based on information from school registers gathered from parents when the child was enrolled in the school. However, some information that was not in the registers (e.g., if the child had premature status, numbers of persons in the child's home and child's health status) was filled in by the parents. The second part of the questionnaire had twelve questions that aimed at providing an understanding of the child's home environment situation related to reading and writing skills acquisition. This was filled in by the parents. The third part of the questionnaire, filled in by the teachers, had thirteen questions aimed at understanding the child's school environment in connection with reading and writing skills acquisition. All of the items, N=25, (12 for the home environment and 13 for the school environment), were scored on a 3-point scale.

5.7.2 The individual test battery

The individual test battery (Appendix 5) includes measures of phonological awareness, measures of reading, comprehension and writing. Measures of phonological awareness include only one subscale, namely the pseudo-word Test. Measures of reading include three subscales, which are the reading words aloud, actual text reading, and one minute reading. The individual test battery also includes a writing test. The total is 6 subscales. All the subscales in the individual test battery were also selected based on the literature review and according to expert advice. The scales were also judged to be manageable when administering them to the 64 sampled pupils. The Individual test battery is given as Appendix 6. In the following part I present the schedule of administering the Individual test and a detailed description of the individual test subscales.

The schedule and administration of the individual test battery

The administration of the individual test battery was also done in four days from 15th August 2007 to 18th August 2007. As in the administration of the group test battery, two sessions were also used in each of the four sampled schools. The first one ran from 8.00am until 10.00am and the second one began at 10.30 and ended at 12.30 noon. For the two sessions 16 children were individually tested, that is 8 children were covered. Table 12 provides the schedule that was followed.

Table 12: Schedule for administering the individual test

Date	***School***	***Session1*** *08.00- 12.00*	***Session2*** *12.15- 04.15*	***Total***
15/08/2007	Kigurunyembe	8 tested	8 tested	16
16/08/2007	Kilakala	8 tested	8 tested	16
17/08/2007	SUA	8 tested	8 tested	16
18/08/2007	Mwande	8 tested	8 tested	16
Total		**32**	**32**	**64**

Description of the individual test battery

As in the group test, before the individual testing, each child was first given the relevant instructions that were required by the subscale. Concerning the reading tests, the children were first allowed to try out reading before they were ready to actually read for assessment. The 'reading words aloud' and the 'actual text reading' tests were untimed, but each child was allocated approximately ten minutes for each scale except for the 'one minute reading' test. However, the children were instructed to read as fast they could and to read all the words correctly. During the course of the reading the children were informed that they were allowed to make three trials to read a difficult word before it was judged. Two teachers and the researcher (i.e. examiners) co-listened to a child reading live. There was no tape recording of children's reading for fear of power cuts (which are very frequent) at the time of the study. Also when the recording was tried some children appeared to be shy to be recorded. Hence, each of the examiners registered in a note-book the words read correctly and incorrectly. An agreement was reached on consensus by the three examiners whether a child read a word correctly or incorrectly. The procedure of allowing children to make three trials, co-listening to a child reading and rating by consensus was followed in all the individual subscales.

For the writing test the children were instructed to write their name on paper provided. They were also instructed to write the 24 letters used in the Kiswahili alphabet, ten words of varying lengths and two simple sentences. The children wrote the items after they were pronounced by the researcher. The researcher repeated an item three times before proceeding to the next. The writing test was untimed, but the children were instructed to work as fast as they could, to write every item and to make sure that each letter and word was written correctly.

Measure of phonological awareness

Pseudo-Word: This subscale measures phonological awareness in terms of reading correctly a list of fifty nonsense words, which were constructed following the Kiswahili phonological structure, and which are decodable based on Kiswahili alphabetic coding. These words were chosen based on the results of qualitative items analysis during the pilot study and upon expert judgment. The words are arranged in order of increasing difficulty. Short, nonsense words with

only vowels (e.g., *ui, eo,)* or with an alternation of syllable and vowel (e.g., *memu*) were considered easy. Lengthy words (e.g., *wangechenjwa* and *mwatembedwe*), words with consonant clusters (e.g., *kinyame*and *sidhambe),* and words with king'ong'o (nasal) (e.g., *king'odo)* were considered to be difficult. The pupils were required to read aloud the non-words from the list. They were instructed to read loudly, correctly and as fast as they could. The scale was scored based on the total number of non-words decoded accurately. The scores were converted into a percentage. The minimum score was (0) and the maximum score (50). The scores were converted into percentages. The mean score of the pseudo-word subscale was 49 (SD = 39) and the internal reliability was .97 (Cronbach's alpha).

Measures of reading

Reading Words Aloud: This subscale measures word recognition in terms of reading correctly 30 given words arranged in order of increasing difficulty. Short words with only vowel combinations (e.g., *oa, ua*), single consonant and two vowels (e.g., *saa, taa*) or consonant and vowel (e.g., *bubu,nuna*) were considered easy. Lengthy words, words with consonant clusters and words with king'ong'o (nasal) plus consonant (e.g., *mwanafunzi, ng'ombe*) were considered difficult. However, all the words are drawn from the list of high frequency words obtained through a word count of three books used in grade one. The pupils were instructed to read the words loudly, correctly and as fast as they could. The score was based on the number of correctly read words as a proportion of the total words in the text and converting the score into a percentage. The minimum score was (0) and the maximum score (30). The scores were converted into percentages. The mean score of the reading words aloud test was 50 (SD = 32) and the internal reliability was .97 (Cronbach's alpha).

Actual Text Reading: This subscale involved reading an actual text. The test measured word identification and fluency in terms of the number of words read correctly against the total words in the text. The scale was not timed but the children were instructed to read as fast as they could without making mistakes. The actual text used was taken from a grade one class reader and had a total of 52 words. The raw score consists of the number of words read correctly as a proportion of the total words in the text. The minimum score was (0) and the maximum was (52). The scores were converted into percentages. The mean score of the actual text reading subscale was 50 (SD = 30) and the internal reliability was .97 (Cronbach's alpha).

One-Minute Reading: This subscale measures word identification in terms of rapid identification and naming of words in a text. The child's task was to read as many words as possible in one minute (Brus & Voeten, 1973). Three trials for familiarization purpose (i.e. on a difficult word for the child) were allowed to the children before they could actually be assessed. The text read was a part selected from a grade one class reader, and had a total of 76 words. The raw scores consist of the total number of list items (words) read correctly as a proportion of

the total list items (words in the text) in one minute. The score was then converted into a percentage. The minimum score was (0) and the maximum was (76). The scores were converted into percentages. The mean score of the one-minute reading subscale was 35(SD = 19) and the internal reliability was .98 (Cronbach's alpha).

Writing: This subscale measures writing ability in terms of ability to write own name, all the 24 letters of the Kiswahili alphabet, given words and simple sentences. Each child was given a prepared piece of paper on which they were asked to write their first name. Each child was then asked to write all the 24 letters of the Kiswahili alphabet (a, b, ch, d, e, f, g, h, i, j, k, l, m, n, o, p, r, s, t, u, v, w, y, z) after listening to the examiner pronouncing them randomly. The next task for the child was to write three prepared words in order of increasing difficult after they were pronounced. The words used were familiar to the child and taken from the list of high frequency words. Each child was also required to write a two word sentence and a three word sentence after they were pronounced. The test comprised 30 items in all. The minimum score was (0) and the maximum was (30). The scores were converted into percentages. The mean score of the writing subscale was 48(SD = 36) and the internal reliability was .97 (Cronbach's alpha).

5.7.3 The follow-up study

The follow-up study was initiated in mid-July, 2008. This was the *fourth phase* of the whole study. Four new group subscales were created. (see Appendix 6). They were created with a similar procedure as the original tests administered as group test in the second phase. The new scales for the follow-up study included spelling, picture-word, initial sound and word-chain. These scales were selected based on expert judgment and on how the scales performed when administered in grade one. The schedule of administering the follow-up test battery and a detailed description of the follow up subscales are presented in this part.

The follow-up test was administered in mid-July, 2008, when the pupils were half-way through grade two. Some pupils who had taken the group test during the second phase of the study, however, could not take the test because they had dropped out of school. But some new students in grade two who had not taken the group test during the second phase of the study took the follow-up test. These were pupils who had either been transferred to the schools from other schools in the country or from the Morogoro municipality and those who had repeated grade two. Some children who had taken the group test during the second phase of the study but now were repeaters in grade one also took the follow-up test. Table 13 provides the schedule for administering the follow-up test battery.

Table 13: Schedule for administering the follow-up test

Date	***School***	***Session1*** *08.00- 10.00*	***Session2*** *10.30- 12.30*	***Total***
18/07/2008	Kigurunyembe	39 tested	37 tested	**76**
19/07/2008	Kilakala	40 tested	39 tested	**79**
19/07/2008	SUA	53 tested	54 tested	**107**
20/07/2008	Mwande	33 tested	35 tested	**68**
Total		**165**	**165**	**330**

Description of the follow-up test battery

The follow-up test battery included measures of phonological awareness (spelling and initial sound); word recognition (picture-word and word-chain). These tests were selected based on expert judgment and performance when administered in grade one. The follow-up test battery is presented as Appendix 6.

Spelling: This subscale measured spelling and phonological awareness. The spelling subscale consisted of 25 items, which varied in terms of phonological structure, word length and difficulty. However, all the items were sampled from the bank of words listed as high frequency; following the word count of three commonly used Kiswahili books for grade one. Examples of items include *kuni (firewood), gauni (dress), vichwa (heads), mwanahabari (journalist) and nyang'au (fierce lion).* The pupils were required to write down the spelling of each word as it was pronounced by the researcher (Invernizzi, Meier & Juel, 2005). The researcher pronounced a word three times before proceeding on to another. The test was scored on the basis of the total number of words spelled accurately. The minimum score was (0) and the maximum was (25). The scores were then converted into percentages. The mean score of the spelling subscale was 54 (SD = 23) and the internal reliability was .94 (Cronbach's alpha).

Picture-Word: This subscale measured word recognition in terms of identifying one correct word among four words presented corresponding to a given picture. The pictures used were familiar to the children. The scale was used as a variation of the picture-word interference task introduced by Rosinski et al., (1975) and following expert advice. The child was required to identify ten words by drawing an 'X' mark on the correct word for each of the ten pictures given. The words were given in lower case. The minimum score was (0) and the maximum was (10). The scores were the converted into percentages. The scale was not timed. However, it was observed that some children were able to complete the task in 5 to 10 minutes while it took 20 minutes or a little more for others. The mean score of the picture-word subscale was 57 (SD=25) and the internal reliability was .95 (Cronbach's alpha).

Word-Chain: The word-chain subscale consisted of 25 word chains with a total of 60 words to recognize. The scale was intended to measure word recognition. All the words used for this scale were listed high frequency ones. They were obtained after the word count of three common Kiswahili books used in grade one. They were selected on the basis of item analysis during the pilot study and following expert judgment. The words were given in lower case. The chains were constructed and listed in order of increasing difficulty. Examples of the chain words are *oga/maji (bath/water), viatu/ndoo (shoes/bucket),* and *mswaki/pwani/bweni (toothbrush/coast/dormitory).*The scale demands reading of word chains of two to three words without space between them. The pupil's task was to recognize and separate the words by making dividing lines between them. It was untimed. The scale was scored based on the number of correctly recognized words. The minimum score was (0) and the maximum was (60). The scores were then converted into percentages. The mean score of the word-chain subscale was 48 (SD=23) and the internal reliability was .84 (Cronbach's alpha).

Initial Sound: This subscale measures phonological awareness (Schatskneider et al., 2004). The initial sound scale comprised ten picture items that were familiar to the children and used in their class readers. The pictures were also selected based on an items analysis during the pilot test and following expert judgment. For each of the ten pictures, four other different pictures were given alongside each one of them. Only one among the four pictures matched the initial sound of the given target picture. The pupils' task was first to give a proper name to the target picture. Then they had to identify the one correct picture from the four which matched the initial sound of the target picture; for example, a list of four pictures of *meza (table*), *mdomo (mouth*), *sahani (plate)* and *chungwa (orange*) was presented with a target picture of *saba (seven*). For Kiswahili, the correct response in this example is *sahani* which corresponds to the first sound in *saba.* For English, of course, there is no response that corresponds to *seven*. In this scale, however, all the target words had correct corresponding responses. The score was based on the total of correctly matched initial sounds. The minimum score was (0) and the maximum was (10). The scores were then converted into percentages. The scale was not timed. However, it was observed that some children completed the task in less than 12 minutes while it took more than 25 minutes for others. The mean score of the initial sound subscale was 46 (SD = 25) and the internal reliability was .86 (Cronbach's alpha).

5.8 Data analysis

All statistical analyses, with the exception of analyses for content validity, were conducted by employing SPSS program. The analyses were based on determining the performance of children in the group test that was the screen measure, and on the other three measures used for validation, i.e. school marks, individual test and follow-up test. The analyses were also targeted to determine the validity and reliability of the group test as a screen measure. In this section I

present the issues of reliability and validity and how they were analyzed in relation to the demands of this study.

In determining the performance of the children on the various measures used in this study, descriptive analyses were conducted. The means, standard deviations, maximum and minimum scores, medians and skewness were used to explore the children's performance and to explore further if there was enough amount of variance in the data to warrant further analysis. Further analyses related to finding the technical adequacy of the group test as a screening instrument for identifying children at risk of reading and writing difficulties. This involved finding the validity and reliability of the group test. Four types of validity were relevant and addressed in this study, namely content validity, construct validity, criterion validity and predictive validity (see Crocker & Algina, 1986; Cronbach, 1971; Messick, 1989; Spache, 1976; Springer et al., 2002).

Validity refers to the consistency with which a set of test scores measures what it is intended to measure. Content validity was assessed through a non-statistical process. This involved determining how the test actually represents the phenomena to be measured (i.e. the systematic examination of the test content to determine whether it covers a representative sample of the behavior domain to be measured). A test has content validity built into it by careful selection of which items to include. Items are chosen so that they comply with the test specification which is drawn up through a thorough examination of the subject domain (Anastasi, 1982; Hudson, 1982; Nunally & Bernstein, 1994; Salvia & Ysseldyke, 2001; Springer et al., 2002; Anastasi & Urbina, 1997, p.114).

Content validity was also ensured by employing a panel of experts to scrutinize the test items. It has been suggested that by using a panel of experts to review the test specifications and selection of items, the content validity of a test can be improved. The experts will be able to review the test items and comment on whether the items cover a representative sample of the behavior (Foxcraft, Patterson, Le Roux & Herbs (2004, p. 49).

Construct validity, which is most closely tied to theoretical considerations and concerned with the theoretical relationship of a variable to other variables, was assessed in two ways in this study: Firstly, by conducting an item analysis in order to establish the internal consistency for both the group test and the follow-up test; and secondly, by conducting a factor analysis in order to find the structure of the group test (see Crocker & Algina, 1986; Salvia & Ysseldyke, 2001; Springer et al., 2002).

Item analysis evaluates the respondents' responses to individual test items for the purpose of assessing the quality of the items and the test as a whole. Item analysis is important and useful in improving items to be used again in future tests. It can also be used to identify and eliminate too easy or too difficult items and also ambiguous or misleading items. In addition, item analysis is useful for increasing skills in test construction and identifying specific areas of content which need greater emphasis and clarity. The quality of the test as a whole is

assessed by estimating its internal consistency. The quality of individual items, on the other hand, is assessed by comparing respondents' item responses with their total test scores (see Anastasi, 1982; ScorePak, 2005; Statsoft, 2008).

Items can be analyzed qualitatively in terms of the items' content and form, which includes the consideration of content validity and the evaluation in terms of effective item writing procedures. Items can also be analyzed quantitatively in terms of their statistical properties, which include principally the measurement of item difficulty and item discrimination. In this study, both qualitative and quantitative item analyzes were employed, since both the validity and reliability of any test depend ultimately on the characteristics of its items (see Anastasi, 1982).

Qualitatively, the items were analyzed based on the establishing of content validity by choosing items so that they complied with the test specification, which was drawn through examination of the subject domain. The items used had a source from the grade one curriculum and were based on three commonly used class readers in Kiswahili orthography. After the construction of the tests, they were tried out in small groups and refined and used in the pilot study. The items were further refined by expert judgment. Necessary improvements were made before the items were used in the actual individual and follow-up studies. The detailed description of the process of qualitative item analysis and the improvements made is presented in the methods part, procedure section, of this study.

Quantitatively, the analysis was conducted by computing an item difficulty index (*p-value*). The item discrimination index (D) was found through computing total-item correlation (r *pbis*). Item and total raw scores were used in the computation of item difficulty, item discrimination and item-total correlation. The quantitative item analysis was an important back-up of the qualitative item analysis as observed in the content validation process

Item difficulty is calculated as the proportion of percentage of individuals choosing the right answer (Anastasi, 1982; Hotiu, 2006). In this case it is also equal to the item mean. The easier the item, the larger will this percentage be. This statistic is known as item difficulty or facility, and termed *p-value.* However, one shortcoming with item difficulty or facility, as in all percentages, is that the scale implied cannot be regarded as calibrated in equal intervals. The difference in intensity of difficulty between items with facilities of 0.40 and 0.50 is not the same; rather, it is somewhat smaller than the difference between items with facilities of 0.10 and 0.20. This elasticity in the scale makes for misjudgment in comparing facilities and rules out simple statistical manipulations of *p-value*, except for summation, which results in an estimate of the total test mean.

Item discrimination, on the other hand, is calculated as the difference between the percentage of respondents in the top one third of the class who correctly

answered that test item, and the percentage in the bottom one third who correctly answered the item (Anastasi, 1982; Hotiu, 2006). This index should be viewed in the context of the overall difficulty of the question. For example, the index of discrimination (*D-value*) might not be useful for a question that only 10 percent of the overall class answered correctly. In most circumstances, however, viewing the correct percentage in each one third of the class is useful in and of itself.

Generally, it has been found and accepted that items with an item difficulty index (*p-value*) between 0.15 and 0.85 are considered good items to be included in a test (Wagner, Torgesen & Rashotte, 1999). Items with item discrimination index (*D-value*) of 0.30 and above are usually considered acceptable to be included in a test (Aiken & Groth-Marnat, 2006). Hence, items for both the group test and the follow-up test were analyzed based on these guidelines. Generally, all the subscales had good enough items to be chosen, except for the word-chain, which had the highest number of items not meeting the criteria: it had five items, 2 very easy for most children and 3 too difficult. The next was the spelling scale, which had 3 items not meeting the criteria: it had 1 item too easy and 2 items too difficult. The picture-word and the word-picture subscales each had only one item that was too easy.

For the context of the study the sample size (N=337) was adequate or good enough to conduct factor analysis, which is about data reduction for the purpose of assessing data suitability and for determining how many underlying factors there are in a set of variables. In order to assess the dimensionality of the screening tool in this study, principal component analysis was used instead of confirmatory factor analysis. Principal component analyses are traditional 'exploratory' factor analyses which analyze all the variance in the items. Data is reduced for simplification by summarizing the variance associated with many items down to just a few factors. (Denton, Ciancio & Fletcher, 2006).

Criterion-related validity refers to the extent of correlation between the test and the criterion. In order to assess criterion validity, descriptive analyses, and correlations of the group, individual and follow up subtests and correlations of the screen with criterion measures were used. Internal consistency was determined by Cronbach's alpha. Predictive validity was sought in two ways. First, by examining the extent to which the children's results from the group test taken in mid-July, 2007, in grade one correlated with the results in the follow-up test taken a year later, in grade two. Secondly, evidence of predictive validity was established based on cluster analysis. This was done by cross-tabulating grade one and grade two clusters of children and analyzing how many of the at-risk children in grade two were already found to be at risk with the screening instrument in grade one.

K-means cluster analysis was used for classification and finding the children at risk of getting difficulties. The analysis was also used in determining the cut-off point on the summary score in the group test for finding the at risk group. Cross tabulation of the group test clusters with school status, gender and age was used

to find out the distributions in the clusters. The analyses also included multivariate GLM procedures MANOVA in determining the effects of school status, gender and age on the test results; and also for determining the effects of home and school factors.

Ethical issues

Ethical issues were paid attention to in this study for ensuring credibility and authenticity. A number of issues were taken into consideration as the time for data collection approached. For salient validity the author visited possible areas for the study to create awareness and obtain consent of the schools to conduct research there. The parents were notified by the teachers and were informed of the intention of the study. They were asked if they could give their time and voluntarily respond to a questionnaire. They gave their consent and the children were notified of the tests beforehand; hence, they were prepared. Formal letters from the Ministry of Education and Vocational Training and from Morogoro municipality were obtained by way of seeking approval, permission and clearance to conduct the research. The Inspectorate and the District Education Officer's office were also used for information about the performance of the schools in the municipality. It is hoped that when the research is complete, feedback will be given to all involved.

6 Results

The aim of this study was to create and validate a group-based screening tool, in Kiswahili, for identifying Tanzanian first graders at risk of developing reading and writing difficulties. The results are presented in this chapter, which is divided into four sections.

The first section includes a presentation of the psychometric properties of the screen. The children's performances in the group test, individual test and follow-up test and their school marks are presented by descriptive statistical data. The validity of the screen is analysed. Content validity was sought by using items based on the content of the schoolbooks which follow the curriculum in grade one. Construct validity was determined by item and factor analyses. Criterion validity was sought through correlating the group test with the criterion measures; and predictive validity through the correlations and cluster analyses, using grade one and grade two data. The reliability of the group test is also presented in the first section.

The second section deals with a presentation of the results involving the classification and identification of children at risk through K-Means cluster analysis. This section also deals with assessing the stability of the classification. Another aspect presented in this section involves the setting of the cutoff point determined by the classification results with the K-means cluster analysis. The section ends with a presentation of the distribution of children in clusters in relation to gender, age, school status and school attendance based on cross tabulation.

In the third section the effects of gender, age and home and school factors are analyzed by multivariate GLM procedures. The fourth and last section is concerned with the explanatory power of environmental determinants. The influence of some selected home and school factors on the group test performance are presented. In the end a general summary of this chapter is offered.

6.1 Psychometric properties of the group test

The psychometric properties of the created group test for identifying children at risk of reading and writing difficulties are presented in this section. The descriptive statistics for the group test and for the tests which were used for validation, i.e. individual test, follow-up test and school marks are presented first.

6.1.1 Descriptive data of the scales of the group test, individual test, follow-up test and school marks

The children's (N=337), summarized performance results in the four batteries (group test, individual test, follow-up test and school marks) are presented in Table 14. The maximum and minimum scores, means, medians, standard

deviations, reliability coefficients and the skewness of all the scales are presented.

Minor negative skewed distributions were apparent. The subscales of the group test had a skewness ranging from -.002 (initial sound) to -.410 (word-picture). The subscales of the individual test ranged from -.009 (writing) to -.942 (one minute reading). The school marks had -.059 and the subscales of the Follow-up test ranged from -.244 (follow-up initial sound) to -.656 (follow-up picture-word). These skewness indices, however, are within the acceptable limit range of +2 to -2. Skewness can be expected when the tasks are designed for the purpose of screening to identify children with difficulty in reading and writing. The low scores are the main point of interest.

The descriptive data in this study provide enough information to discriminate among the children. The scores are in percentages: the 0 to 100 range of the scores for almost all of the subscales, and the 0 to 88 range for the follow-up word chain suggest a good range of variation. The group test means for the 7 subscales range from 47 to 60 (SD=33 to 35); the individual test, 5 subscales, means range from 35 to 50 (SD=19 to 39); the follow-up test, 4 subscales, means range from 46 to 57 (SD=23 to 25) and school marks mean, 51 (SD=29). The amount of variance was therefore clearly sufficient to justify further analyses of the underlying structural relations. Principal component analysis was used for structural analysis, since it is robust regarding deviations from normality (Engen & Høien, 2002; Høien, Lundberg, Stanovich & Bjaalid, 1995).

The descriptive results indicate that there are some children, in mid-grade one, and in mid-grade two, who have not yet mastered the simple tasks used when assessing word and letter knowledge, phonological awareness and spelling. This is revealed by the minimum score (0) in all the variables. The results also demonstrate that there are some children in mid-grade one who still find it difficult to read pseudo words and real words (with simple vowel and consonant arrangement or those of high frequency). At this point in time, there are children who cannot read an actual text based on their common class reader. Some children during this time are also struggling to read with fluency. Further still, it is evident that some children cannot write their names, the letters of the alphabet, simple words and simple sentences as measured by the spelling and writing tasks.

Table 14: Descriptives: group test, individual test, school marks, and follow-up test. All scores are in %. All subscales: Max=100, Min=0; except Follow-up Word-Chain: Max=88; Min=0.

Gtest	Mean	Med	SD	α	Skew	Foltest	Mean	Med	SD	α	Skew
PiWo	60	70	34	.92	-.392	FoPw	57	70	25	.95	-.656
WoPi	59	60	35	.93	-.410						
PiLe	57	67	34	.94	-.412						
LePi	56	67	34	.94	-.382						
InSo	48	50	33	.91	-.002	FinSo	46	60	25	.86	-.244
WoCh	50	52	33	.97	-.025	FoWch	48	68	23	.84	-.522
Spell	47	44	34	.97	-.213	FoSp	54	72	23	.94	-.497
Smks	51	52	29	.99	-.059						
Indtest	**Mean**	**Med**	**SD**	**α**	**Skew**						
Pseudo	49	50	39	.97	-.032						
Aloud	50	50	32	.97	-.308						
Text	50	63	30	.97	-.470						
Minute	35	42.5	19	.98	-.942						
Write	48	50	36	.97	-.009						

Note: Gtest=Group test; Smks=School marks; Indtest=Individual test; Foltest=Follow-up test; PiWo=Picture-Word; WoPi=Word-Picture; PiLe=Picture-Letter; LePi=Letter-Picture; InSo=Initial Sound; WoCh=Word-Chain; Spell=Spelling; Pseudo=Pseudo-Word; Aloud=Reading Words Aloud; Text=Actual Text Reading; Minute=One Minute Reading; Write=Writing; FoPw=Follow-up Picture-Word; FinSo=Follow-up Initial Sound; FoWch=Follow-up Word-Chain; FoSp=Follow-up Spelling.

6.1.2 Validity

Content validity

This section deals with the description of how content validity was ensured in this study. The content validity of the instruments created for measuring reading and writing ability was firstly ensured by using regular schoolbooks as basis in the process. The letters and words that formed the test items for the group, individual and follow-up tests were those used in common Kiswahili books used by and familiar to grade one children at school. This was verified by doing a manual word count of three common Kiswahili books by Kihampa (1997), Mkinga (2000) and another by the Tanzania Institute of Education (TIE, 2000) in order to identify a list of the most frequently seen and used words. The list was further checked if it was included in the grade one syllabus.

Secondly, the tests were scrutinized and evaluated by a panel of experts, including two competent and experienced grade one teachers and a Kiswahili expert from a teacher training college for (a) appropriateness of content and the reading texts, (b) familiarity of the words and pictures to the children's context, (c) appropriateness of language for the comprehension and writing scales, (d) distinction of easy and difficult words, and (e) clarity of instructions. The panel was given the objectives of the tests, a description of the intended sample and the test items for the task. Most of the items were judged to be appropriate, with

very minor changes and modification especially where some pictures were assessed as being unfamiliar to the children.

Construct validity

In order to ensure construct validity of the instruments, the following steps were taken. Firstly, an item analysis was performed in order to ensure internal consistency. Secondly, the dimensionality was assessed through conducting a principal component factor analysis. In the following section, a presentation of the results of the item analysis are given first, and then followed by the results of factor analysis.

Item analysis

The process of item analysis involved determining the reliability, item difficulty and item discrimination for the 7 subscales of the group test, which had a total of 104 items, and of the follow-up test with 4 subscales, including 70 items. The scales and the number of items are presented in Table 15.

Table 15: Group test and follow-up test subscales; number of items

Group test (N=337)		**Follow-up test (N=330)**	
Subscale	**Items**	**Subscale**	**Items**
PiWo	10	FoPw	10
WoPi	10		
PiLe	12		
LePi	12		
Inso	10	FinSo	10
WoCh	25	FoWch	25
Spell	25	FoSp	25
Total	**104**		**70**

Note: PiWo=Picture-Word; WoPi=Word-Picture; PiLe=Picture-Letter; LePi=Letter-Picture; Inso=Initial Sound; WoCh=Word-Chain; Spell=Spelling; FoPw=Follow-up Picture-Word; FinSo=Follow-up Initial Sound; FoWch=Follow-up Word-Chain; FoSp=Follow-up Spelling.

A detailed description of the different items, including difficulty and discrimination indexes for the group test and follow-up test is given in Appendices 7 and 8.

A summary of the item analysis for Group test items is presented in Table 16. The group test has high reliability, with its items having alphas ranging from r=.91 to .97. The p-values for most of the items lie within acceptable levels of difficulty (p=.15 to .85). The exceptions are the easiest items of picture word, word picture, word chain subscales (nr 1), and of word chain (nr 2); all of them ranging from .86 - .91. Six difficult items were noticed: nr 12 of letter picture, nr 23-25 of word chain, and nr 24 and 25 of spelling, all ranging from p=.07 to .12. The subscales of picture letter and initial sound had neither easy nor difficult items. The D-values for all the items in all the subscales lie within acceptable ranges of good discrimination.

Table 16: Summary of the results of the item analysis for the group test

Subscale	Items	Mean	SD	Alpha	p-value	D-value
PiWo	10	6.0	3.4	.92	.21 – .90	.51 – .82
WoPi	10	5.9	3.5	.93	.26 – .88	.55 – .83
PiLe	12	6.7	4.1	.93	.15 – .85	.42 – .85
LePi	12	6.7	4.0	.93	.10 – .85	.38 – .87
Inso	10	4.8	3.3	.91	.27 – .85	.53 – .81
WoCh	25	12.4	8.3	.97	.07 – .91	.39 – .87
Spell	25	11.7	8.6	.97	.10 – .88	.47 – .86

Note: Mean refers to number of items correct.
PiWo=Picture-Word; WoPi=Word-Picture; PiLe=Picture-Letter; LePi=Letter-Picture; Inso=Initial Sound; WoCh=Word-Chain; Spell=Spelling

Table 17 summarizes the item analysis results for the follow-up test. The reliability is high, with alphas ranging from .84 to .95. Some easy items were identified in spelling (nr 1, 2, and 3, p=0.95), picture word (nr 1, p=0.95, nr 2, p=0.91 and nr 3, p=0.88), initial sound and word chain (nr 1 and 2, p=0.95 and 0.88). Some rather difficult items were found in spelling (nr 21-24, p=0.08-0.2), picture word (nr 9-10, p=0.10-0.02), initial sound (nr 8-10, p=0.09-0.01) and word chain (nr 8-10, p=0.09-0.01).

Most of the items discriminated well. However, each subscale, except for spelling, had items which did not meet the criteria for discrimination: picture word nr 10 (D=.20), initial sound nr 9 (D=0.21), and 10 (D=0.19), word chain nr 22 (D=0.21) and 23-25 (D=0.11).

Table 17: Summary of the item analysis results for the Follow-up test

Subscale	Items	Mean	SD	Alpha	p-value	D-value
FoSp	25	13.4	6.3	.95	.02 – .95	.43 – .83
FoPw	10	5.7	2.5	.86	.02 – .95	.20 – .78
FinSo	10	4.6	2.3	.84	.01 – .95	.19 – .75
FoWch	25	12.1	5.8	.94	.00 – .94	.14 – .81

Note: Mean refers to number of items correct.
FoPw=Follow-up Picture-Word; FinSo=Follow-up Initial Sound; FoWch=Follow-up Word-Chain; FoSp=Follow-up Spelling.

Dimensionality of the group test

The previous part was concerned with presenting the results of item analysis necessary for establishing the construct validity of the screen. This part deals with the results of factor analysis. The dimensionality was checked by principal component analysis (PCA). Prior to performing PCA, the suitability of the data for factor analysis was assessed. The correlation matrix (Table 18) revealed high correlation coefficients. The Kaiser-Meyer-Oklin value was .94, exceeding the

recommended value of .6 (Kaiser, 1974), and the Bartlett's Test of Sphericity (Bartlett, 1954) reached statistical significance (p=.000). All correlations are highly significant.

Table 18: Correlations between the group test subscales

	Subscale	1	2	3	4	5	6	7
1	PiWo		.95	.93	.92	.91	.88	.89
2	WoPi			.93	.93	.92	.90	.90
3	PiLe				.96	.91	.90	.90
4	LePi					.92	.91	.91
5	Inso						.93	.91
6	WoCh							.92
7	Spell							

Correlation is significant at the 0.01 level (2-tailed)
Note: PiWo=Picture-Word; WoPi=Word-Picture; PiLe=Picture-Letter; LePi=Letter-Picture; Inso=Initial Sound; WoCh=Word-Chain; Spell=Spelling

The results of principal component analysis are presented in tables 19 and 20. Based upon eigenvalue greater than 1, a very powerful (one-factor) 'latent' dimension emerged. This factor explained 93% of the variance. The factor was named *initial literacy factor*. The high correlations and the factor structure indicate that a summary score can be counted and used in further analysis. The results also indicate that any subset of the scales in the group test could be used to form a reliable test for screening. Hence, a shorter version of the group test was created and is presented as Appendices 9 and 10.

Table 19: Factors, eigenvalues, and percentage of variance accounted for

Factors	eigenvalue	% Variance	Cumulative % Variance
1st	6.70	92.81	92.81
2nd	.17	2.38	95.19
3rd	.10	1.44	96.62
4th	.08	1.17	97.79
5th	.06	.92	98.71
6th	.05	.75	99.46
7th	.03	.54	100.000

Table 20: Factor loadings in Principal Component Analysis

Subscale	Factor 1
PiWo	.961
WoPi	.970
PiLe	.969
LePi	.971
Inso	.964
WoCh	.955
Spell	.953

All factor loading > 0.50
Note: PiWo=Picture-Word; WoPi=Word-Picture; PiLe=Picture-Letter; LePi=Letter-Picture; Inso=Initial Sound; WoCh=Word-Chain; Spell=Spelling

Criterion validity evidence

Criterion-related validity refers to the extent of correlation between the test and the criterion. Concurrent validity refers to the data indicating the similarities and differences of a new test from other known instruments (Springer, et al., 2002; Salvia & Ysseldyke, 2001; Crocker & Algina, 1986). In the following section, the correlations between the group test, and the three measures used as criterion for validity, i.e the individual test, school marks and follow-up test are presented.

As there are no standardized screening tests for identifying first graders at risk of reading and writing difficulties in the Kiswahili language, criterion validity was established using the individual test, the school marks, and the follow-up test as criterion measures. The correlation coefficients between the seven subscales of the group test and the measures in the three criterion tests are presented in Table 21 and Table 22. The intercorrelations of the scales in the individual test are high and significant. All the scales in the group test correlate significantly with the scales in the individual test, with coefficients ranging from r = .81 to .96 and with the school marks ranging from .76 to .96. The correlation between the scales in the individual test and the school marks are also significant, varying from .78 to .95.

Table 21: Correlations between group test, individual test and school marks

		Group test							Individual test					Smks
		1	2	3	4	5	6	7	8	9	10	11	12	
1	PiWo		.95	.93	.92	.91	.88	.89	.90	.89	.88	.84	.91	.83
2	WoPi			.93	.93	.92	.90	.90	.92	.91	.91	.88	.92	.85
3	PiLe				.96	.91	.90	.90	.91	.89	.92	.85	.95	.83
4	LePi					.92	.90	.91	.90	.88	.92	.87	.94	.83
5	Inso						.93	.91	.92	.90	.90	.84	.95	.85
6	WoCh							.92	.88	.88	.87	.81	.91	.80
7	Spell								.90	.87	.86	.80	.92	.86
8	Pseudo									.93	.90	.83	.94	.85
9	Aloud										.94	.90	.91	.82
10	Text											.93	.91	.84
11	Minute												.85	.78
12	Write													.91

Correlation is significant at the 0.01 level (2-tailed
Note: PiWo=Picture-Word; WoPi=Word-Picture; PiLe=Picture-Letter; LePi=Letter-Picture; Inso=Initial Sound; WoCh=Word-Chain; Spell=Spelling; Pseudo=Pseudo-Word; Aloud=Reading Words Aloud; Text=Actual Text Reading; Minute=One Minute Text Reading; Write=Writing; Smks=School marks.

The intercorrelations of all the subscales in the follow-up test are high and significant. The scales of the group test and the follow-up test correlated also significantly (.83 - .96). The relationship between the school marks (grade one) and the follow-up results were slightly weaker, although significant.

Table 22: Correlation between the group test, follow-up test and school marks

	Group test								Follow-up test				Smks
		1	2	3	4	5	6	7	8	9	10	11	
1	PiWo		.95	.93	.92	.91	.88	.89	.91	.88	.88	.89	.83
2	WoPi			.93	.93	.92	.90	.90	.91	.89	.89	.89	.85
3	PiLe				.96	.91	.90	.90	.90	.88	.87	.88	.83
4	LePi					.92	.91	.91	.88	.86	.85	.86	.83
5	Inso						.93	.91	.88	.86	.87	.88	.85
6	WoCh							.92	.85	.83	.86	.87	.80
7	Spell								.87	.85	.86	.87	.86
8	FoSp									.96	.93	.93	.82
9	FoPw										.93	.93	.79
10	FinSo											.96	.76
11	FoWch												.77

Correlation is significant at the 0.01 level (2-tailed).
Note: Gtest = Group test; Foltest = Follow-up test; PiWo=Picture-Word; WoPi=Word-Picture; PiLe=Picture-Letter; LePi=Letter-Picture; Inso=Initial Sound; WoCh=Word-Chain; Spell=Spelling; Fpw=Follow-up Picture-Word; FinSo=Follow-up Initial Sound; FoWch=Follow-up Word-Chain; FoSp=Follow-up Spelling; Smks=School marks.

In summary, it can be concluded that the scales of the group test correlated significantly ($p<0.01$) with the three criterion measures (the individual test, school marks and follow-up test). Based on the results, it can be concluded that the criterion validity of the created group test is good.

Predictive validity evidence

Predictive validity refers to the degree to which scale scores predict criterion measures that will be made at some point in the future and it requires that participants are followed during a period of time (Springer et al., 2002). In this study predictive validity of the group test was determined by examining the extent to which children's results from the group test in mid-July 2007, when the children were in grade one, correlated with the results in the follow-up test one year later, when they were in grade two. The results already presented in Table 22 revealed a strong relationship between the measures from the two school years. High correlations between the subscales in the group test and the follow-up were noticed.

Strong predictive validity is also evident based on the cluster analysis described in Chapter 6.2.1. By cross tabulating grade one and grade two clusters of children it was found that 94% of the at-risk children in grade two were already identified in the group test in grade one.

6.1.3 Reliability of the tests

In this study, achieving good reliability was a high priority; as the aim was to produce the first Tanzanian screen for finding children at risk of getting into reading and writing difficulties in grade one in Tanzanian schools. The previous chapters include descriptions of the strategies for achieving a reliable screen. The process included literature analysis, careful creation and piloting of the scales, expert judgment, and careful administration, marking and scoring of the tests. In addition, the scoring was double checked in Finland in a 10% random sample of the original test sheets. The reliability indicated by Cronbach's alpha was very high in all the scales in the subscales of the tests. The figures of the scales are given within the description of the group test, the individual test and the follow-up. The item analysis includes the specific figures. The reliability scores (Cronbach's alpha) were extremely high in all the tests (group test, 7 scales: .91 to .97; individual test, 5 scales: .97 to .98 and for the follow-up, 4 scales: .84 to .95), which also could be expected based on the high intercorrelations. The apparent explanation is the one-dimensionality of the scales and the careful process of creating and administrating the tests.

6.2 Classification and finding children at risk

Cluster analysis: A cluster analysis was conducted in order to identify groups of children sharing certain similar characteristics in relation to their reading ability, but being as different as possible from children in other groups. The results of the analysis are presented in this section. The interest was to see whether the children could be meaningfully grouped based on their scores in reading and writing ability in the group test. The analysis is useful in developing hypotheses concerning the nature of the data, for classifying data into manageable groups and for reducing population information to specific subgroups. As

recommended, the sample for this study (N=337) was large enough in exceeding the recommended N=250. Three methods were available for choice. First was the Wards method, which is hierarchical. In this analysis observations are required to remain together once they have joined in a cluster. This method is most appropriate for small samples, less than 250.

The K-means method algorithm, on the other hand, uses Euclidian distance. This method is nonhierarchical and appropriate for larger samples. In contrast to the Wards method, the K-means cluster analysis does not require observations to remain together once they have joined in a cluster. The researcher in the K-means cluster analysis has to pick a number of clusters to run. It was appropriate for this study. However, it was reasoned that two-step cluster analysis be used first, as it is suitable as an exploratory or so-called 'blind' analysis, and capable of handling very large samples. It is also capable of handling either continuous and categorical variables, or attributes, requiring only one data pass in the procedure. The two-step method was used to assist on deciding the number of clusters. The criterion used was Akaike's Information Criterion (AIC). It was based on deviance scores and a number of clusters is selected such that it predicts the data well but is not too large, in order to find a balance: detailed, but not too detailed. Hence, the four cluster solution was found to be meaningful and also manageable by the teachers in assisting children struggling to learn to read.

In the first step of the procedure, you pre-cluster the records into many small sub-clusters. In the second step, you cluster the sub-clusters from the pre-cluster step into the desired number of clusters. If the desired number of clusters is unknown, the SPSS Two Step Cluster component will find the proper number of clusters automatically (SPSS inc. 2001; Garson, 2009). In this study, cluster solutions ranging from three to six clusters were tested. The final four-cluster solution was preferred as it provided greatest contrast between the groups and was regarded as being meaningful and manageable when related to screening.

The 337 participants were classified into four clusters, with cluster sizes ranging from N=75 to N=99 children. The final cluster centers are displayed in Table 23. The table gives a schematic and simplified overview of the characteristics concerning the one factor that emerged from the factor analysis and of the proportion of each cluster.

Table 23: Cluster centres and cases (group test)

Subscale	Clusters				Total
	1	2	3	4	
PiWo	12	48	76	96	
WoPi	10	47	75	96	
PiLe	9	45	74	92	
LePi	10	43	72	92	
Inso	7	34	57	87	
WoCh	9	32	61	88	
Spell	6	28	53	90	
N	**82**	**75**	**81**	**99**	**337**
%	**24.3**	**22.3**	**24.0**	**29.4**	**100**

Note: Gtest = Group test; PiWo=Picture-Word; WoPi=Word-Picture; PiLe=Picture-Letter; LePi=Letter-Picture; Inso=Initial Sound; WoCh=Word-Chain; Spell=Spelling

Table (23) indicates that until the middle of grade one 157 children (46.6%) in clusters 1 and 2 are struggling, with those in cluster 1 (24.3%) being more at risk. The children in both clusters are still struggling to acquire the basic reading and writing skills. On the one hand, 180 children (53.4%), in clusters 3 and 4 can read and write, with 99 children (29.4%) in cluster 4 being good readers.

The clusters formed were named based on the observed characteristics in relation to the children's performance on the group test. This was done to facilitate the understanding of mathematically derived clusters. By using cluster names one can easily associate the clusters with their depicted analogue characteristics. Table 24 contains brief descriptions of the characteristics (behaviors) and names given to the clusters.

Table 24: Description of characteristics of children in clusters and cluster names

Cluster	Characteristics	Name
1	Very low scores in all subscales	*At-risk*
2	Moderate low scores in many subscales	*Strugglers*
3	Moderate high scores in many subscales	*Readers*
4	Very high scores in all subscales	*Good readers*

The four clusters described in Table 24 were retained and named as follows: 1=*At-risk*, 2=*Strugglers*, 3=*Readers* and 4=*Good readers*.

At-risk children (24%) had difficulties in all areas, with very poor children facing more problems. *Strugglers* (22%) had some difficulties with initial sound and spelling. They also had problems in letter-picture and more problems than the former group with word chain and spelling. *Readers* (24%) were generally rather good in decoding but scored lower in initial sound, word chain and

spelling compared to *good readers* (29%). The children in the good readers group had high scores in all the group subtests, indicating that they were able to decode without problems. Their phonological awareness was also good.

For comparing with the Group test (grade1) clustering, it was found justified to conduct a K-Means cluster analysis based on the follow-up test (grade 2). The results are presented in Table 25. The results indicates that there are 52 (15%) *at-risk* children (cluster 1), 77 (23.3%) *strugglers* (cluster 2), 118 (35.8%) *readers* (cluster 3), and 83 (25.1%) *good readers* (cluster 4). These results were cross-tabulated with those of the group test to determine the stability of the classification. It should be considered that in grade two only 330 children participated in the follow-up test. Seven children were 'lost' from grade one to grade two as a result of dropping out of school. All these seven children belonged to the *at-risk* group in grade one.

Table 25: Cluster groupings and cases in the clusters (follow-up test)

Subscale	**Clusters**				**Total**
	1	2	3	4	
FoSp	11	38	64	80	
FoPw	13	43	66	83	
FinSo	9	32	53	72	
FoWch	9	35	56	73	
N	**52**	**77**	**118**	**83**	**330**
%	**15.8**	**23.3**	**35.8**	**25.1**	**100**

Note: FoPw=Follow-up Picture-Word; FinSo=Follow-up Initial Sound; FoWch=Follow-up Word-Chain; FoSp=Follow-up Spelling.

6.2.1 Classification stability

In this part the results of the cross tabulation between the group test clusters and the follow-up test clusters are presented based on the K-means cluster analysis. These results, presented in Table 26, indicate that the majority of the 330 pupils were found in the same clusters in the follow-up as in grade one.

Table 26: Classification stability grade 1 and grade 2 (Group test vs. Follow-up test)

Clusters grade 1	Clusters grade 2									
	1		2		3		4		Total	
	N	%	N	%	N	%	N	%	N	%
1	49	65	26	35	0	0	0	0	75	100
2	3	4	45	60	27	36	0	0	75	100
3	0	0	6	7	64	79	11	14	81	100
4	0	0	0	0	27	27	72	73	99	100
Total	**52**	**15.8**	**77**	**23.3**	**118**	**35.8**	**83**	**25.2**	**330**	**100**

Of the 75 *at-risk* children (7 had dropped out) in grade one, 65% were still found in the *at-risk* group in grade two, while 35% had migrated one step to the *Strugglers* group. No children had moved to the two highest clusters. The interpretation of the cross tabulation pattern is similar for the children in all clusters. Most of the children (60-80%) in the grade one clusters are found in the same clusters in grade two. The migrations found are mostly to a higher cluster, which should be regarded as a positive move after one year in school. Three children from the *Strugglers*' group and six from the *Readers'* group were found in the next lower cluster in grade two. Of the *Good readers*, 27% had moved down to the *Readers*' cluster, a result which is naturally not desired, but not problematic from a risk perspective of becoming a poor reader. Apparently, no factors studied in this study could be based to explain this migration. Other factors could be speculated as reasons, e.g., sickness or lack of motivation at the time of taking the test. In summary, it can be said that the children's situation in relation to reading and writing risk problems remained rather stable from grade one to grade two.

The presented results reveal that some of the children in the lowest cluster managed to move to a higher cluster during the first school years, and this is naturally the desirable development. Most at-risk children identified by the screen are, however, in need of support.

It is relevant and of special interest to analyze how many of the children found in the *at-risk* group in grade two were identified by the screen in grade one to be at risk of developing problems. In grade two, there were 52 children in the *at-risk* group, and of these children 49 were identified as being at risk already in grade one. This means that 94% of the *at-risk* children in grade two were already identified in the group test in grade one. All in all, 16.8% of the children in grade two belonged to the *at-risk* group.

6.2.2 Determining the cut-off point for the four clusters

A summary variable was created and a summary score was counted based on the seven scales of the Group test (Mean=377.2, SD=227.48, Maximum score 700, Minimum score=0, Median=404 and Mode=0). The analysis of the distribution reveals that 24.3% (N=82) of the children, a proportion equivalent to the amount

of *At-risk* children discussed above, have scores below 166. The suggestion is that children scoring lower than 166 belong to the *At-risk* group and need intensive support. The 75 children (22.3%) scoring between 166 and 368 also belong to a group needing attention and observation. They can somehow manage, but they are struggling. With extra attention and some support most of them can probably manage well. Children scoring above 368 are doing fine. When using the test, however, attention should be paid to the fact that many background factors have an effect on the test results.

6.2.3 Distribution of children in clusters in relation to *gender, age*, *school status* and *school attendance*

Further analyses aimed to find how the children were distributed in clusters in relation to *gender, age, school status,* and *school attendance*. Cross tabulations are presented.

Slight *gender* differences can be seen in the cluster pattern (Table 27). The proportion of boys and girls was about the same in the risk cluster, while there were 10% more girls than boys in the *Good readers* cluster.

Table 27: Distribution in clusters by gender

Gender	1		2		3		4		Total	
	N	%	N	%	N	%	N	%	N	%
Girls	37	23.7	32	20.5	33	21.2	54	34.6	156	100
Boys	45	24.9	43	23.8	48	26.5	45	24.9	181	100
Total	**82**	**24.3**	**75**	**22.3**	**81**	**24.0**	**99**	**29.4**	**337**	**100**

Some age differences are evident (Table 28). About 40% of the older students (age group 9 to 12) compared to about 20% of the younger students are found in the *at-risk* and *strugglers* clusters (1 and 2). The pattern is naturally opposite in the *readers* and *good readers* clusters (3 and 4).

Table 28: Distribution in clusters by age

Age group	1		2		3		4		Total	
	N	%	N	%	N	%	N	%	N	%
9-12	11	42.3	10	38.5	4	15.4	1	3.8	26	100
6-8	71	22.8	65	20.9	77	24.8	98	31.5	311	100
Total	**82**	**24.3**	**75**	**22.3**	**81**	**24.0**	**99**	**29.4**	**337**	**100**

There is a clear difference in the cluster pattern between the schools (Table 29). The highest proportion of at risk children are found in the *low* and *periphery average performance* schools (37% and 35%), and the lowest in the *high performance* school (8%), whereas over 40% of the children in the *high performance* school are found in the *good readers* cluster.

Table 29: Distribution in clusters by school status

School Status	1		2		3		4		Total	
	N	%	N	%	N	%	N	%	N	%
LP	29	36.7	17	21.5	23	29.1	10	12.7	79	100
PA	25	35.2	19	26.8	9	12.7	18	25.4	71	100
GP	20	25.0	15	18.8	20	25.0	25	31.3	80	100
HP	8	7.5	24	22.4	29	27.1	46	43.0	107	100
Total	**82**	**24.3**	**75**	**22.3**	**81**	**24.0**	**99**	**29.4**	**337**	**100**

Note: LP=low performance; PA=periphery average; GP=good performance; HP=high performance;

An analysis of the age groups in relation to school status showed that 46% of the 26 children aged 9 to 12 came from the *low performance* school and 42% from the *good performance* school. No older children were found in the *peripheral average performance* school and three were in the *high performance* school. The children in the younger age group (6 to 8) were more evenly distributed in the schools.

In order to check if the older age groups affected the results and conclusions of the cluster analysis, an analysis was performed after omitting the children aged 9 to 12. There was still the same amount of children (24%) in the *at-risk* group and in the *strugglers* group (23%).

School attendance was considered to be a critical factor in the study. It was not unusual that children did not come to school every day. Long distances and poverty in the family can be mentioned among the reasons behind the low school attendance of many children. In the following (Table 30), a picture of the situation in relation to the risk perspective is given. There were 46 children who had been in school for less than one third of the schooldays during the first term (98 days) in grade one. Of these children, 39 (85%) were categorized as *at-risk* children and 13% as *strugglers*. Of the children having been at least two thirds of the time in school, 9% were *at-risk* children and 14% were *strugglers* according to the analysis.

Table 30: Distribution in clusters by school attendance (half of the school year, 98 school days)

Attendance days	1		2		3		4		Total	
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
1-33	39	84.8	6	13.0	1	2.2	0	0	46	100
34-66	24	32.0	39	52.0	12	16.0	0	0	75	100
67-98	19	8.8	30	13.9	68	31.5	99	45.8	216	100
Total	**82**	**24.3**	**75**	**22.3**	**81**	**24.0**	**99**	**29.4**	**337**	**100**

6.2.4 The influence of gender and age

Further analyzes involved the influence of *gender* and *age* on the performance in the group test (summary score), and on school marks. Multivariate GLM - MANOVA was used in the analysis. The results are presented in this section.

Gender had no significant effect (Table 31) on the group test or on school marks (Wilks' lambda multivariate test; Wilks' lambda=.989, F (3,324) =1.247, p >.05, partial eta squared=.011). The univariate tests also expose the situation.

Significant effects were, however, found on *age* (Wilks' lambda=.950, F (3,324) =5.673, p < .05, partial eta squared=.050). The univariate main effects were significant (p < .05) but the effect sizes (partial eta-squared) were weak. Only 3% of the variance on the group test was explained, whereas the explanation of the variance on school marks was 4.5%.

Table 31: Gender and age effect on group test and school marks

Factors	**Tests**	**Wilks Lambda**	**F(3,324)**	**p-value**	**Partial Eta Squared**
Gender		.989	1.247	.293	.011
			F(1,326)		
	Gtest		.288	.634	.001
	Smks		.301	.584	.001
			F(3,324)		
Age		.950	5.673	.001	.050
			F(1,326)		
	Gtest		9.939	.002	.030
	Smks		15.451	.000	.045

Pair-wise comparisons, including means, standard errors and confidence intervals for *gende*r and *age* are presented for the group test and school marks (Table 32).

The differences between boys and girls were not significant in any of the tests, although boys seemed to outperform girls in the group test, but not in school marks, where the situation was the opposite. Significant differences were found between younger children (mean=396) and older children (Mean=247), both on the group test and on school marks (means =53 vs. 30). The younger children outperformed the older ones, a pattern which was noticed already in the earlier described results. The reason is considered to be related to the late enrollment of some children, mainly due to context factors.

Table 32: Gender and age differences on the Ggoup test and school marks

Factor	Test	N	Mean	Std. error	95% Confidence interval	
					Lower bound	Upper bound
Gender	**Gtest**					
	2	181	333.0	30.7	272.6	393.4
	1	156	310.4	36.0	239.5	381.3
	Smks					
	2	181	40.2	3.9	32.5	47.9
	1	156	43.6	4.6	34.5	52.6
Age	**Gtest**					
	2	311	396.3	12.6	371.5	421.1
	1	26	247.1	45.6	157.3	336.8
	Smks					
	2	311	53.8	1.6	50.6	56.9
	1	26	30.0	5.8	18.6	41.5

Note: Gender: 2=boys, 1=girls; age: 2=6 to 8 years, and 1=9 to 12 years.

6.3 The explanatory power of environmental determinants

The environmental factors of interest in this study included some home background factors and some school related factors. For the analysis the following home background factors were chosen: *parents´ reading ability, time for school work at home, number of books at home and getting support at home*. The chosen school background factors included *school status, attending nursery school, and school attendance*. The analysis was conducted by using the sum score of the scales in the group test as well as by the school marks. Multivariate GLM -MANOVA was used in the analysis.

The reason for choosing the indicated environmental factors was based on their expected relationship with reading and writing ability and school achievement, based both on earlier research and on knowledge about the studied context. This section provides the results.

6.3.1 Home factors

The analysis revealed significant multivariate main effect for all the analyzed home background factors (Table 33). The univariate main effects were also significant (Table 34). *Parents' reading ability* explained 15% of the variation in the group test, and 7% of the variation in the school marks. The effect of how much schoolwork children did at home was slightly smaller (10% for group test and 3% for school marks), but followed the described pattern for the effect of *parent's reading ability*. *Number of books at home* was the strongest factor, explaining 38% of the variance in the group test, and 21% of the school marks. *Receiving support at home* explained 10% of the variance in the group test. No significant effect was noticed on school marks.

Table 33: Home background effects on group test and school marks

Factors	Tests	Wilks Lambda	F(6,596)	p-value	Partial Eta Squared
Parents´ reading ability		.841	8.973	.000	083
			F(2,300)		
	Gtest		26.580	.000	.151
	Smks		10.946	.000	.068
			F(6,596)		
Doing Schoolwork at home		.889	6.032	.000	.057
			F(2,300)		
	Gtest		17.325	.000	.104
	Smks		4.558	.011	.029
			F(6,596)		
Number of books		.583	30.816	.000	.237
			F(2,300)		
	Gtest		90.514	.000	.376
	Smks		38.716	.000	.205
			F(6,596)		
Receiving support at home		.856	8.050	.000	.075
			F(2,300)		
	Gtest		17.074	.000	.102
	Smks		1.611	.201	.011

Significant mean differences across the scale levels were found for most of the home background factors in both the group test and the school marks (Table 34). In families where the parents were good readers and where many books were found, the children received high group test scores, as well as high school marks, compared to children from other families. There was also a difference in the group test results between children doing and not doing schoolwork at home, but not between those doing it only sometimes, and those who had indicated that they do schoolwork at home. School marks did not differ significantly between the children, according to the indicated activity of doing schoolwork at home, although slight differences in the means were noticed.

Table 34: Home background factors scale level differences in the group test and school marks

Factor	Test	N	Mean	Std. error	95% Confidence interval	
					Lower- bound	Upper-bound
Parents´ reading ability	**Gtest**					
	Good	222	442.5	16.5	409.9	474.9
	Fair	82	335.1	23.9	288.2	382.1
	None	33	128.4	42.3	45.1	211.7
	Smks					
	Good	222	59.1	2.9	53.4	64.8
	Fair	82	50.4	4.2	42.2	58.7
	None	33	22.3	7.4	7.6	36.9
Doing school-work at home	**Gtest**					
	Yes	47	405.1	28.1	349.7	460.4
	Sometimes	198	375.1	20.9	333.9	416.2
	No	92	236.6	19.3	198.6	274.7
	Smks					
	Yes	47	53.1	4.9	43.4	62.8
	Sometimes	198	52.9	3.7	45.6	60.1
	No	92	39.4	3.4	32.7	46.1
Number of books	**Gtest**					
	Many	96	554.8	24.8	506.0	603.5
	Few	152	353.2	21.8	310.3	396.2
	None	89	92.0	24.1	44.6	139.4
	Smks					
	Many	96	70.5	4.4	61.9	79.0
	Few	152	53.4	3.8	45.9	61.0
	None	89	18.4	4.2	10.1	26.8
Receiving support at home	Gtest					
	Yes	39	447.0	30.1	388.3	506.6
	Sometimes	198	350.0	18.8	313.6	387.5
	No	93	243.0	20.5	203.2	284.0
	Smks					
	Yes	39	56.0	5.3	46.0	66.8
	Sometimes	198	47.0	3.3	41.1	54.1
	No	93	45.0	3.6	38.0	52.1

6.3.2 School factors

All the school related factors included in the study had effects on the results (Table 35). A significant multivariate main effect was noticed. Univariate main effects were significant for all the variables in the group test and on school marks (Table 36). School status explained 6% of the variance on the group test, and 14% on school marks. The effect of having been in nursery before entering the school was evident: 17% of the variance on the group test, and 16% of the school marks was explained. The effect of school attendance was strong: 21% of the variance on the group test was explained. For school marks, the figure was lower (8.5%).

Table 35: Effect of school factors on group test and school marks

Factors	Tests	Wilks Lambda	F(9,903)	p-value	Partial Eta Squared
School status		.708	12.306	.000	.109
			F(3,301)		
	Gtest		6.476	.000	.061
	Smks		16.804	.000	.143
			F(3,299)		
Attending nursery		.812	23.030	.000	.188
			F(1,301)		
	Gtest		61.996	.000	.171
	Smks		57.253	.000	.160
			F(6,598)		
School attendance		.762	14.540	.000	.127
			F(2,301)		
	Gtest		40.134	.000	.211
	Smks		13.962	.000	.085

Significant mean differences on group test and school marks were found between the groups in several variables (Table 36). The role of school status on the results of the group test was noteworthy. The good performance school (Mean=376) outperformed the high performance school (Mean=327) in the group test. Similarly, the low performance school (Mean=278) outperformed the periphery average school (Mean=234). However, both the high and good performance schools had comparatively higher means than the other two schools. Regarding the school marks, the order was different. The children in the high performance school had the highest mean, whereas the children in the low performing school had the lowest.

The effect of nursery was very clear, with very high scores in the group test (Mean=406) and on school marks (Mean=55.5) for children having been in nursery school. The role of school attendance was very strong, especially in the group test. The children with high school attendance had a mean of 420, compared to 304 and 88 for children with lower attendance. The effect on school marks followed the pattern, but the difference between the two groups with higher attendance was not significant.

Table 36: School factors scale level differences in the group test and school marks

Factor	Test	N	Mean	Std. error	95% Confidence interval	
					Lower bound	Upper bound
School status	**Gtest**					
	4	107	326.6	25.0	277.5	375.7
	3	80	376.1	20.2	336.4	415.9
	2	71	234.3	29.5	176.2	292.4
	1	79	277.6	23.3	231.6	323.5
	smks					
	4	107	60.5	3.3	54.0	66.9
	3	80	43.1	2.7	37.9	48.3
	2	71	37.1	3.9	29.5	45.0
	1	79	29.2	3.1	23.1	35.2
Attending nursery	**Gtest**					
	2	186	406.0	16.1	374.2	437.5
	1	151	210.1	19.0	172.8	247.2
	Smks					
	2	186	55.5	2.1	51.3	59.7
	1	151	30.7	2.5	25.8	35.6
School attendance	**Gtest**					
	3	110	428.2	22.6	383.7	472.6
	2	181	303.9	17.6	269.2	338.6
	1	46	88.2	30.5	28.2	148.2
	Smks					
	3	110	48.6	3.0	42.8	54.5
	2	181	45.3	2.3	40.7	49.8
	1	46	23.5	4.0	15.6	31.4

Note: School status: 4=high performance, 3= ood performance, 2=periphery average, 1=low performance; Attending nursery: 2: attended, 1=not attended; School attendance: 3=very good, 74-98 days, 2=good, 34-73 days, 1=poor, 33 days and less.

Summary

The preceding chapter presented the psychometric features of the created screening tool for identifying children at risk of developing reading and writing difficulties. It also presented the explanatory power of the environmental determinants on the outcome measures.

In summary, the screening tool seems to function well. The descriptive data indicates a large range of variation in all the scales and good discrimination between the children.

The validity analysis of the screen included four types of validity: content validity, construct validity, criterion validity, and predictive validity. In order to achieve *content validity*, the test items were chosen directly from high frequency words from grade one regular school books in reading and writing. In addition, a panel of experts, including qualified grade one Kiswahili teachers and Kiswahili experts from a teacher training college and Tanzania Institute of Education were used for scrutinizing and evaluation of the items and the scales. Most items were judged to be appropriate.

The analysis of *construct validity* included, firstly, an item analysis aiming at ensuring the internal consistency of the scales. The reliability coefficients (Cronbach's alpha) of the scales were high. Secondly, a principal component factor analysis was conducted. The group test was found to be one-dimensional. It included one strong factor, which explained 93% of the variance. The factor was named initial literacy.

Based on the one-dimensional structure and the strong inter-correlations, a summary score of the scales was created for use in further analysis. The finding also justifies the use of a subset of the seven scales in the group test to create a reliable short version to be used as a screening test. (see Appendices 9 and 10).

Criterion validity was analyzed by correlating the scales in the group test with three criterion measures: the individual test, school marks, and the follow-up test. All the scales in the group test correlated significantly with the criterion measures (coefficients range from r=.80 to .92 with the individual test .80 to .86 with school marks and .83 to .91 with the follow-up test). Predictive validity was very strong. It was analyzed by correlating the measures from the two school years and also by cross-tabulating the two school years' clusters.

The processes of ensuring validity also had the purpose of boosting the *reliability* of the group test. The reliability scores (Cronbach's alpha) were very high in all the tests (group test, 7 scales: r=.91 to .97; individual test, 5 scales: r=.97 to .98; and follow-up test, 4 scales: r=.84 to .95). The high scores support the idea of creating a short version of the screening instrument.

In order to find children who are at risk of developing reading and writing difficulties a K-Means cluster analysis was performed. The 337 grade one children were classified in four clusters. The four clusters were named *at-risk*

children (24%), *strugglers* (22%), *readers* (29%) and *good readers* (29%). Grade one clustering was compared to grade two clustering. The majority of children in grade two (60 to 80 %) were found in the same clusters as in grade one. Looking at the situation from grade two to grade one, we find that 94% (49 of 52) of the *at-risk* children in grade two were already identified in the group test in grade one. All in all 16.8% of the children in grade two belonged to the *at-risk* group. This indicates good predictive validity of the group test.

Cut-off points for the groups, equivalent to the four clusters, were set based on the distribution of the summary score of the scales of the group test. The analysis revealed that 24.3 % (N=82) children (the *at-risk* group) had scores below 166. These children were considered to require intensive support. Also the 75 children (22.3%), having scores between 166 and 368 (the *strugglers*) were regarded as needing attention and observation. Children scoring above 368 were considered to manage well without special attention.

Children's distributions in cluster by gender, age, school status and school attendance was determined through cross-tabulations. Slight gender differences were seen. Some more boys than girls were found among the *at-risk* children and among the *strugglers*. Age differences were evident: 40% of the older students compared to 20% of the younger students were found in the *at-risk* cluster. There is a clear difference in the cluster pattern between the schools. The highest proportion of *at-risk* children was found in the low and average performance schools (37% and 35%). School attendance was a critical factor in the study. A total of 46 children had been in school for less than one third of the school days during the first term (98 days) in grade one. The majority of these children was found in the *at-risk* cluster (85%) and in the *strugglers'* clusters (13%).

The effect of gender and age was analyzed by Multivariate GLM-MANOVA. Gender had no significant effect on the group test, follow-up test or school marks. The age effects were significant ($p < .05$), but weak. Only 1.9 % of the variance on the group test and 4.5 % on school marks were explained by age.

In the study there was also an interest in analyzing the explanatory power of certain environmental determinants on the children's performance on the group test and on the school marks. The environmental determinants included a selection of home and school factors. Multivariate GLM -MANOVA was used in the analyses.

The home factors examined were *parents' reading ability*, *doing school work at home*, the *number of books available at home* and *receiving support at home*. The effect of all the factors was significant on the group test. The *number of books at home* was the strongest factor, explaining 38% of the variance. *Parents' reading ability* was also a strong factor, explaining 15% of the variation. *Doing school work at home* and *receiving support at home* explained 10% of the variance. The explanation effect of all the factors was slightly lower on school

marks than on group test. It can be noticed that receiving support at home had no significant effect on school marks.

The school factors examined were *school status*, *attending nursery school*, and *school attendance*. All the factors had significant effects on the results. The effect of *school attendance* was very strong in that it explained 21% of the variance on the group test. The effect of having been in nursery before entering school was also strong: 17% of the variance on the group test was explained. *School status*, although significant, explained 6% of the variance on the group test.

The explanation effect of the three school factors on school marks was also significant. Some differences compared to the explanation effect on the group test are of interest. The effect of *school attendance* on school marks was clearly lower (8.5%) compared to the effect on group test (21%). The effect of *attending nursery* was about the same for school marks as for group test. *School status* should also be noticed. The explanation effect of the variance of school marks was 16%, which is more than double the effect on the results of the group test (6%).

7 Discussion

7.1 Summary of the process

The aim of this study was to create a group-based screening tool in Kiswahili, which has a transparent orthography, for identifying beginning readers in grade one in Tanzania at risk of developing reading and writing difficulties. The need for early screening for the purpose of providing early literacy intervention for children in need of support has been discussed in the introductory part of this thesis. The main theories that brought light and insight into the creation of the screening tool include the simple view of reading, the componential model of reading, as opposed to discrepancy theory. The tool also fits the Tanzanian context, where Kiswahili language is not only the national and official language, but also a medium of instruction in all public primary schools and taught as a subject in secondary schools. English is also an official language in Tanzania. It is used as a medium of instruction in public secondary schools and in primary and secondary private schools. In private schools Kiswahili is taught as a subject.

The nature of the study is quantitative. All the test results are numeric and the questionnaire responses were quantified. The analyses include item analysis, correlations, factor analysis, cluster analysis, and multivariate analysis with GLM procedures. The design of the study is longitudinal. In mid-July 2007, 337 grade one children undertook a group test of seven subscales with cognitive linguistic indicators measuring *word identification, letter identification, phonological awareness and spelling* in order to find the children at risk of reading and writing difficulties. One month later, a smaller sample of 64 children drawn from the main sample participated in an individual test including six subscales measuring pseudo-word reading, actual word reading, actual text reading, one-minute reading and writing. In mid-July 2008, when the children (N=330) were in grade two, they participated in a follow-up test, including a selection of four subscales from the test in grade one. The individual test, follow-up test, and school marks from mid-grade one served the purpose of validation. No other criterion measures could be found in Tanzania at the time of the study.

In this chapter, I firstly concentrate on discussing the challenges in creating the instrument. Secondly, I discuss the psychometric properties of the screening tool in the light of justifying the validity and reliability. Thirdly, I discuss the identification of children at risk and the determination of the cutoff point. Fourthly, the effects of gender, age, home and school factors on the group test are discussed. The last sections include a discussion on the limitations and on the significance of the study, and recommendations and suggestions for further research. The chapter closes with a conclusion.

7.2 Challenges in the construction of the group test

Several challenges had to be addressed in the process of constructing and validating the screening instrument. Some of these challenges are exposed in the methods chapter as *matters to consider in the construction of the instrument.* The challenges encountered are discussed in this section.

Most researchers agree that the screening of children at risk of developing reading and writing difficulties should take place before children start grade one (see e.g., Davis, Lindo & Compton, 2007; Jenkins & O'Connor, 2002; Ritchie & Speece, 2004). This suggestion is better effected in contexts where all children, or rather most, pass through nursery schools in formalized preschool conditions. For Tanzania this would not be feasible, since not all children go through nursery school. For this study, it was therefore decided to screen children in grade one. Screening at this level would still not be late for children to be identified and provided with appropriate early intervention. Research findings show that children up to grade three with difficulties in reading and writing can be easily assisted. Beyond this point it becomes difficult to assist them (Rathvon, 2006).

In most studies, screens involve large and extensive samples. Screens are basically surveys (see Rescorla, 1989). In this study, the ambition was to get a representative sample appropriate to generalize the findings to the wider Tanzania. Tanzania is an expansive country with inhabitants living both in urban and rural settings. Hence, the ideal was to have a sample representing urban and rural schools. Due to lack of practical and financial possibilities to cover a wider area of Tanzania, the sample for this study was purposively selected. It was reasoned that Morogoro municipality could be representative enough, based on its geographical, socioeconomic and educational status. Morogoro is located close to the coast and close to the hinterland.

The schools selected, although they are in the municipality, have children from both urban and rural settings. After deciding on what grade level of children to screen and deciding on the sample area, the school status (i.e. *high performance*, *good performance*, *peripheral average performance* and *low performance* schools) based on school inspectors' reports was considered and included as an important basic variable guiding the selection of the sample. The aim of using *school status* in the selection was to ensure variation representing the population. Other variables considered important for describing the sample were *age* and *gender*, and certain home and school factors anticipated to be of value for learning to read and write (see Johnson, Pool & Carter, 2009). These factors included, e.g. *parents´ reading ability*, *the number of books at home*, *attending nursery school,* and *school attendance*. Information about home and school factors was collected through a questionnaire.

Screening instruments are available to assess children on either an individual or group basis. Individual assessment is possible and purposeful if there is a good amount of personnel available for the task. This is, however, mostly not the case

in contexts where there is only one teacher and where the class sizes are large. Hence, creating a group-based screen was considered ideal for the Tanzanian situation, where class sizes are very large in most public schools. Class sizes in Tanzania may surpass the set standard size of 45 pupils. Class sizes from 80 to 100 pupils are rather common and found in many places. This context makes it difficult to plan for individualized assessments, which may involve too costly, time-consuming and cumbersome procedures for implementation (see Jenkins & Johnson, 2007; Jenkins, Hudson & Johnson, 2007). In terms of this problem it was necessary to select predictors and design tasks that could be administered at group level in grade one.

Which cognitive linguistic predictors would be most suitable and relevant for identifying at risk children using a transparent orthography? How could these be translated into tasks that could be effectively managed in a group situation? These questions were addressed by examining the available literature and complementing this with expert judgment. Before deciding on the linguistic components of the screen, it was above all necessary to find an operational definition of reading and writing, and from that to determine the implication of writing difficulties (Jenkins & Johnson, 2007). The literature on the topic reveals a variety of reading disability (difficulty) definitions (Kerka, 1998; Siegel & Ryan, 1989). The resulting instrument was constructed with an embodiment of crucial cognitive linguistic predictors judged to be most significant to later reading achievement: in this study, *phonological awareness* was measured by initial sound and spelling; *letter identification* was measured by picture-letter and letter-picture tasks; *word identification* was measured by picture-word, word-picture and word chain tasks. *Spelling* was also a task on its own in the group test.

Phonological awareness was selected, as it has been found to have a unique impact on reading acquisition in previous studies. This linguistic indicator is important for reading achievement also in a highly transparent language. It has also been observed to account for unique variance, especially in the middle of grade one and at the end of grade two. Its predictive power, however, decreases as a function of time. Hence, it becomes less effective in grade two and beyond (Sprugevica & Høien, 2003). *Letter identification*, on the other hand, was selected as it has been found to be a more powerful factor than *phonemic awareness* when explaining variance in early reading acquisition (Sprugevica & Høien, 2003). *Word identification* was selected following the simple view of reading, which stipulates that decoding, together with comprehension, are the most vital components of reading. Spelling was also selected as an indicator in this study for its vital contribution to early reading acquisition and its unique impact, especially in a transparent language (Alcock et al., 2000; Aro, 2004; Lyytinen & Erskine, 2006). These indicators and their related tasks are graphically illustrated to show the conceptual model of reading and writing in this study (see Figure 3).

In order to establish the content validity of the instrument, there was a need to identify the source of the items used to design the related tasks of the measures. Many available screens have adopted items used in available criterion measures. Because of a lack of such corresponding measures for the screen in this study, there was a need to have an original source for the items. This source had to contain items relevant to the children's reading and writing curriculum. Three authorized class reading books for grade one were selected. A manual word count of the three books was made and lists of high and low frequency words were drawn up. From these lists, items were pulled out to construct the tasks.

The acquiring of initial literacy skills means that the reader is able to decode fluently and accurately. In this study, *phonological awareness, word identification, letter identification,* and *spelling* have been shown to be vital for learning to decode. *Phonological awareness* has been extensively studied and shown to be an indicator and contributing agent in reading in opaque languages, English in particular, but has also shown in this study to be an effective indicator in the transparent language, Kiswahili. Some authors (e.g. Alcock, et al., 2000; Aro, 2004; Lyytinen & Erskine, 2006) have suggested that *letter knowledge* could be the single indicator of reading and a contributor to reading and writing, especially in regularly spelled languages. These indicators were also suggested for this study by a group of experts.

The cognitive linguistic indicators selected for this study (i.e. *phonological awareness, word identification, letter identification,* and *spelling)* were adequate enough for the purpose of finding children at risk of reading and writing difficulties (see Jenkins & Johnson, 2007). The performance of the children in the aspect of initial literacy skills was quite informative in relation to the objectives of the study. The scores were discriminating, although a little skewed. The results showed that by halfway in grade one some children were still struggling to acquire the initial reading skills to be able to decode, while others had managed to break the code. Further, the follow-up in grade two showed that a large number of children facing problems in reading and writing in grade one continued to face problems in grade two. This implies that the predictive aspect of the group test is rather good, as there had not been any special intervention, which could have been expected to reduce the problems considerably. It was, however, found that some children who had been in the *good readers* group in grade one had migrated to the lower *readers* group. A decrease of reading performance with increasing grade has also been noticed by, e.g., Johnson et al., (2009). No explanation for the migration of the children to a lower level was found when analyzing the background of these migrated children. However, it can be speculated that a low level of motivation in the test situation or being sick at the time of taking the test might be the cause. From the risk point of view, however, these children were still not out of danger of being at risk.

7.3 Psychometric properties of the group test

The results of this study indicated that the group test battery has adequate psychometric properties required for a screen. The reliability coefficients (Cronbach's alpha) ranged from r=.91 to .97. All the seven subscales intercorrelated highly. The high correlations (r=.88 to .96) can be explained from the one factor structure that was extracted, indicating that all the tests were carefully constructed and that they seemed to measure the same construct. Another explanation is based on the very well-established content validity.

Evidence of content validity was established by theory and following advice from a panel of experts. The theory basis was on the simple view of reading and the componential model of reading. The experts also suggested what cognitive or linguistic predictors would be most sensitive to reading and writing difficulties in a transparent language like Kiswahili. The process through which the screen was constructed and administered also contributed to the enhancement of content validity. The test items for the subtests were drawn from the authorized instruction and reading textbooks used in grade one in Tanzania. The words were drawn from a high frequency list. Hence, the test items (words and pictures used) were familiar to the children. The administration and invigilation of the tests was, generally, carefully planned and executed.

In the validation process of the group test, subscales of the individual test, follow-up test and school marks were used as criterion measures. Concurrent validity was determined by the children's (N = 337) school marks. This validation design was adopted due to lack of a Tanzanian standardized screening tool which could have served as criterion measure.

Linguistic components known from research, based on individual testing, to be sensitive predictors of reading and writing skills were chosen for validating the created group test. The following components measuring reading fluency and writing were chosen for individual testing: pseudo-word reading (Baker, Park & Baker, 2010; Good, Baker & Peyton, 2009), reading words aloud and actual text reading (Adams, 1990; National Reading Panel, 2000a; Snow et al, 1998; Stanovich, 2000), one-minute reading (Brus & Voeten, 1973), and writing (Moats, 2006). These dimensions are regarded to be sensitive in predicting which children, in the beginning of acquiring literacy skills, have acquired the enabling skills and which have not, and are still struggling and therefore in need of effective intervention.

The school marks and follow-up test were also used for validating the group test. The correlations were also high. The school marks originate from an average score of three teacher-made tests of *reading*, *writing* and *language,* taken halfway in grade one. The school marks were also used to sample 64 children who took the individual test one month after the group test was administered. The follow-up test was administered one year later, halfway through grade two. The use of the individual test, school marks and follow-up test for validation has been an effective strategy where there has been a lack of a criterion measure.

Predictive validity was established by examining the extent to which children's results from the group test in grade one correlated with the results in the follow-up test in grade two. Correlations between the two tests were high. Cross-tabulation between grade one and grade two clusters were also used to establish predictive validity.

Construct validity was established by item and factor analyses. Item analysis was conducted to determine the internal consistency of both the group test and follow-up test. Item analysis was first established by content validity, which has already been described. Secondly, it was quantitatively analyzed by computing the p-value to get item difficulty indices and the D-value to determine the discrimination indices for each item used. Most of the items were within an acceptable range of difficulty (i.e. p-value was between 0.15 and 0.85. The items also discriminated well. The D-value was from 0.30 and above. In this respect, most of the items were retainable. The few items that did not meet the criteria were, however, judged as also important for motivating the struggling readers, i.e. the easy ones; or for challenging the good readers, i.e. the difficult ones. Construct validity was further determined by principal component analysis. The results indicated a one-dimensional structure, named *initial literacy factor*, for the group test. Similar results are reported with the get ready to read screening tool (Whitehurst, 2003), although this is a screen for children in the pre-kindergarten year.

7.4 Identifying at risk children and cut-off point

The K-Means cluster analysis resulted in four distinct groups of children based on their mean scores in the different scales in the group test. The groups were named: *at-risk, strugglers, readers,* and *good readers.* These group names reflect the characteristics of the children in relation to their acquisition of the early skills of reading and writing. The *at-risk group*, which is the target group in this study, is characterized by obtaining very low scores, or none in all the scales. Conversely, the *good readers* group comprises children scoring highly in all the measures. This grouping pattern seems to be consistent with earlier findings by Burrow, Cabell, Konold and Invernizzi (2010) and also with the findings of Huang, Ford and Invernizzi (2011). For example, Huang and his colleagues (2011) found four clusters, with the highest cluster operating above average in all the predictive indicators used (i.e. *beginning sound, rhyme awareness, alphabet recognition letter sounds* and *invented spelling*). The second highest cluster operated above average in all the predictors, but slightly lower compared to the highest group. The third cluster was average in two indicators (*beginning sounds* and *rhyme*) and below average in three predictors (*alphabet cognition, letter sound* and *invented spelling*). The fourth cluster had children performing below average in all the measures. It should be noticed that *beginning sound, rhyme, letter sound and spelling* are measures of *phonological awareness,* while *alphabet cognition* is a measure of *letter knowledge.* The

clusters and the dimensions in the present study were similar to the outcomes in the study by Huang et al. (2011).

In order to set up the cut-off point on the group test, a summary variable was created and a score was counted based on the seven scales of the group test. The analysis of the distribution revealed that 24.3% (N= 82) of the children, an amount comparable to the *at-risk* group from cluster analysis, had scores below 166. The suggestion is consequently that children scoring lower than 166 belong to the *at-risk* group and need intensive support. The 75 children (22.3%) scoring between 166 and 368 also belong to a group needing attention and observation. They can somehow manage, but they are struggling. With extra attention and some support, most of them can probably manage well. Children scoring above 368 are doing fine. When using the test, however, notice should be given to the fact that many background factors have an effect on the test results. It should be noted that due to the influence of various factors that affect the test scores, different cut-off scores could be set for the risk groups in different kinds of schools (see Jenkins & Johnson, 2007; Jenkins, Hudson & Johnson, 2007).

7.5 Gender and age effects

This part includes short comments on the effects of *gender* and *age.* The effect of *gender* on the group test scores was not significant. However, the boys had on average slightly higher means than the girls. This is in accordance with results from recent Tanzanian studies, where similar kinds of literacy skill components have been studied in a low income area of Dar-es-Salaam (Kumburu, 2011) and in a poor rural area in Eastern Tanzania (Ngorosho, 2011). However, the finding is in contrast with results from a large country-wide survey in Tanzania (Uwezo, 2010). In this survey, the girls slightly outperformed the boys, a finding that has been common in many literacy skill studies in other parts of the world, e.g., the PISA 2000 study reported by (Organization for Economic Co-operation and Development, OECD, 2001) and also the International Association for the Evaluation of Educational Achievement, IEA Reading Literacy study 1985-1994 reported by (Elley, 1994). More studies about gender differences in learning are needed in Tanzania.

The results regarding *age* should be noticed. An *age* effect was observed, and the younger children outperformed the older children significantly. This is inconsistent with earlier studies reported in developed countries, where older children usually perform better than younger children; for example, on the Brigance K & 1 screen (Mantzicopoulos, 1999). This is not the case for older children in Tanzania. Many parents in Tanzania cannot send their children to school at the proper age for various reasons, including poor economic situation, far distances and lack of transport. Most of these children come from a low socioeconomic environment, where the home background does not favor their learning effectively. At home, the older children mostly remain idle, are sent on errands or engage in petty business. At school, these children might find themselves unable to keep up with the younger ones in learning. Their

frustration probably increases their learning difficulties. The implication is that at risk children are more vulnerable with increasing *age*.

7.6 Environmental determinants

This section includes comments and reflection on the effects of home background factors, and a discussion on the effects of certain school related factors of special importance.

The home factors used in the study were the following: *parents' and guardians' reading ability, the child doing schoolwork at home, the number of reading books at home, and the support which the child receives on school work at home*. These factors were used as indicators in the analysis of the effect of children's home background factors on the performance on the group test. All these factors had a significant effect. The number of reading books at home accounted for not less than 38% of the variance on the group test results. Effect of home environment factors on children's literacy skills in Tanzania has also been shown in other recent studies (see Ngorosho, 2011). Parents' education and books for school subjects at home and some other variables (e.g. housing) were identified to be important descriptors of home environment in rural Tanzania. Father's education and books at home emerged as strong predictors on a variety of literacy skill components.

The importance of literacy related home factors for children's literacy development has also been revealed in many international studies (e.g. PIRLS, 2001 reported in NFER, 2003). According to PIRLS, *number of reading books at home* accounted for 10-15% of the variance on reading skills. There seems to be much higher influence of the home environment on early reading in Tanzania than in the countries included in PIRLS. (see also Martin, Mullis & Kennedy, 2003). *Parents reading ability* was the next strong factor in the present study, accounting for 15% of the variance. Other important factors were *doing school work at home* and *receiving support at home*, which explained 10% of the variance.

This finding reflects at least partly the socioeconomic status of the family. In well-to-do families (both economically and academically) children can enjoy the support of their parents both materially and literary; and they can also get good models for learning. Findings by Uwezo (2010) specifically suggest that the mother's education is important for children's schooling at all levels. In Ngorosho's study in the rural area, the father's education was a stronger predictor than the mother's. The explanation is that very few of the mothers had more than the lowest level of formal education and thus the variance was low. In any case, parent's education has an impact on the development of the children's literacy skills in many ways. Educated parents can, e.g. afford the time to stay at home and help their children with schoolwork or pay a tuition fee to someone to teach their child privately. In these homes children can get books to read at home. Mason and Allen (1986) contend that homes that have a large number of

books tend to be homes where parents read to their children on a consistent basis.

In poor families parents cannot afford to buy schoolbooks for their children, and due to scarcity, after classes school books are left at school and not allowed to be brought home. Parents in poor families, moreover, may be too busy trying to provide for the family. In such contexts, children lack the attention and support of their parents. The children may also not have time to stay at home and do schoolwork; rather, they will be on the streets engaged in petty business or laboring in the neighborhood. Hence, in high socioeconomic status families, children are more likely to be stimulated in learning compared to children in low socioeconomic status families (see, e.g. Leroy & Symes, 2001; Knizer & Klein, 2007).

School factors in this study included *attending nursery school, school status,* and *school attendance*. Having attended nursery explained 16% vs. 17% of the results in group test and school marks. The effect was expected and it is in accordance with earlier findings. Attending nursery school makes a difference (e.g. Aboud & Hossain, 2011; Foster & Miller, 2007; Gillian & Ziggler, 2000; Magnuson, Meyers, Ruhm & Waldfogel, 2004; Nonoyama-Tarumi & Bredenberg, 2009; PIRLS, 2005/2006; Sylva, Siraj-Blatchford & Taggart, 2003; Valenti & Tracey, 2009).

On *school status*, the results indicated that the good performance school outperformed the *high performance* school and the *low performance* school outperformed the *peripheral average performance* school. The findings indicated that *school status* had an effect on the achievements in the group test, but the explained variance was only 6%. The differences were only seen between the two highest and the two lowest status schools. Regarding school marks, 14% of the variance was explained. The school with the highest status outperformed the school with the second highest status and the third one in order outperformed the fourth one. These findings need some comments.

The achievements of the children measured by the group test did not differ clearly between the all the schools with different status, which was the case regarding school marks. The explanation might be that the teachers make different tests in the different schools and these are not standardized. The school marks seem to follow and reflect the status which the school has obtained in the school inspectors´ assessment. It can be speculated that the status of the school also affects the teachers' perception of the children's achievement levels and that it thus can have consequences for giving the school marks. This means that there can be a tendency among the teachers to give marks according to the tradition in the school and that this tradition can be related to the status of the school. Hence, it is possible that there can be a tradition where children in high and good performing schools receive higher marks than those in low and periphery average schools and not necessarily strictly according to the performance. The status of the schools is based on many factors, as given by the school

inspectorate, not only on the performance of the pupils. All in all, the effect of *school status* on school achievement among children in Tanzania is significant. School differences with consequences for children's achievement are found in developing as well as in developed countries.

School attendance had a clear multivariate and univariate main effect on the results. The findings indicate that school attendance explained 21% of the results of the group-test, which could be expected. The reasons for differences in school attendance are many and mostly related to home- and other context factors. There was also a clear effect on school marks. However, only 8.5% of the variance in school marks was explained by *school attendance*. The effect on school marks discussed above in relation to *school status* is regarded to be relevant in this case also.

School attendance needs further discussion. The cluster analysis showed very clearly that *school attendance* is a critical factor in this study. About 14% of the children spent less than one third of the schooldays in school during the first half of the school year. Of these children, as many as 85%, belonged to the *at-risk* group for reading and writing difficulties. Looking at the whole group of children (36%) who attended school on up to two thirds of the schooldays, we find that 77% belonged to the *at-risk* group. Of the children attending school more than two thirds of the school days, we found about 9% belonging to the *at-risk* group. All in all, it was rather common that the children did not come to school every day or every week. The reasons for low attendance are not analyzed in this study, but long and unsafe walking distances, especially during the rainy season, in addition to poverty in the family, are some of the likely reasons. Other reasons are related to the education of the parents and the view of the importance of schooling for the future of the child.

It seems logical to think that poor school attendance is the reason for low reading and writing ability and being at risk of developing reading and writing difficulties. Negative relationships between school attendance and academic performance have been reported in earlier studies (e.g. Alexander, Entwistle & Horsey, 1997; Caldas, 1993; Lamdin, 1996; National Center for Education Statistics, NCES, 2007). However, one also has to think the other way around. With low literacy support in the family, where the parents might not know how to read and write, the preconditions for learning in the school might not be good. In a school with big classes, where 3-5 children share a desk, where 5-10 children might have to share one textbook during the lessons and no one is allowed to take the book home, it is not unusual that the children have difficulties in following the instruction and meeting the requirements of the curriculum. If a child experiences that he or she does not understand, does not learn, does not do well enough and get good marks, why then go to school? Experience of low achievement can in many cases be the reason for low attendance and not only the other way around.

The conclusion is, in any case, that these two factors are related and that low scores in the group test in this study could be used in finding the *at-risk* group, which is large among children with *low school attendance*. However, a risk group of almost 9% is also found among children with *high school attendance*, which indicates that *low school attendance* is not the only explanatory factor for low achievement and risk status in the study.

In summarizing this section it can be said that the findings on the effect of the environmental determinants (i.e. the selected home and school factors) are very marked. It seems obvious that there is a tendency of higher influence of home- and school factors in this study on school achievement compared to findings in other countries, especially in the western world. Hence, these findings should stimulate more studies in developing countries so as to gain more knowledge on the influence of environmental determinants on school achievement.

7.7 Limitations

The intention was to screen 400 or more children representing a variety of schools in both urban and the rural areas. Although the schools were carefully selected, the sample included more children from *high performance* and *good performance* schools than from *periphery average* and *low performance* schools. However, considering the aim of the study, the sample is considered to be representative of the schools in the area. For generalization to the whole country, the sample is, however, not good enough, although Morogoro, where the study was undertaken, has many basic features in common with other parts of the country.

One aspect to mention is that it was not practically possible to administer the tests to all students in the four schools on the same day. There was, nevertheless, no risk of an effect on the performance between the schools. The schools were located far apart, and therefore speculation can be ruled out that children in one school would have known beforehand, from their counterparts, what they were going to be tested on when their turn came.

There is also one issue related to the measurements which needs comment. In analyzing the scores on the items, it became evident that some children scored zero on items with four options towards the end of some of the scales. This was not expected. Not knowing the correct answer and thus guessing between four alternatives would have had a consequence of at least 25% correct answers on the items experienced to be difficult. The explanation is that the children got tired or did not want to guess when they were unsure about the correct answer, and thus they did not complete the task, although the instruction was that all tasks should be done. Because of the high number of items of different difficulty level in scales correctly filled in and completed, the conclusion is that the described problem has no significant consequence for the functioning of the group test. The observation and the problem encountered have been taken into account in designing a short version of the screening instrument.

Despite the limitations noted, it can be concluded that the use of an individual test, with a smaller sample drawn from the main sample, and a follow-up test as criterion measures for validation provides the strength of the created group test in identifying children at risk of developing reading and writing difficulties.

7.8 Significance of the study

The successful completion of this study is firstly expected to create awareness among educational stakeholders of the existence of children with learning difficulties, who have not been catered for in Tanzania, specifically children with reading and writing difficulties. Secondly, the study ventures to provide parents, teachers and other interested parties with a reliable and valid instrument for screening students in order to identify those at risk of developing reading and writing difficulties. Hence, a culture of early screening of children in Tanzania may be established. Thirdly, teachers and other educationists will be able to determine the appropriate steps to take or be led to an appropriate remedial instruction program for children with disabilities. Fourthly, this study is expected to be a way of preparing regular teachers for effective participation and contribution to the focused inclusion set-up in Tanzania. In a long-term focus the study thus implies the promotion of reading and writing among beginning readers.

The psychometric findings suggest that the group test developed in this study can be soundly used as a screening device in identifying beginning readers at risk of developing reading and writing difficulties. This device is especially efficient where Kiswahili language is concerned. However, it can also serve as a basis for developing adaptations in other transparent orthography realms. Further, the questionnaire (described in the methods part) would be very useful in obtaining information about the children's home and school background factors that could affect their acquisition of reading and writing skills. It is anticipated that teachers in grade one can easily use the screen to identify children in grade one requiring immediate intervention or further assessment.

7.9 Recommendations and suggestions for further research

Further research is suggested to focus on gaining experience of the created screen (the short version) in predicting different kinds of outcomes in school achievement after grade one, There is also a need for obtaining experiences of different kinds of samples, e.g., in different areas of urban and rural Tanzania. The function of the test could also be studied in other East African countries such as Kenya, Uganda, Ruanda and Burundi with Kiswahili speaking children.

There is also a need to educate teachers carefully in using the test in order to achieve reliable results. Although the use of the test is simple, it requires some training in the technicalities of its administration and in the interpretation of the results. Pre-service and in-service training of the teachers is needed.

Implementation of regular screening in grade one for finding children at risk of developing reading and writing difficulties is suggested. This can be easily and effectively done by explicitly including early and regular screening in school curricula. It is recommended that the implementation should be followed by systematic research. Supporting strategies and intervention programs for children found through the screening are needed. Evaluation processes should also be developed. Decision makers and educational planners can support this by seeing the possibility of allocating more resources on classroom evidence based teacher practices and children's general learning, and particularly the acquisition of literacy skills.

8 Conclusion

The results provide evidence that the group test created can be used for identifying first graders at risk of developing reading and writing difficulties in the Kiswahili language. However, the method followed in the creation and validation of the screen can be used universally, especially where the language is of transparent orthography and where there are problems in identifying criterion measures for validation. In this study validation was done through the use of individual testing, school marks and a follow-up study.

With the group test created it is possible to identify *at risk readers, struggling readers* and *good readers*, half way through grade one. Following the good reliability and validity, the test is recommended as being useful for teachers, parents and other stakeholders interested in children's reading and writing process and wanting to know who is progressing well and who is facing problems. The created instrument in combination with home- and school background information is also useful for obtaining information about factors which are important for student achievement. It has been apparent in this study that school attendance needs to be emphasized among beginning readers alongside the learning of the basic literacy skills. It is suggested that parents should also be informed of the importance of improving the home environment so that it becomes literacy conducive.

Performing the test does not take much time and it is relatively easy to administer and score. It can be administered following a short training and guiding instructions. It is also cost effective in the sense that it does not require a lot of staff and expensive material to administer.

From the results of the item and factor analyses the indication is that it is possible to have a shorter version of the group test for screening. The shorter version can enable teachers to use less time, money and energy in screening. The shorter version has been constructed and evaluated. It is attached as Appendices 9 and 10).

Summary of the study

Aim and context of the study

The major focus of this study is on the identification of children at risk of reading and writing difficulties. Specifically, the aim was to create a theory-based group test in Kiswahili, to be used for identifying Tanzanian children in first grade who are at risk of developing reading and writing difficulties. The validity of the grade one group test was studied by comparing with an individual test and through a follow-up test in grade two. The test was created and its function is presented based on a sample of first graders from the municipality of Morogoro in Tanzania. In studying the function of the test there was special interest in the explanatory power of certain contextual factors related to the home (e.g. literacy) and the school (e.g. school attendance). These contextual effects are of interest because the study is conducted in Africa and there is little knowledge about the roles of schooling and home related determinants outside Europe and USA, or generally in developing countries.

Many children in Tanzania complete primary education without the skills of reading and writing (Chonjo, 1966; Malekela, 2003; Uwezo, 2010). High dropout and repetition rates and poor school attendance, especially in the lower grades are large problems in Tanzanian schools today. The reasons are many and complex, but it is evident that some of them are related to low reading ability in the school leavers. There are also children with learning problems, including reading and writing difficulties which are not identified and not aided. The described situation needs to be addressed. Reading skills are important for the children's learning of all subjects at school, and for their later socioeconomic life. This need is well stipulated in the National Policy on Disability (MLYDS, 2004, p. 4):

There is a need for early identification of children with disabilities. Early identification followed by appropriate intervention has a chance of eliminating the occurrence of a disability or minimizing its impact later in life. Unfortunately, Tanzania does not have a national programme for early intervention which would assist the identification of children with disabilities in their respective communities. The fact that parents hide their children with disabilities pushes them further away from accessing appropriate services.

Enabling skills

There is broad agreement among researchers in the science of reading, that deficit in the phonological system is an indicator of reading and writing difficulties (e.g. Catts & Kahmi, 2004; Høien & Lundberg, 2000; Kerins, 2006; National Reading Panel, NHI, 2000; Snowling, 2000; Vellutino et al., 2004). Deficits in core phonological processes seem to be central to many reading problems. These processes are often categorized into three areas: phonological awareness, rapid naming and phonological memory.

Phonological awareness involves an individual's awareness of the sound structure of spoken language. Phonological awareness is demonstrated when one can hear, remember and manipulate sound units within words, syllables and syllable fragments, i.e. onsets/rimes and phonemes (Bailet et al., 2009; McGuiness, 2004). Rapid automatized naming is the time it takes to name objects, e.g. letters or syllables aloud as quickly as possible. Rapid naming speed has successfully been used in revealing phonological awareness deficits and it is a strong predictor of reading skills. The importance of rapid automatized naming was already identified a quarter of a century ago by Denkla and her colleagues (Denkla & Rudel, 1972). Phonological memory is the third process implicated in the acquisition of language and reading skills. This refers to the coding or imprinting of phonological information for temporary storage in the working memory prior to storage in the long-term memory (Kerins, 2006; Wagner, Torgesen & Rashotte, 1999). The phonological loop associated with working memory provides a brief, verbatim storage of auditory information (Baddeley, 1986; Baddeley, 1992; Gathercole, Willis & Baddeley, 1991).

The acquiring of initial literacy skills means that the reader is able to decode fluently and accurately. *Phonological awareness* has been extensively studied and shown to be an indicator and contributing agent in reading in opaque languages, English in particular, but has also shown in this study to be an effective indicator in the transparent language, Kiswahili. Some authors (e.g. Alcock, et al., 2000; Aro, 2004; Lyytinen & Erskine, 2006) have suggested that *letter knowledge* could be the single indicator of reading skills and an important contributor to reading and writing skills, especially in regularly spelled languages.

Creating a group-based screening test

Creating screening tests is a demanding task, both theoretically and empirically. In the creation of screens three steps should be taken (Jenkins & Johnson 2007; Jenkins, Hudson & Johnson, 2007). The first step is to define the future outcome the screen seeks to predict, the second is to identify early predictors and the third is to set a cut-off point for identifying children at risk of failing the future criterion test.

In this study a screening test to be used in groups was created. The test was constructed to screen first graders in reading and writing. The essence was to identify children at risk of reading and writing difficulties. The group test is founded on current understanding of reading and writing and the related difficulties. The conceptual model from which the subscales and the test items were conceived is based on ideas from the simple view of reading, the componential model of reading and the response to intervention model. In the simple view of reading, the understanding is that reading is a product of decoding and comprehension. However, the complexity of reading is explained in the other various intertwined aspects involving the cognitive linguistic mind, and psychological, motivational and environmental factors.

Study design

The design of the study is longitudinal. In mid-July 2007, 337 grade one children aged 6 to 12 (Mean = 7.25), from four purposively selected primary schools in the municipality of Morogoro, undertook a group test measuring cognitive linguistic skills with seven subscales concerning *word identification, letter identification, phonological awareness and spelling.* The aim was to identify children at risk of reading and writing difficulties. One month later, a smaller sample of 64 children drawn from the main sample participated in an individual test including six subscales measuring *pseudo-word reading, actual word reading, actual text reading, one-minute reading and writing*. In mid-July 2008, when the children (N=330) were in grade two, they participated in a follow-up test, including a selection of four subscales from the test in grade one. The individual test, follow-up test, and school marks from mid-grade one served the purpose of validation. No other criterion measures could be found in Tanzania at the time of the study.

The schools selected, although they are in the same municipality, include children from both urban and rural settings. In addition to grade level and the sampling area, school status (i.e. *high performance*, *good performance*, *peripheral average performance* and *low performance* schools), was considered an important basic variable guiding the selection of the sample. The aim of using *school status* in the selection was to ensure variation representing the population. School status was based on school inspectors' reports. The description of the sample includes, in addition to age and gender, certain home and school factors anticipated to be of value for learning to read and write (see Johnson, Pool & Carter, 2009). The selected home factors included *parents´ reading ability*, *the number of books at home* and *receiving support on school work at home;* and school factors included *attending nursery school, school status* and *school attendance*. Information about home and school factors was collected through a questionnaire filled in by parents and teachers.

Psychometric properties of the group test

The descriptive data indicates a large range of variation in all the scales and good discrimination between the children. The group test battery has adequate psychometric properties required for a screen. In the validation process of the group test, subscales of the individual test, follow-up test and school marks were used as criterion measures.

Construct validity of the screen was established by item and factor analyses. Item analysis was conducted to determine the internal consistency of both the group test and follow-up test. The content validity of the items was first secured qualitatively. The dimensions behind the scales were theory-based and the test items were drawn from a high frequency list of words counted from the authorized instruction and reading textbooks in grade one. Secondly, the items were quantitatively analyzed by computing the p-value to obtain item difficulty

indices and the D-value to determine the discrimination indices for each item used. Most of the items were within an acceptable range of difficulty (i.e. p-value was between 0.15 and 0.85). The items also discriminated well. The D-value was from 0.30 and above. In this respect, most of the items were qualified to be included when developing further versions of the screen.

Construct validity was further determined by principal component analysis. The results indicated a one-dimensional structure, named *initial literacy factor*. This means that all the scales measured the same construct. All the seven subscales intercorrelated highly.

The reliability coefficients (Cronbach's alpha, r =.91 to .97) of the screen were high, which should be expected based on the high intercorrelations, one-dimensional structure and good content validity described above. The high reliability also reflects that the test was carefully constructed and performed.

In validating the screen (the group test) the individual test was used as the first important criterion measure. Linguistic components known to be sensitive predictors of reading and writing skills were chosen for validating the created group test. The following components measuring reading fluency and writing were chosen for individual testing: pseudo-word reading (Baker, Park & Baker, 2010; Good, Baker & Peyton, 2009), reading words aloud and actual text reading (Adams, 1990; National Reading Panel, 2000a; Snow et al, 1998; Stanovich, 2000), one minute reading (Brus & Voeten, 1973), and writing (Moats, 2006). As the individual test was performed rather close to the screening test (the group test) it reflected concurrent validity. The children's school marks from mid-grade one were also used as a measure of concurrent validity.

Predictive validity was analyzed by correlating the measures from the two school years and by cross-tabulating the two school years' clusters. The predictive validity was very strong. All the correlations were positive and very high, and 94% of the at-risk children in grade two were already identified in the group test in grade one.

The psychometric properties of the group test justify the use of a subset of the seven scales in the group test to create a reliable short version to be used as a screening test. The short version is created and attached. (see Appendices 9 and 10).

Identifying at risk children

K-Means cluster analysis was used in the process of identifying the children at risk of reading and writing difficulties. The analysis resulted in four distinct groups of children based on their mean scores in the different scales in the group test. The groups were named: *at-risk, strugglers, readers,* and *good readers*. These group names reflect the characteristics of the children in relation to their acquisition of the early skills of reading and writing. The *at-risk group*, which is the target group in this study, is characterized by obtaining very low or zero

scores in all of the scales. Conversely, the *good readers* group comprises children scoring highly in all the measures.

In order to set the cut-off point on the group test, a summary variable was created and a score was counted based on the seven scales of the group test. The analysis of the distribution revealed that 24.3% (N= 82) of the children, an amount comparable to the *at-risk* group from cluster analysis, had scores below 166. The suggestion is consequently that children scoring lower than 166 belong to the *at-risk* group and need intensive support. The 75 children (22.3%) scoring between 166 and 368 also belong to a group needing attention and observation. They can somehow manage, but they are struggling. With extra attention and some support, most of them can probably manage well. Children scoring above 368 are doing fine. When using the test, however, notice should be given to the fact that many background factors have an effect on the test results. It should be noted that due to the influence of various factors that affect the test scores, different cut-off points could be set for the risk groups in different kinds of schools (see Jenkins & Johnson, 2007; Jenkins, Hudson & Johnson, 2007).

Gender and age

The effect of gender and age on the group test scores was analyzed. No significant gender differences were found. However, the boys had on average slightly higher means than the girls. This is in accordance with results from recent Tanzanian studies, where similar kinds of literacy skill components have been studied in a low income area of Dar-es-Salaam (Kumburu, 2011) and in a poor rural area in Eastern Tanzania (Ngorosho, 2011). However, the finding is in contrast with results from a large country-wide survey in Tanzania (Uwezo, 2010). An age effect was observed, and the younger children outperformed the older children significantly. This is inconsistent with earlier studies reported in developed countries, where older children usually perform better than younger children; for example, on the Brigance K & 1 screen (Mantzicopoulos, 1999). The children in the studied context seem to be more vulnerable with increasing *age*. A late school start is rather common in the context of the study. The reasons are mainly related to the socioeconomic background.

Environmental determinants

The home factors used in the study were the following: *parents' and guardians' reading ability, the child's schoolwork at home, the number of reading books at home, and the support which the child receives with school work at home.* All these factors had a significant effect on children's performance in the group test. The explanatory power of some of the home and school factors were very strong. *The number of reading books at home* accounted for not less than 38% of the variance on the group test results. The effect of home environment factors on children's literacy skills in Tanzania has also been shown in other recent studies (see Ngorosho, 2011). Parents' education and books for school subjects at home and some other variables (e.g. housing) were identified to be important

descriptors of home environment in rural Tanzania. Father's education and books at home emerged as strong predictors on a variety of literacy skill components. The importance of literacy related home factors for children's literacy development has also been revealed in many international studies (e.g. PIRLS, 2001 reported in NFER, 2003). According to PIRLS, *number of reading books at home* accounted for 10-15% of the variance on reading skills. There seems to be much higher influence of the home environment on early reading in Tanzania than in the countries included in PIRLS. *Parents´ reading ability* was the next strongest factor in the present study, accounting for 15% of the variance. Other important factors were *time for school work at home and receiving support at home*, which explained 10% of the variance.

Of the school factors, having attended nursery explained 16% of the results in group test and 17% of the school marks. The effect is in accordance with earlier findings (e.g. Aboud & Hossain, 2011; Foster & Miller, 2007; Gillian & Ziggler, 2000; Magnuson, Meyers, Ruhm & Waldfogel, 2004; Nonoyama-Tarumi & Bredenberg, 2009; PIRLS, 2005/2006; Sylva, Siraj-Blatchford & Taggart, 2003; Valenti & Tracey, 2009).

On *school status*, the results indicated that the good performance school outperformed the high performance school and the low performance school outperformed the peripheral average performance school. The findings indicated that *school status* had an effect on the achievements in the group test, but the explained variance was only 6%. The differences were only seen between the two highest and the two lowest status schools. Regarding school marks, 14% of the variance was explained. The school with the highest status outperformed the school with the second highest status and the third one in turn outperformed the fourth one.

School attendance explained 21% of the results of the group-test, which could be expected. The reasons for variation in school attendance are many and mostly related to home- and other context factors. There was also a clear effect on school marks. However, only 8.5% of the variance in school marks was explained.

Significance of the study

The successful completion of this study is firstly expected to create awareness among educational stakeholders of the existence of children with learning difficulties, who have not been catered for in Tanzania, specifically children with reading and writing difficulties. Secondly, the study ventures to provide parents, teachers and other interested parties with a reliable and valid instrument for screening students in order to identify those at risk of developing reading and writing difficulties. Hence, a culture of early screening of children in Tanzania may be established. Thirdly, teachers and other educationists will be able to determine the appropriate steps to take or be led to an appropriate remedial instruction program for children with disabilities. Fourthly, this study is

expected to be a way of preparing regular teachers for effective participation and contribution to the focused inclusion set-up in Tanzania. In a long-term focus the study thus implies the promotion of reading and writing among beginning readers.

The psychometric findings suggest that the group test developed can be soundly used as a screening device in identifying beginning readers at risk of developing reading and writing difficulties. This device is especially efficient where Kiswahili language is concerned. However, it can also serve as a basis for developing adaptations in other transparent orthography realms. Further, the questionnaire (described in the methods part) would be very useful in obtaining information about the children's home and school background factors that could affect their acquisition of reading and writing skills. It is anticipated that teachers can easily use the screen to identify children in grade one requiring immediate intervention or further assessment. Performing the test does not take much time and it is relatively easy to administer and score. It can be administered following a short training and guiding instructions. It is also cost effective in the sense that it does not require a lot of staff and expensive material to administer.

Conclusion

With the original group test and the short version presented it is possible to identify *at risk readers, struggling readers* and *good readers*, half way through grade one. Following good reliability and validity, the test is recommended as being useful for teachers, parents and other stakeholders interested in children's reading and writing process and wanting to know who is progressing well and who is facing problems. The created instrument in combination with home- and school background information is also useful for obtaining information about factors which are important for student achievement. It has been apparent in this study that school attendance needs to be emphasized among beginning readers alongside the learning of the basic literacy skills.

Further research is suggested to focus on gaining experience of the created screen (the short version) in predicting different kinds of outcomes in school achievement after grade one, There is also a need for obtaining experiences of different kinds of samples, e.g., in different areas of urban and rural Tanzania. The function of the test could also be studied in other East African countries such as Kenya, Uganda, Ruanda and Burundi with Kiswahili-speaking children.

Sammanfattning

Avhandlingens syfte och kontext

Temat för denna avhandling är identifiering av barn med risk för läs- och skrivsvårigheter. Det specifika syftet var att skapa och validera ett teoribaserat grupptest på swahili för identifiering av barn med risk för läs- och skrivsvårigheter. Målgruppen var elever i årskurs ett i Tanzania. Testet validerades genom jämförelse med resultat från ett individuellt test och från en uppföljning i årskurs två. Testet skapades och dess funktion presenteras på basen av ett urval elever från årskurs ett från Morogoro kommun i Tanzania. Betydelsen av kontextuella faktorer för läs- och skrivfärdigheten, undersökt med det skapade testet, rönte ett speciellt intresse i studien. Vår kunskap om betydelsen av hem- och skolrelaterade faktorer för läs- och skrivutvecklingen härstammar främst från Europa och USA. Forskning om temat från afrikanska länder och över huvudtaget från utvecklingsländerna existerar i mycket begränsad utsträckning.

Många barn i Tanzania går ut grundskolan utan att ha lärt sig läsa och skriva (Chonjo, 1966; Malekela, 2003; Uwezo, 2010). Hög avbrytarprocent, hög skolfrånvaro och det faktum att många barn måste gå om klassen, särskilt i de lägre årskurserna, är stora problem i de tanzanianska grundskolorna. Orsakerna är många och komplexa, men det är tydligt att flera faktorer är relaterade till svag läsförmåga. Det finns också barn med inlärningssvårigheter, t.ex. läs- och skrivsvårigheter, som inte identifierats. Dessa barn har heller inte fått hjälp. Behovet av åtgärder är uppenbart. Läsfärdigheter är viktiga för barnens inlärning inom alla ämnesområden i skolan och även för hela deras framtid. Behovet av stöd har lyfts fram i ett nationellt handikappdokument The National Policy on Disability (MLYDS, 2004, p.4):

Det finns ett behov av tidig identifiering av barn med olika funktionsnedsättningar. Tidig identifiering följd av lämpliga åtgärder kan förhindra uppkomsten av funktionsnedsättningar eller minimera effekterna senare i livet. Tanzania har dock inte något nationellt program för tidigt ingripande som kunde underlätta identifieringen av barn med funktionsnedsättning på kommun- och samfundsnivå. Det faktum att föräldrarna gömmer undan barn med funktionsnedsättningar gör det svårare för dem att få ändamålsenlig hjälp.

Grundläggande färdigheter

Många läsforskare är överens om att brister i det fonologiska systemet är en indikator på läs- och skrivsvårigheter (e.g. Catts & Kahmi, 2004; Hoien & Lundberg, 2000; Kerins, 2006; National Reading Panel, NHI, 2000; Snowling, 2000; Vellutino et al., 2004). De grundläggande fonologiska processerna är centrala för läsutvecklingen. Brister i processerna förefaller leda till lässvårigheter.Dessa processer katergoriseras ofta i tre områden: fonologisk medvetenhet, benämning (rapid naming) och fonologiskt minne.

Fonologisk medvetenhet innefattar medvetenheten om ljudstrukturer i talat språk. En fonologiskt medveten individ kan höra, minnas och manipulera ljudenheter i ord, stavelser och stavelsefragment, dvs. onset rime och fonem (Bailet et al., 2009; McGuiness, 2004). Snabb automatiserad benämning är den tid det tar att nämna saker vid namn, till exempel att säga bokstäver eller stavelser högt så fort som möjligt. Snabb benämning har använts framgångsrikt för att upptäcka brister i fonologisk medvetenhet och den är en stark prediktor för läsfärdighet. Betydelsen av snabb automatiserad benämning identifierades redan för tiotal år sedan av Denkla och hennes kollegor (Denkla & Rudel, 1972). Fonologiskt minne är den tredje processen som behövs vid förvärvandet av läs- och skrivfärdigheter. Med denna process avses kodningen av fonologisk information för temporär lagring i arbetsminnet före det lagras i långtidsminnet (Kerins, 2006; Wagner, Torgesen & Rashotte, 1999). Den s.k. fonologiska loopen associerad med arbetsminnet ger en kort ordagrann lagring av auditiv information (Baddeley, 1986; Baddeley, 1992; Gathercole, Willis & Baddeley, 1991).

Att förvärva läs- och skrivfärdigheter innebär att läsaren kan avkoda flytande och exakt. Fonologisk medvetenhet har undersökts utförligt och visat sig vara en indikator och samverkande faktor vid läsning av oregelbundna språk, engelska i synnerhet, men har också visat sig - i denna avhandling – vara en utslagsgivande indikator för det regelbundna språket swahili. Några författare (t. ex. Alcock, et al., 2000; Aro, 2004; Lyytinen & Erskine, 2006) har föreslagit att bokstavskännedom kunde fungera som enda indikator på läsfärdigheter och vidare att den är en viktig bidragande faktor till läs- och skrivfärdighet, speciellt i regelbundet stavade språk.

Att skapa ett grupptest för screening

Att skapa screeningtest är en krävande uppgift, både teoretiskt och empiriskt. När man skapar screeningtest bör ett trestegsförfarande iakttas (Jenkins & Johnson 2007; Jenkins, Hudson & Johnson, 2007). Det första steget är att definiera vad som skall prediceras, det andra är att identifiera tidiga prediktorer och det tredje att bestämma ett gränsvärde för att skilja åt de barn som bedöms höra till riskgruppen.

I denna studie skapades ett screeningtest avsett att användas som grupptest i årskurs ettAvsikten var att identifiera barn med risk för läs- och skrivsvårigheter. Grupptestet grundar sig på aktuell kunskap om läsning och skrivning och läs- och skrivsvårigheter. Den begreppsmodell som utgjort grunden för konstruktionen av testuppgifterna, baseras på läsning som avkodning + förståelse (the simple view of reading), läsning som avkodning + förståelse + motivation (the componential model of reading) samt modellen Response to Intervention . Enligt den förstnämnda modellen (simple view of reading) är läsning en produkt av avkodning och förståelse. Emellertid är läsning ett komplext fenomen som innefattar förutom avkodning och förståelse även kognitiva, lingvistiska, psykologiska, motivationsrelaterade och miljömässiga faktorer.

Avhandlingens design

Studien är longitudinell. I mitten av juli 2007 deltog 337 barn i årskurs ett, i åldrarna 6-12 år (medeltal 7,25 år), från fyra grundskolor i Morogoro kommun, i ett grupptest. Testet mäter kognitiva lingvistiska färdigheter och innefattar ordkunskap, bokstavskännedom, fonologisk medvetenhet och stavning. Målet var att identifiera barn med risk för läs- och skrivsvårigheter. En månad senare deltog ett mindre urval, 64 barn ur den första undersökningsgruppen, i ett individuellt test som mäter läsning av nonsensord, riktiga ord, läsning av text, läsning på en minut och skrivning. I mitten av juli 2008, när barnen (N=330) var i årskurs två, deltog de i ett uppföljande test som inkluderade ett urval av fyra deltest från testbatteriet i årskurs ett. Det individuella testet, uppföljningstestet och skolbetyget från mitten av årskurs ett användes i valideringen. Inga andra test som kunde användas som kriteriemått i valideringen existerade i Tanzania vid tiden för undersökningen.

De valda skolorna var belägna i staden Morogoro men hade elever både från stadskärnan och från den omgivande rurala miljön. Förutom årskurs och bostadsområde ansågs skolstatus vara en central grundläggande variabel för valet av undersökningsgrupper. Skolorna hade klassificerats i fyra kategorier: högpresterande skolor, skolor som presterade över genomsnittet, medelpresterande och lågpresterande skolor. Syftet med att använda skolstatus som grund för urvalet var att säkra en spridning som representerar befolkningen. Skolstatusen baserades på skolinspektörernas rapporter. Beskrivningen av undersökningspersonerna omfattar förutom ålder och kön också vissa hem- och skolrelaterade faktorer som bedömdes vara av betydelse för läs- och skrivinlärning (se Johnson, Pool & Carter, 2009). De valda hemrelaterade faktorerna inkluderade föräldrarnas läsförmåga, antalet böcker i hemmet och hemmets stöd i skolarbetet. De skolrelaterade faktorerna inkluderade daghemserfarenhet, skolstatus och skolnärvaro. Information om hem- och skolfaktorer samlades in genom ett formulär som fylldes i av föräldrar och lärare.

Grupptestets psykometriska egenskaper

Det deskriptiva datamaterialet visar på stor variation i alla test. Grupptestet har de psykometriska egenskaper som krävs för ett screeningstest. I valideringsprocessen användes deltesten från det individuella testet, uppföljningstestet samt skolbetyget som kriteriemått.

Begreppsvaliditet undersöktes genom item-analys och faktoranalys. Item-analysen genomfördes för att bestämma den interna konsistensen i grupptestet och uppföljningstestet. innehållsvaliditeten säkrades först kvalitativt. Dimensionerna bakom deltesten var teoribaserade och testuppgifterna baserades på högfrekventa ord från godkända läro- och textböcker för årskurs ett. Svårighetsgraden (p-värde) och diskriminationsförmågan (D-värde) för varje testuppgift uträknades. De flesta uppgifterna hade en acceptabel svårighetsgrad (p-värdet var mellan 0,15 och 0,85). Uppgifterna diskriminerade också väl. D-värdet var från 0,30 och uppåt. De flesta uppgifterna bedömdes kunna användas i fortsatt utveckling av screeninginstrumentet.

Begreppsvaliditeten undersöktes vidare med hjälp av principalkomponentanalys. Strukturen var endimensionell och den faktor som identifierades benämndes tidig läs och skrivfärdighet. Alla deltest mätte alltså samma underliggande fenomen. Alla sju deltest korrelerade högt.

Screeningtestets reliabilitetskoefficienter var höga (Crohnbach's alpha, r = .76 - .95), vilket kunde förväntas på basen av de höga korrelationerna, den

endimensionella strukturen och den ovan beskrivna goda innehållsvaliditeten. Den höga reliabiliteten visar också att testet var omsorgsfullt konstruerat och utfört.

I valideringen av screeningstestet användes det individuella testet som det första viktiga kriteriemåttet. Det individuella testet innehöll sådana lingvistiska komponenter, som i tidigare forskning visat sig vara sensitiva prediktorer för läs- och skrivfärdigheter. Följande komponenter som mäter läsflyt och skrivning ingick i testet: läsning av nonsensord (Baker, Park & Baker, 2010; Good, Baker & Peyton, 2009), högläsning av ord och läsning av faktisk text (Adams, 1990; National Reading Panel, 2000a; Snow et al, 1998; Stanovich, 2000), en minuts läsning (Brus & Voeten, 1973), och skrivning (Moats, 2006). Eftersom det individuella testet genomfördes tidsmässigt ganska nära screeningen (grupptestet) reflekterade det samtidig validitet. Skolbetyget från medlet av årskurs ett var ett annat mått på samtidig validitet.

Prediktiv validitet analyserades genom att korrelera mätningarna från de två skolåren och genom korstabulering av klustren från årskurs ett och två. Den prediktiva validiteten var mycket stark. Alla korrelationer var positiva och väldigt höga. Hela 94 % av de barn som klassificerades som ”riskbarn” i årskurs två identifierades med hjälp av grupptestet i screeningen redan i årskurs ett.

De beskrivna psykometriska egenskaperna av grupptestet möjliggör användningen av en kortversion av testet. Den korta versionen har skapats och analyserats och är bifogad som Appendix 7.

Att identifiera barn i riskzonen

K-Means klusteranalys användes i processen att identifiera barn i riskzonen för läs- och skrivsvårigheter. Analysen resulterade i fyra distinkta grupper av barn, baserat på resultaten i grupptestet. Grupperna benämndes enligt följande: *barn i riskzonen, kämpare, läsare och goda läsare*. Gruppen barn i riskzonen, som är målgrupp i denna avhandling, karaktäriseras av att de erhöll väldigt låga eller noll poäng i alla deltest. De goda läsarna fick höga poäng i alla mätningar.

För att bestämma gränsvärdet på grupptestet skapades en summavariabel av de sju deltesten i grupptestet. Utgående från fördelningen valdes 24,3 procent (N=82) av barnen (alla som hade en poängsumma under 166) till riskgruppen motsvarande resultatet från klusteranalysen. Barnen i riskzonen anses behöva omfattande stöd. De 75 barn (22,3%) som fick mellan 166 och 368 poäng hör

också till en grupp som behöver uppmärksamhet och observation. De kan klara sig någorlunda, men de får kämpa. Med extra uppmärksamhet och stöd kan de flesta av dem klara sig bra. Barn som får minst 368 poäng klarar sig fint. När man använder testet bör man dock notera att många bakgrundsfaktorer har effekt på testresultaten. Man bör också notera att eftersom olika faktorer inverkar på testsiffrorna så kan man överväga att använda olika gränsvärden för riskgrupperna i olika slag av skolor (se Jenkins & Johnson, 2007; Jenkins, Hudson & Johnson, 2007).

Kön och ålder

Effekten av kön och ålder analyserades. Inga signifikanta könsskillnader noterades. Dock hade pojkarna i snitt en aning högre medeltal än flickorna. Detta stämmer överens med resultaten från några aktuella tanzanianska studier, där liknande läs- och skrivfärdighetskomponenter har studerats i ett låginkomstområde i Dar-es-Salaam (Kumburu, 2011) och i ett fattigt ruralt område i östra Tanzania (Ngorosho, 2011). Motsatta resultat erhölls dock i en stor landsomfattande undersökning i Tanzania nyligen (Uwezo, 2010).

En ålderseffekt observerades. De yngre barnen presterade märkbart bättre än de äldre barnen, vilket är oförenligt med forskningsresultat från industrialiserade länder, där äldre barn vanligen presterar bättre än yngre (se t.ex. Mantzicopoulos, 1999). De äldre barnen i den studerade kontexten verkar vara mer sårbara än de yngre. En sen skolstart är ganska vanligt i den miljö där studien genomfördes. Orsakerna är huvudsakligen relaterade till den sociokonomiska bakgrunden.

Miljörelaterade determinanter

De hemrelaterade faktorer som användes i avhandlingen var följande: föräldrarnas eller vårdnadshavarnas läsfärdigheter, barnets skolarbete hemma, antalet skolböcker hemma, och stödet för skolarbetet som barnet erhöll hemma. Alla dessa faktorer hade signifikant effekt på barnens prestation i grupptestet. Förklaringsgraden för en del av de hemrelaterade och skolrelaterade faktorerna var väldigt stark. Antalet skolböcker hemma stod för inte mindre än 38 % av variationen i grupptestresultatet. Effekten av hemmiljöfaktorer på barns läs- och skrivfärdigheter i Tanzania har också påvisats i andra nyutkomna studier (se Ngorosho, 2011). I Ngoroshos studie identifierades föräldrarnas utbildning och antalet skolböcker i hemmet samt några andra variabler (t. ex. husets byggnadsmaterial) som viktiga beskrivande faktorer för hemmiljön i rurala Tanzania. Pappans utbildning och tillgången på böcker hemma var starka

prediktorer för flera läs- och skrivfärdighetskomponenter. Betydelsen av hemrelaterade faktorer för läs- och skrivfärdigheten har också framkommit i många internationella studier (t. ex. PIRLS, 2001 rapporterad i NFER, 2003). Enligt PIRLS har antalet textböcker i hemmet stått för 10-15 % av variansen på läsfärdighetsmätningarna. Hemmiljöns inverkan förefaller vara mycket större på tidig läsinlärning i Tanzania än i länderna som medverkat i PIRLS. Föräldrarnas läsfärdigheter var den näststarkaste faktorn i denhär avhandlingen och den stod för 15 % av variansen. Andra viktiga faktorer var tid för läxläsning/skolarbete hemma och stödet i hemmet, vilka förklarade 10 % av variansen.

De skolrelaterade faktorerna var också av stor betydelse. Daghemsvistelse stod för 16 % av variansen i grupptestet och 17 % av variansen på skolvitsorden. Betydelsen av att ha varit på daghem har också konstaterats i tidigare forskning. (t. ex. Aboud & Hossain, 2011; Foster & Miller, 2007; Gillian & Ziggler, 2000; Magnuson, Meyers, Ruhm & Waldfogel, 2004; Nonoyama-Tarumi & Bredenberg, 2009; PIRLS, 2005/2006; Sylva, Siraj-Blatchford & Taggart, 2003; Valenti & Tracey, 2009).

Skolstatus hade signifikant effekt på resultatet i grupptestet, men den förklarade endast 6 % av variansen. Skillnaderna sågs endast mellan de två skolorna med högst och lägst status. Det noterades att leverna i skolan som presterade över genomsnittet hade bättre resultat än eleverna i den högst presterande skolan och att eleverna i den lägst presterande skolan hade bättre resultat än eleverna i den medelpresterande skolan. Effekten av skolstatus på skolbetygen var tydlig med en förklaring av variansen på 14 %. Rangordningen följde rangordningen av skolstatus.

Skolnärvaro förklarade 21 % av resultaten i grupptestet, vilket var väntat. Orsakerna till variationen i skolnärvaro är många och mestadels kopplade till hem- och andra kontextrelaterade faktorer. Effekten på skolbetygen var också signifikant, dock var förklaringsgraden inte så hög (8,5 %).

Avhandlingens betydelse

Denna avhandling förväntas för det första skapa medvetenhet hos olika parter inom utbildningsväsendet om att det finns barn med inlärningssvårigheter och att dessa inte uppmärksammats i Tanzania. Speciellt gäller detta barn med läs- och skrivsvårigheter. För det andra strävar avhandlingen till att förse föräldrar, lärare och andra berörda parter med ett tillförlitligt screeninginstrument med vars hjälp det är möjligt identifiera barn med risk för läs- och skrivsvårigheter. Förhoppningen är att en screeningkultur utvecklas i Tanzania. För det tredje kan

studien bidra till att lärare och andra representanter för utbildningsväsendet får en grund för bedömning av vilka åtgärder som bör vidtas och till att finna lämpliga stödjande undervisningsprogram för barn med funktionsnedsättning. För det fjärde väntas studien kunna bidra till att förbereda allmänundervisningens lärare för ett effektivt och fokuserat utvecklingsarbete för inkluderande skolor i Tanzania. På lång sikt förväntas avhandlingen bidra till att höja nivån på läs- och skrivfärdigheten bland nybörjarläsare

Det framtagna grupptestets psykometriska karaktäristika visar att testet lämpar sig väl som screeninginstrument i identifieringen av nybörjarläsare med risk för läs- och skrivsvårigheter. Instrumentet är speciellt avsett för barn som talar swahili. Det kan dock också fungera som bas för att utveckla test på andra språk med regelbunden ortografi. Det använda frågeformuläret föreslås också kunna användas för att inhämta information om sådana faktorer i hemmet och skolan som är av betydelse för läs- och skrivutvecklingen. Lärare förväntas lätt kunna använda screeninginstrumentet för att identifiera barn i årskurs ett som behöver omedelbar hjälp eller vidare utvärdering. Att genomföra testet och analysera resultaten är relativt lätt och tar inte lång tid. Testet kan användas efter kort skolning och med vägledande instruktioner. Testet är också kostnadseffektivt. Det förutsätter varken mycket personal eller kostsamma material.

Konklusion

Med det ursprungliga grupptestet och den korta versionen som presenterats är det möjligt att identifiera läsare i riskzonen, läsare som kämpar och goda läsare, halvvägs genom årskurs ett. Testet har god reliabilitet och validitet och föreslås kunna användas av lärare, föräldrar och andra som är intresserade av barns läs- och skrivprocesser, och som vill veta vilka som klarar sig bra och vilka som stöter på problem. Det skapade instrumentet i kombination med information om hem- och skolbakgrund är också användbart då man behöver få information om faktorer som är viktiga för studieprestationer. Det har tydligt framkommit i denhär avhandlingen att nybörjarläsares skolnärvaro bör betonas jämsides med inlärningen av baskunskaper i läsning och skrivning.

Ytterligare forskning behövs för att samla erfarenhet av hur väl det skapade screeningtestet (den korta versionen) fungerar när man önskar predicera skolframgång efter årskurs ett. Erfarenhet av testet på olika elevsampel behövs också, till exempel i olika områden av urbana och rurala Tanzania. Vidare skulle det vara av intresse att studera hur testet fungerar på swahilitalande barn i andra östafrikanska länder såsom Kenya, Uganda, Ruanda och Burundi.

References

Aaron, P. G., Joshi, R. M., Gooden, R. & Bentum, K. E. (2008). Diagnosis and treatment of reading disabilities based on the component model of reading. An alternative to the discrepancy model of LD. Journal of Learning disabilities, 41(1), 67-84.

Aboud, F. E. & Hossain, K. (2011). The impact of preprimary on primary school achievement in Bangladesh. *Early Childhood Research Quarterly*, 26(2), 237-246.

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Aiken, L. R. & Groth-Marnat, G. (2006). *Psychological testing and assessment* (12th ed.). New York: Allyn & Bacon.

Ainscow, M. & Lewis, I. (2005). Inclusive networking: The example of the enabling education network (EENET). A paper presented at the Inclusive and Surppotive Education Conference (ISEC), 1st-4th August, 2005. Glasgow: Scotland.

Alcock, K. J., Nokes, K., Ngowi, F., Musabi, C., Mbise, A., Mandali, R., Bundy, D. & Baddeley, A. (2000). The development of reading tests for use in a regularly spelled language. *Applied Psycholinguistics*, 21, 525-555.

Alexander, K. L., Entwistle, D. R. & Horsey, C. S. (1997). From first grade forward: Early foundations of high school dropout. *Sociology of Education*, 70, 87-107.

Anastasi, A. (1982*). Psychological testing* (5th ed.). New York, NY: Macmillan.

Anastasi, A. & Urbina, S. (1997). Psychological issues in ability testing. In A. Anastasi & S. Urbina, (Eds.), Psychological *testing* (7th ed., pp. 331-346). New Jersey: Prentice Hall.

Anthony, J. L**.** & Francis, D. (2005). Development of phonological awareness. *Current Directions in Psychological Science*, 14, 255-259.

Armbruster, B. B., Lehr, F. & Osborn, J. (2001). *Put reading first: The research building blocks for teaching children to read: Kindergarten through grade 3.* Jessup: National Institute for Literacy.

Arnold, C. L., Davis, T. C., Frempong, J. O., Humiston, S. G., Bocchini, A., Kennen, E. M. & Lloyd-Puryear, M. (2006). Assessment of newborn screening parent education materials. *Pediatrics*, 5, 320-325.

Aro, M. (2004). *Learning to read. The effect of orthography*. Academic dissertation. Jyväskylä Studies in Education, Psychology and Social Research 237, University of Jyväskylä.

Baddeley, A. D. (1986). *Working memory*. London: Oxford University Press.

Baddeley, A. D. (1992). Working memory. *Science*, 255, 556-559.

Bailet, L. L., Repper, K. K., Piasta, S. B. & Murphy, S. P. (2009). Emergent literacy intervention for prekindergarteners at risk for reading failure. *Journal of Learning Disabilities*, 42(4), 336-355.

Baker, S. K., Gersten, R., Haager, D. & Dingle, M. (2006). Teaching practice and the reading growth of first-grade English learners: Validation of an observation instrument. *Elementary School Journal*, 107(2), 199-220.

Baker, S. K. & Baker, D. L. (2008). English learners and response to intervention. In E. L. Grigorenko (Ed.), *Educating individuals with disabilities*: IDEA 2004 and beyond (pp. 249-273). New York: Springer.

Baker, D. F., Park, Y. & Baker, S. K. (2010). Efffect of initial status and growth in pseudoword reading on Spanish reading comprehension at the end of first grade. *Psicothema*, 22(4), 955-962.

Bartlett, M. S. (1954), "A note on multiplying factors for various chi-squared approximations", *Journal of the Royal Statistical Society, Series B*, vol. 16, 296-298.

Bast, J. & Reitsma, P. (1998). Analyzing the development of individual differences in terms of Matthew effects in reading: Results from a Dutch longitudinal study. *Developmental Psychology*, 34, 1373-1399.

Basic Education Development Committee-BEDC (2001) Educational sector development program: Primary education development plan (2002-2006). Government of Tanzania.

Beech, J. R. & Singleton, C. H. (1997). *Psychological assessment of reading*. London: Routledge.

Beers, C. S. & Beers, J .W. (1980). Early identification of learning disabilities: Facts and fallacies. *The Elementary School Journal*, 81, 67-76.

Beery, A., Barrett, T. C. & Powell, R. W. (1969). *Elementary reading instruction: Selected materials*. Boston: Allyn and Bacon Inc.

Berninger, V. W., Stage, S. A., Smith, D. R. & Hildebrand, D. (2001). Assessment for reading and writing intervention: A three-tier model for prevention and remediation. In J. W. Andrews, D. H. Saklofske, & H. L. Janzen (Eds.): *Handbook of psychoeducational assessment: Ability, achievement, and behaviour in children* [A volume in the educational psychology series] (pp. 195-223). San Diego, CA: Academic Press.

Berninger, V. W., Abbott, R. D., Vermeulen, K., Ogier, S., Brooksher, R., & Zook, D. (2002). Comparison of faster and slower responders to early intervention in reading: Differentiating features of their language profiles. *Learning Disability Quarterly*, 25, 59-76.

BEST (2008). Basic Education Statistics in Tanzania. Ministry of Education and Vocational Training.

Bhalalusesa, E.P. (2003). Education for all: Is Tanzania on track? NORRAG.

Bishop, A. G. (2003). Prediction of first-grade reading achievement: A comparison of fall and winter kindergarten screenings. *Learning Disability Quarterly*, 26, 189–200.

Bishop, A. G. & League, M. B. (2006). Identifying a multivariate screening model to predict reading difficulties at the onset of kindergarten: A longitudinal analysis. *Learning Disabilities Quarterly, 29*, 235–252.

Blachman, B. A. (1997). *Foundations of reading acquisition and dyslexia, implications for early intervention*. Mahwah, NJ: Erlbaum.

Blackman, J., Bruck, M., Herbert, M. & Seiderberg, M. (1984). Acquisition and use of spelling-sound correspondences in reading. *Journal of Experimental Child Psychology*, 38, 114-133.

Bloomfield, L. & Barnhart, C. L. (1961). *Let's read: A linguistic approach*. Detroit, IM: Wayne State. University Press.

Bowers, P. G. (1995). Tracing symbol naming speeds unique contributions to reading disabilities overtime. *Reading and writing: An Interdisciplinary Journal*, 7, 189-216.

Bradley, L., & Bryant, P .E. (1983). Categorising sounds and learning to read: A causal explanation. *Nature,* 301,419-421.

Bradley, R., Danielson, L. & Doolittle, J. (2005). Response to intervention: 1997. *Journal of Learning Disabilities*, 38, 8-13.

Bryant, P. E. & Bradley, L. (1985). *Children's reading problems*. Oxford, UK: Blackwell.

British Dyslexia Association (1968). Definition of dyslexia. BDA.

Brock-Utne, B. (2001) Education for all - in who's language? *Oxford Review of Education*, 1, 115-134.

Bruce, D. (1964). An analysis of word sounds by young children. *British Journal of Educational Psychology*, 34, 170.

Brus, B. T. & Voeten, M. J. M. (1973). *Een minuut test. (One minute test)*. Vorm A en B. Nijmegen: Berkhout.

Burrow, L., Cabell, S., Konold, T., & Invernizzi, M. (2010, April). Using cluster analysis to explain heterogeneity in Hispanic English language learners at kindergarten entry. Invited poster session at the International Reading Association Conference, Chicago, IL.

Caldas, S. J. (1993). Reexamination of input and process factor effects on public school achievement. *Journal of Educational Research*, 86, 206-214.

Carrillo, M. (1994). Development of phonological awareness and reading acquisition: A study in Spanish language. *Reading and Writing*, 6, 279–298.

Casey, A. & Howe, R. (2002). Best practices in early literacy skills. In A. Thomas & J. Grimes (Eds.), *Best Practices in Psychology IV*, (pp. 721-735). Bethesda, MD: NASP.

Castles, A. & Colheart, M. (1993).Varieties of developmental dyslexia. *Cognition*, 47, 149-180.

Catts, H. & Kahmi, A. (2004). *Language and reading disabilities*. Boston: Allyn and Bacon.

Catts, H. W. (2006). School-wide screening. Presentation at the National SEA Conference on Responsiveness to Intervention: Integrating RTI within the SLD Determination Process, Kansas City, MO, 19-21 April 2006.

Cavanaugh, C. L., Kim, A. H., Wanzek, J. & Vaughn, S. (2004). Kindergarten reading interventions for at-risk students: Twenty years of research. *Learning Disabilities Journal: A Contemporary Journal, 2*(1), 9–21.

Chapman, J. W. (1992). *Understanding learning difficulties: A guide for teachers and parents*. (Australian edition). Palmerston North: ERDC Press.

Children's Book Project-Tanzania, CBP (1991). Retrieved 10th June, 2009, from http://www.cbp.or.tz/activities.asp.

Children's Educational Services Inc. (1987). *Test of Reading Fluency*. Minneapolis, MN: Author.

Chonjo, P. N. (1996). Primary education as preparation for life and further education. *Journal of the Faculty of Education*, No 17. Dar es Salaam: Adult Education Press.

Cohen, L. & Manion, L. (1994). *Research methods in education*. (4th ed.). London: Routledge.

Coker, D. (2006). Impact of first grade factors on the growth and outcomes of urban school children's primary-grade writing. *Journal of Educational Psychology*, 3, 471-488.

Compton, D. L., Fuchs, D., Fuchs, L. S. & Bryant, J. D. (2006). Selecting at-risk readers in first grade for early intervention: A two year longitudinal study of decision rules and procedures. *Journal of Educational Psychology*, 98, 394-409.

Compton, D. L., Fuchs, D., & Fuchs, L. S. (2007). *The course of reading and mathematics disability in first grade: Identifying latent class trajectories and early predictors*. Manuscript submitted for publication.

Comrey, A. L., & Lee, H. B. (1992). *A First Course in Factor Analysis* (2nd ed.), Hillsdale, NJ: Lawrence Erlbaum Associates.

Cook, G. B., Cameron, L. D. & Tankersley, M. (2007). Inclusive teachers' attitudinal ratings of their students with disabilities. *The Journal of Special Education*, 4, 230-238.

Cossu, G., Shankweiler, D., Liberman, I., Katz, L., & Tola, G. (1998). Awareness of phonological segments and reading ability in Italian children. *Applied Psycholinguistics*, 9, 1–16.

Coyne, M. D., Kame'enui, E., Simmons, D. C. & Harn, B. A. (2004). Beginning reading intervention as inoculation or insulin: First grade reading performance of strong responders to kindergarten instruction. *Journal of Learning Disabilities*, 37, 90-104.

Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. Fort Worth, TX: Holt, Rinehart & Winston.

Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed., pp. 443-507). Washington, DC: American Council on Education.

Cutting, L. E. & Denkla, M. B. (2001). The relationship of rapid serial naming and word reading in normally developing readers. An exploratory model. *Reading and Writing: An Interdisciplinary Journal*, 14, 673-705.

Dallman, M., Rouch, R. L., Chang, L. Y. C. & Deboer, J. J. (1974). *The teaching of reading*. (4th ed.). New York: Holt, Rinehart & Winston.

Davidson, E. (2006). The progress of the primary education development plan (PEDP) in Tanzania 2002-2004. HakiElimu Working Paper Series. info@hakielimu.org and www.hakielimu.org

Davis, G. N., Lindo, E. J. & Compton, D. (2007). Children at-risk for reading failure: Constructing an early screening measure. *Teaching Exceptional Children, 39* (5), 32–39.

Delgado, C. E. F. & Scott, K. G. (2006). Comparison of referral rates for pre-school children at risk for disability using information obtained from birth certificate records. *The Journal of Special Education*, 1, 28-35.

Denkla, M. B. & Rudel, R.G. (1972). Colour naming defects in dyslexic boys. *Cortex*, 8, 164-176.

Denton, C. A. & Mathes, P. G. (2003). Intervention for struggling readers: Possibilities and challenges. In B. R. Foorman (Ed.), *Preventing and remediating reading difficulties: Bringing science to scale*, (pp. 229-251). Timonium, MD: York Press.

Denton, C. A., Hasbrouck, J. E., Weaver, L. R. & Riccio, C. A. (2000). What do we know about phonological awareness in Spanish? *Reading Psychology, 21*, 335–352.

Denton, C. A., Ciancio, D. J. & Fletcher, J. M. (2006). Validity, reliability and utility of the observation survey of early literacy achievement. *Reading Research Quarterly*, 41, 8-34.

Deshler, Ellis & Lenz (1996). *Teaching adolescents with learning disabilities*. Deuver, CO: Love Publishing Company.

Downey, D. B., von Hippel, P.T. & Broh, B. (2004). Are schools the great equalizer? Cognitive inequality during the summer months and the school year. *American Sociological Review, 69* (5), 613-635.

Downing, J. (1973). *Comparative reading: Cross-national studies of behavior and processes in reading and writing.* New York: The Macmillan Company.

Duffy-Hester, A. M. (1999). Teaching struggling readers in elementary school classrooms: A review of classroom reading programs and principles for instruction. *The Reading Teacher*, 52, 480-495.

Duke, N., Pressley, M., Fingeret, L., Golos, D., Halladay, J., Hilden, K., et al., (2006, July). *Revisiting the simple view of reading*. Paper presented at the annual meeting of the Society for the Scientific Study of Reading,Vancouver, BC, Canada.

Durgunoglu, A., Nagy, W. E. & Hancin-Bhatt, B. J. (1993). Cross-language transfer of phonological awareness. *Journal of Educational Psychology,* 85, 453-465.

Elbers, C., Lanjouw, J.O. & Peter Lanjouw, P. (2003). "Microlevel estimation of poverty and inequality", *Econometrica*, 1, 355-64.

Elbro, C., Boström, I. & Petersen, D. K. (1998). Predicting dyslexia from kindergarten: The importance of distinctness of phonological representations of lexical items. *Reading Research Quarterly*, 33, 36-60.

Elley, W. (Ed.). (1994). *The IEA study of reading-literacy: Achievement and instruction in 32 school systems.* Oxford: Pergamon Press.

Elkonin, D. B. (1973). USSR. In J. Downing (Ed.), *Comparative reading*. (pp. 551-579). New York: McMillan.

Engen, L. & Høien, T. (2002). Phonological skills and reading comprehension: Reading and writing. *An Interdisciplinary Journal*, 15, 613-631.

Felton, R. H. (2001). Students with three types of severe reading disabilities: Introduction to the case studies. *The Journal of Special Education*. 3, 122-133.

Fletcher, J. M., Coulter, W. A., Reschly, D. J. & Vaughn, S. (2004). Alternative approaches to the definition and identification of learning disabilities: Some questions and answers. *Annals of Dyslexia*, 54, 304-331.

Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C. & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology, 90*, 37-55.

Foster, W.A. & Miller, M. (2007). Development of the literacy achievement gap: A longitudinal study of kindergarten through third grade. *Language, Speech & Hearing Services in Schools*, 38(3), 173–181.

Foxcroft, C.D., Patterson, H. Le Roux, D. & Herbs, D. (2004). *Psychological Assessment in South Africa: A Needs Analysis*. Final Report of the Human Sciences Research Council, Pretoria, South Africa.

Frith, U. (1985). Beneath the surface of developmental dyslexia. Are comparisons between developmental and acquired disorders meaningful? In K. E. Patterson, J. C. Marshall & M. Coltheart (Eds.), *Surface dyslexia*: *Neuropsychological and cognitive studies of phonological reading* (pp. 301-330). Hillsdale, NJ: Erlbaum.

Fuchs, L. S., Fuchs, D. & Maxwell, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education, 9*(2), 20-29.

Fuchs, L. S., Fuchs, D., Hosp, M. & Jenkins, J. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading, 5,* 239–256.

Fuchs, L. S., Fuchs, D. & Speece, D. L. (2002). Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disabilities Quarterly*, 25, 33-45.

Fuchs, D., Mock, D., Morgan, P. L. & Young, C. L. (2003). Responsiveness to intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research and Practice*, *18*, 151–171.

Fuchs, D., Fuchs, L. S. & Compton, D. L. (2004). Identifying reading disabilities by responsiveness-to-instruction: Specifying measures and criteria. *Learning Disability Quarterly, 27*, 216–227.

Gakuru, M., Kang'ethe, F. & Ngugi, K. (2004). Some essential features in developing a text to speech system in Kiswahili. Paper presented at the LLSTI workshop held meeting in Lisbon, Portugal, 22 May 2004.

Garson, G. D. (2009). Cluster analysis. Statnotes from North Carolina State University, Public Administration. Retrieved on 31st October, 2009, from http://faculty.chass.ncsu.edu/garson/PA765/cluster.htm#2step Gathercole, S. E., Willis, C. S. & Baddeley, A. D. (1991). Differentiating phonological memory and awareness of rhyme. Reading and vocabulary development in children. *British Journal of Psychology*, 82, 387-406.

Gersten, R. & Santoro, L. E. (2006). Advances in research on teaching students who experience difficulties in learning: Grappling with the issue of access to the general curriculum. In Lani Florian (Ed), *The SAGE Handbook of Special Education.* University of Aberdeen: Sage Publications Ltd.

Gilliam, W. S. & Zigler, E. F. (2000). A critical meta-analysis of all evaluations of state funded preschool from 1977 to 1998: Implications for policy, service delivery and program evaluations. *Early Childhood Research Quarterly*, 15, 441-473.

Good, R., Baker, S. & Peyton, J. (2009). Making sense of nonsense word fluency: Determining adequate progress in early first grade reading. *Reading & Writing Quarterly, 25*, 33-56.

Good, R. H. & Kaminski, R.A. (2000-2001). Dynamic Indicators of Basic Early Literacy Skills (DIBELS). Retrieved in March, 2001, from http://dibels.oregon. edu

Good, R. H. & Kaminski, R. A. (Eds.). (2002). *Dynamic Indicators of Basic Early Literacy Skills* (6th ed.). Eugene, OR: Institute for the Development of Education Achievement.

Good, R. H., Simmons, D. C. & Kame'enui, E. J. (2001). The importance of decision making utility of a continuum of fluency based indicators of foundational reading skills for third grade high stakes outcomes. *Scientific Studies of Reading*, 5(3), 257-288.

González, J. E. J. & Garcia, C. R. H. (1995). Effects of word linguistic properties on phonological awareness in Spanish children. *Journal of Educational Psychology,* 87(2), 193–201.

Goswami, U. (2003). Phonology, learning to read and dyslexia: Across-linguistic analysis. In V. Csepe (Ed.). *Dyslexia: Different brain, different behavior*, (pp.1-40). New York: Kluwer.

Goswami, U. & Bryant, P. E. (1990). *Phonological skills and learning to read*. Hillsdale, NJ: Erlbaum.

Gough, P. B. & Tunmer, W. E. (1986). Decoding, reading and reading disability. *Remedial and Special Education*, 7(1), 6-10.

Graham, L. & Bailey, J. (2007). Learning disabilities and difficulties. An Australian conspectus? Introduction to the special series. *Journal of Learning Disabilities*. 40(5) 386-391.

Gredler, G. R. (1997). Issues in early childhood screening and assessment. *Psychology in the Schools*, 34, 99–106.

Grigorenko, E. L. (2001). Developmental dyslexia. An update on genes, brains and environments. *Journal of Child Psychology and Psychiatry*, 42(1), 91-125.

Gunning, T. G. (2009). *Assessing and correcting reading and writing difficulties*. Boston: Allyn & Bacon.

Haats, R. A. (2002). *Predictive validity of kindergarten screener for young children with reading difficulties*. A Research Paper. The Graduate School University of Wisconsin-Stout.

Hall, L.A. (2007). Understanding the Silence: Struggling readers discuss decisions about reading expository text. *Journal of Educational Research, 100,* 132-141.

Hallahan, P. D. & Kauffman, J. M. (1994). *Exceptional children* (6th ed.). Boston: Allyn and Bacon.

Hamilton, S. S. & Glascoe, F. P. (2006). Evaluation of children with reading difficulties. *American Family Physician*. Retrieved on 25th April, 2008, from http://www.aafp.org/20061215/2079.html

Harrison, A. G. & Nichols, E. (2005). A validation of the dyslexia adult screening test (DAST) in a post-secondary population. *Journal of Research in Reading* 28 (4), 423–434.

Hinnesbusch, T. J. (1979). "Swahili". In T. Shopen, (ed.). *Languages and their status*, (pp. 209-293). Cambridge, MA: Winthrop Publishers.

Hinnesbusch, T. J. & Mirza, S. M. (1979). *Kiswahili: Msingi wa kusema, kusoma na kuandika*. Lanham, MD: University Press of America.

Høien, T. & Lundberg, I. (1988). Stages of word recognition in early reading development. *Scandinavian Journal of Educational Research, 32*, 163-182.

Høien, T., Lundberg, I., Stanovich, K. E. & Bjaalid, I-K. (1995). Components of phonological awareness. *Reading and Writing*, 7, 171-188.

Høien, T. & Lundberg, I. (2000). *Dyslexia: From theory to intervention*. Dordrecht: Kluwer Academic Publishers.

Høien, T. & Lundberg, I. (2001). Dyslexia and phonology. In: A. J. Fawcett (Ed.), *Dyslexia: Theory and Good Practice.* London: Whurr.

Holopainen, L., Ahonen, T. & Lyytinen, H. (2001). Predicting delay in reading achievement in a highly transparent language. *Journal of Learning Disabilities*, 34(5), 401-414.

Holopainen, L. (2002). *Development in reading and reading related skills: A follow-up study from pre-school to the fourth Grade*. Jyväskylä studies in Education, Psychology and Social Research 200. University of Jyväskylä.

Hoover, W. A. & Gough, P. B. (1990). The simple view of reading: Reading and writing. *An Interdisciplinary Journal*, 2, 127-160.

Horn, W. F. & Packard, T. (1985). Early identification of learning problems: A meta-analysis. *Journal of Educational Psychology*, 77, 597-607.

Hotiu, A. (2006). *The relationship between item difficulty and discrimination indices in multiple-choice tests in a physical science course*. The Degree of Master in Science. Florida Atlantic University.

Huang, F. Ford, K. & Invernizzi, M. (2011). Using early literacy profiles of Hispanic-English language learners to predict later reading achievement. SREE Conference Abstract Template.

Hudson, W. W. (1982). *The clinical measurement package. A field manual*. Homewood, IL: Dorsey Press.

Hurry, J. & Doctor, E. (2007). Assessing literacy in children and adolescents. *Child and Adolescent Mental Health,* 2 (1), 38-45.

International Dyslexia Association Board of Directors (2002). Adopted definition of dyslexia. IDA.

Invernizzi, M., Justice, L. M., Landrum, T. & Booker, K. (2005). Early literacy screening in kindergarten: Widespread implementation in Virginia. *Journal of Literacy Research*, 4, 479-500.

Invernizzi, M., Meier, J. & Juel, C. (2005). *Phonological Awareness Literacy Screening 1-3 Technical Reference*. Charlottesville, VA: University Printing.

Iraki, F. K. (2002). *Lecture pragmatique des morphèmes temporels du Swahili*. Doctoral dissertation, University de Geneve.

Jenkins, J. R. & O'Connor, R. (2002). Early identification and intervention for young children with reading/learning disabilities. In R. Bradley, L. Danielson & D. Hallahan (Eds.), *Identification of learning disabilities* (pp. 99-149). Hillsdale, NJ: Erlbaum.

Jenkins, J. R. & Johnson, E. S. (2007). *Universal screening for reading problems: Why and how should we do this?* Retrieved on 31st October, 2008, from http://www.rtnetwork.org/Essential/Assessment/Universal/ar/ReadingProblems

Jenkins, J. R., Hudson, R. F. & Johnson, E. S. (2007). Screening for service delivery in an RTI framework: Candidate measures. *School Psychology Review*, 36, 560–582.

Jimenez Gonzalez, J. E. (2000). Identifying subtypes among Spanish dyslexic children. Paper presented at the XXVII International Congress of Psychology, Stockholm, 23rd – 28th July, 2000.

Jimenez Gonzalez, J. E. & Hernandez-Valle, (2000). Word identification and reading disorders in the Spanish language. *Journal of Learning Disabilities*, 1, 44-60.

Johnson, E. S., Pool, J. L. & Carter, D. (2009). *Implementing a screening process within an RTI framework: A review of methods and instruments K-12.* National Center on Learning Disabilities RTI Action Network.

Joshi, R. M, & Aaron, P.G. (2000). The Component model of reading: Simple view of reading made a little more complex. *Reading Psychology*, 21, 85-97.

Joshi, R. M. & Aaron, P. G. (2009). Assessment of literacy problems based on the component model of reading. In J. Everett, A. Fawcett, & L. Siegel (Eds.), *The SAGE handbook of dyslexia*, (pp. 268-289). London, UK: Sage Publications.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, 80, 437-447.

Kadege, N. P., Keregero, M., Mlekwa, V. & Mushi, P.A.K. (1992). *Peasants and education. A case study of literacy environment in rural Tanzania*. A Report Commissioned by the Ministry of Education and Culture, Dar es Salaam and Swedish International Development Authority SIDA, Stockholm. Ministry of Education and Culture.

Kahigi, K. K. (2003). *Misingi ya kusoma*. DSM: Mradi wa Vitabu vya Watoto Tanzania.

Kaiser, H. F. (1974). "An index of factorial simplicity". *Psychometrika, 39*, 31-36.

Kalanje, E. S. (2002). *Conception of students with learning disabilities: Regular teachers' views on students with learning disabilities – A Tanzanian perspective*. Unpublished Master's Thesis in Education. Åbo Akademi University, Vasa, Finland.

Kaminski, R. A. & Good, R. H. (1998). Assessing early literacy skills in a problem-solving model: Dynamic indicators of basic early literacy skills. In M. R. Shinn (Ed.), *Advanced application of Curriculum-Based Measurement* (pp. 113-142. New York: Guilford.

Kamwela, A. S. (2000). The primary education programme experience (1993-1996) in improving quality of primary education. In J. C. J. Galabawa, F. E. M. K. Senkoro and A. F. Lwaitama (Eds.), *The quality of education in Tanzania: Issues and experiences.* Institute of Kiswahili research: University of Dar es Salaam.

Kapinga, V. M. (2004). School Improvement and Poverty Alleviation: Challenges in changing school culture. In J.C.J. Galabawa & A. Narman (Eds.), *Education, poverty and inequality*. KAD Associates: Dar es Salaam.

Kentonen, R. & Mulenga, K. (2003). *Basic skills assessment tool: Reading and writing grades 1-2 and children with special needs*. Ministry of Zambia.

Kerins, M. (2006). The effects of systematic reading instruction on three classifications of readers. *Journal of College Reading Association*, 3, 243-260.

Kerka, S. (1998). Adults with learning disabilities. *Eric Digest* No. 189.

Kihampa, M. G. (1997). *Karibu tusome-1 (chapa ya 4)*. Dar es Salaam: Mkuki na Nyota Publishers.

Kirk, S., Gallagher, J. & Anastasiow, N. (2000). *Educating exceptional children*, (9th Ed.). Boston, MA: Houghton Mifflin.

Knizer, J. & Klein, L. G. (2007). *Promoting effective early learning: What every policy maker and educator should know*. National Centre for Children in Poverty. Mailman School of public Health. Columbia University.

Kumburu, S. A. (2010). Promoting reading and writing with children at risk for reading and writing difficulties in grade one in Tanzania. Manuscript.

Kutnick, P., Jules, V. & Layne, A. (1997). *Gender and School Achievement in the Caribbean*. Education Research Paper No. 21, London: Department for International Development.

Kyando, N. M. (2007). Quality education: Do we know what is involved? *Journal of Issues and Practice in Education,* 1, 1-11.

Lamdin, D. J. (1996). Evidence of student attendance as an independent variable in education production functions. *Journal of Educational Research, 89, 155-162.*

Learning Print Associates (2007). *Understanding the No Child Left Behind Act*. Reading (Quick Key 1). Association of Educational Publishers. Retieved from http://www.learningpt.org/Quick Keys/

Lee, L. W. (2008). Development and validation of reading-related assessment battery in Malay for the purpose of dyslexia assessment. *Annals of Dyslexia* 1, 37–57.

Leroy, C. & Symes, B. (2001). Teachers' perspectives on the family backgrounds of children at risk. *McGill Journal of Education,* 1, 45-60.

Lesaux, N. K. & Siegel, L. S. (2003). The development of reading in children who speak English as a second language. *Developmental Psychology*, 6, 1005-1019.

Leung, C., Lindsay, G. & Lo, S. K. (2007). Early identification of primary school students with learning difficulties in Hong Kong: the development of a checklist. *European Journal of Special Needs Education*, 3, 327-339.

Lewis, R. B. & Doorlag, D. H. (1995). *Teaching special students in the mainstream* (4th Ed.). Englewood Cliffs: Prentice - Hall Inc.

Lindeman, J. (1998). Ala-asteen lukutesti. Tekniset tiedot. Åbo Akademis förlag.

Lindsay, G. A. & Wendell, K. (1982). The early identification of educationally 'at risk' children revisited. *Journal of Learning Disabilities*, 15, 212-217.

Lindsay, G. & Desforges, M. (1998). *Baseline Assessment: Practice, problems and possibilities*. London: David Fulton.

Lipson, M. Y. & Wixon, K. K. (1997). *Assessment & instruction of reading and writing disability*. (2nd ed.). New York: Longman.

Locke, J. L., Hodgson, J., Macaruso, P., Roberts, J., Lambretch-Smith, S. & Guttentag, C. (1997). The development of developmental dyslexia. In C. Hulme & M. Snowling (Eds.), *Dyslexia: Biology, cognition and intervention.* London: Whurr.

Lundberg, I., Olofsson, A., & Wall, S. (1980). Reading and spelling skills in the first school years predicted from phonemic awareness skills in kindergarten. *Scandinavian Journal of Psychology,* 21, 159-173.

Lyon, G. R. (July 10, 1997). Report on learning Disabilities Research, Congressional testimony given before the Committee on Education and the Workforce in the U.S. House of Representatives, 10 July 1997.

Lyon, G. R. & Fletcher, J. M. (2001). Early intervention for children at risk for reading failure. *Basic Education*, 46, 12-15.

Lyytinen, H. & Erskine, J. (2006). Early identification and prevention of reading problems. In R. E. Tremblay, R. G. Barr, R. DeV. Peters (Eds.), *Encyclopedia on Early Childhood Development.* Montreal, Quebec: Centre of Excellence for Early Childhood Development, 1-6.

Maugham, B. & Yule, W. (1994). Reading and other learning disabilities. In M. Rutter, E. Taylor & L. Hersov (Eds.), *Child and Adolescent Psychiatry*: Modern Approaches, (3rd ed.). (pp. 647-665). London: Blackwell Scientific Publications.

Marsh, G., Friedman, M., Welch, V. & Desberg, P. (1980). The development of strategies in spelling. In U. Frith (Ed.), *Cognitive strategies in spelling* (pp. 339-353). New York: Academic.

Martin, M., Mullis, I. & Kennedy, A. (Eds.) (2003). *The PIRLS technical report.* Chestnut Hill, MA: Boston College.

McBride-Chang, C. & Manis, F. R. (1996). Structural invariance in the associations of naming speed, phonological awareness, and verbal reasoning in good and poor readers: A test of the double deficit hypothesis. *Reading and writing: An Interdisciplinary Journal*, 8, 323-339.

McCardle, P., Cooper, J.A., Houle, G. R., Karp, N. & Paul-Brown, D. (2001). Emergent and early literacy: Current status and research directions-introduction. *Learning Disabilities Research & Practice*, 16(4), 183-185.

McGuinness, D. (2004). *Early reading instruction: What science really tells us about how to teach reading*. London: A Bradford Book.

MacMaster, K. L., Du, X. & Pétursdóttir, A. (2009). Technical features of curriculum-based measures for beginning writers. *Journal of Learning Disabilities*, 1, 41- 60.

Magnuson, K. A., Meyers, M. K., Ruhm, C. J. & Waldfogel, J. (2004). Inequality in preschool education and school rediness. *American Research Journal*, 4(1), 115-157.

Malekela, G. A. (2000). Quality of secondary eduction in Tanzania. In J. C. J. Galabawa, F. E. M. K. Senkoro and A. F. Lwaitama (Eds.). *The quality of education in Tanzania: Issues and experiences*. Dar es Salaam: University of Dar es Salaam.

Malekela, G. A. & Ndeki, A. S. (2001). Access and equity in basic education in Tanzania. *Journal of the Faculty of Education,* No 21. DSM: Adult Education Press.

Malekela, G.A. (2003). English as a medium of instruction in post-primary education in Tanzania: Is it a fair policy? In B. Brock-Utne, Z. Desai & M. Qorro (Eds.), *Language of instruction in Tanzania and South Africa*. (pp.102-113). Dar Es Salaam: E & D Ltd.

Manis, F. R., Doi, L. M. & Bhadha, B. (2000). Naming speed, phonological awareness and verbal reasoning in good and poor readers: A test of the double deficit hypothesis. *Reading and Writing: An Interdisciplinary Journal*, 8, 323-339.

Marsh, G., Friedman, M., Welch, V. & Desberg, P. A. (1980). The development of strategies in spelling. In U. Frith (Ed.), *Cognitive strategies in spelling* (pp.339-353). New York: Academic.

Maugham, B. & Yule, W. (1994). *Reading and other learning disabilities*. In M. Rutter, E. Taylor & L. Hersov (Eds.), *Child and ddolescent psychiatry: Modern approaches*, (3rd ed., pp. 647-665). London: Blackwell Scientific Publications.

May, D. C. & Kundert, D. K. (1992). Kindergarten screening in New York State: Tests, purposes and recommendations. *Psychology in the Schools*, 29, 35-41.

May, D. C. & Kundert, D. K. (1997). School readiness practices and children at risk: Examining the issues. *Psychology in the schools*, 34, 73-84.

Mayringer, H., & Wimmer, H. (2000). Pseudoname learning by German-speaking children with dyslexia. Evidence for a phonological learning deficit. *Journal of Experimental Child Psychology*, 75, 116-133.

Mboya, M. W. & Possi, M. K. (1996). The Special Child in Tanzania Primary Schools. *Papers in Education and Development,* 17, 52-64.

Mantzicopoulus, P. (1999). Risk Assessment of Head Start Children with the Brigance K&1 Screen: Differential Performance by Sex, Age, and Predictive Accuracy for Early School Achievement and Special Education Placement. *Early Childhood Research Quarterly*, 14(3), 383-408.

Mason, J. M. & Allen, J. (1986). a review of emergent literacy with implications for research and practices in reading. In E. Rothkopf (Ed.), *Review of Research in Education*, 13, 3-47. Washington, DC: American Educational Research Association.

Meisels, S. J., Steele, D. M. & Quinn-Leering, K. (1993). Testing, tracking, and retaining young children: An analysis of research and social policy. In B. Spodek (Ed.), *Handbook of research on the education of young children*. (pp. 279-292). New York: MacMillan.

Meisels, S. J. (1994). *Developmental screening in early childhood: A* guide (4th ed.). Washington, D.C.: National Association for the Education of Young Children.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed.), (pp. 13-103). New York: Mcmillan.

Mkinga, M.G. (2000). *Kusoma na kuandika. Kiswahili darasa la kwanza*. Dar es Salaam: Books Publishers Ltd.

MLYDS (2004). *National policy on disability*. United Republic of Tanzania. DSM.

Mmbaga, D. R. (2002). *The inclusive classroom in Tanzania: Dream or reality?* Stockholm: Stockholm University.

Moats, L. (2006). How spelling supports reading. Retrieved on 05/04/2011 http://www.ldonline.org/article/How_Spelling_Supports_Reading

Ministry of Education and Culture (1995). *Education and training policy*. Dar es Salaam. MOEC.

Ministry of Education and Culture (1995). *Education and training policy*. Dar es Salaam: Adult Education Press.

Ministry of Education and Culture (2001), United Republic of Tanzania, *Education Sector Development Programme: Primary education development plan (2002-2006)*, Basic Education Development Committee, Tanzania.

Minisry of education and Culture (2003). *Basic Statistics in Education* 1999-2003. Dar es Salaam. MoEC

Ministry of Education and Vocational Training (2004-2008). Basic educational statistics in Tanzania, (BEST).United Republic of Tanzania.

Ministry of Education and Vocational Training (2008). *Education sector development program* (ESDP): Education performance report 2007/2008. Education Sector Development Committee. The United Republic of Tanzania. Font size too large

Mohammed, M. A. (2001). *Modern Swahili Grammar*. Nairobi: East African Educational Publishers.

Montogomery, D. (1990). *Special needs in ordinary schools. Children with learning difficulties.* London: Cassell Education Limited.

Morgan, P. L., Fuchs, D., Compton, D. L., Cordray, D. S. & Fuchs, L. S. (2008). Does early reading failure decrease children´s reading motivation? *Journal of Learning Disabilities* 5, 387-404.

Mosha, H. J. (2000). Conceptualizing quality education. In J. C. J. Galabawa, F. E. M. K. Senkoro and A. F. Lwaitama (Eds.), *The quality of education in Tanzania. Issues and experiences*. Institute of Kiswahili Research, University of Dar es Salaam.

Moyle, D. (1968). *The teaching of reading*. (3rd Ed.). London: Ward Lock Educational.

Mswia, R., Lewanga, M., Moshiro, C., Whiting, D., Wolfson, L., Hemed, Y., Alberti, K. G., Kitange, H., Mtasiwa, D. & Setel, P. (2003).Community-based monitoring of safe motherhood in the United Republic of Tanzania. *Bulletin of the World Health Organization*, 2, 87-94.

Munishi, G. K. (2000). Quality of education and the national policy management context in Tanzania. Education quality: Context and conceptualization. In J. C. J. Galabawa, F. E. M. K. Senkoro and A. F. Lwaitama (Eds.), *The quality of education in Tanzania. Issues and experiences*. Institute of Kiswahili Research, University of Dar es Salaam.

Murray, B. (2002). Wanted: politics-free science-based education. *Monitor on Psychology*, 8, 52–54.

Myachina, E. N. (1981). *The Swahili language: A descriptive grammar.* London: Routledge.

National Center for Education Statistics (NCES). (2007). Status and Trends in the Education of Racial and Ethnic Minorities. Washington, DC: U.S. Department of Education.

National Reading Panel (NIH) (2000a). *Teaching children to read: An evidence-based assessment of scientific research literature on reading and its implication for reading instruction.* Washington, DC: National Institute of Child Health and Human Development.

National Research Council. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

NFER. (2003). *Progress in international reading literacy* (PIRLS) *study: Reading all over the world.* Retrieved on 6th April, 2004, from http://www.standards.dfes.gov.uk/literacy/publications/research/63521/

Ngorosho, S. D. (2010). *Literacy skills of Kiswahili speaking children in rural Tanzania. The role of home environment*. Thesis manuscript, Åbo Akademi University, Vasa.

Nicolson, R. I. & Fawcett, A. J. (1998). *The Dyslexia Adult Screening Test*. London: Psychological Corporation.

No Child Left Behind Act of 2001, 20 U.S.C. 70 § 6301 *et seq*. (2002).

Nonoyama-Tarumi, Y. & Bredenberg, K. (2009). Impact of school readiness program interventions on children's learning in Cambodia. *International Journal of Educational Development*, 29, 39-45.

Nunally, J. C., Bernstein, I. H. (1994). *Psychometric theory.* (3rd ed.). New York: McGraw-Hill.

Nurse, D. & Spear, T. T. (1985). *The Swahili: Reconstructing the history and language of an African society*: *800-1500*. Philadelphia: University of Pennsylvania Press.

Nurse, D. & Hinnesbusch, T. J. (1993). *Swahili and Sabaki: A linguistic history*. Berkeley and Los Angeles: University of California.

OECD (2001). *Knowledge and skills for life. First results from PISA 2000*. Paris: OECD Publications.

O'Connor, R. E. & Jenkins, J. R. (1999). The prediction of reading disabilities in kindergarten and first grade. *Scientific Studies of Reading*, 3, 159-197.

O'Connor, R. E., Fulmer, D., Harty, K. R. & Bell, K. M. (2005). Layers of reading intervention in kindergarten through third grade: changes in teaching and student outcomes. *Journal of Learning Disabilities*, 38, 440-455.

Odom, S. L., Brantlinger, E., Gersten, R., Horner, R. H., Thompson, B. & Harris, K. R. (2005). Research in special education: scientific methods and evidence-based practice, *Exceptional Children*, 2, 137–148.

Otieno, N.A.O. (2000). Primary school management and its impact on quality in mainland Tanzania. In J. C. J. Galabawa, F. E. M. K. Senkoro & A. F. Lwaitama (Eds.), *The quality of education in Tanzania. Issues and experiences*. Institute of Kiswahili Research, University of Dar es Salaam.

OXFAM (2007). *Quality education*: Oxfam novib's position. Oxfam Novib.

Pennington, B. F. & Olson, R. K. (2005). Genetics of dyslexia. In M. Snowling, C. Hulme, & M. Seidenberg (Eds.), The science of reading: A handbook (pp. 453-472). Oxford: Blackwell Publishing.

Peterson, N. L. (1987). *Early intervention for handicapped and at risk children*. USA: Love Publishing Company.

PIRLS, (2005/2006). Education counts. Retrieved on 11/04/2011 from http://www.educationcounts.govt.nz/themes/research/pirls/pirls_200506.

Phillips, B. M., Lonigan, C. J. & Wyatt, M. A. (2009). Predictive validity of the Get Ready to Read. *Journal of Learning Disabilities*, 2, 133-147.

Podhajski, B., Mather, N., Nathan, J. & Sammons, J. (2009). Professional development in scientifically based reading instruction. Teacher knowledge and reading outcomes. *Journal of Learning Disabilities*, 5, 403- 417.

Pumphrey, P. D. & Reason, R. (1991). *Specific learning difficulties (Dyslexia): Challenges and responses*. London: Routledge Falmer.

Puolakanaho, A. (2007). *Early prediction of reading-phonological awareness and related language and cognitive skills in children with a familial risk for dyslexia.* Jyväskylä Studies in Education, Psychology and Social Research 317. University of Jyväskylä.

Raforth, M. A. K. (1988). Early identification of learning disabilities using the meeting street school screening tests. *Journal of Learning Disabilities*, 21, 186-188.

Rathvon, N. (2006). Early reading assessment. A practitioner´s handbook. Retrieved on 12th November, 2007, from http://www.Rathvonrathrom.com/era.html

Rayner, K. & Pollatsek, A. (1989). *The psychology of reading*. Englewood Cliffs, NJ: Prentice Hall.

Reading, S. & Van Deuren, D. (2007) Phonemic awareness: When and how much to teach? *Journal of the College Reading association*, 3, 267-286.

Reid, G. & Wearmouth, J. (2009). Identification and assessment of dyslexia and planning for learning. In G. Reid, D. Knight, G. Elbeheri & J. Wearmouth (Eds.), *Dyslexia Handbook*. London: Routledge Falmer.

Rhodes, L. K. & Shanklin, N. L. (1993). *Windows into literacy: Assessing learners K8*. Portsmouth, NH: Heinemann.

Ritchie, K. D. & Speece, D. L. (2004). Early identification of reading disabilities: Current status and new directions. *Assessment for Effective Intervention*, 29(4), 13–24.

Rosinski, R. R., Golinkoff, R. M., Kukish, K. S. (1975). Automatic semantic processing in a picture-word interference task. *Child Development*, 46, 247–253.

Rosner, J. & Simon, D. (1971). The auditory analysis test: An initial report. *Journal of Learning Disabilities,* 4, 40-48.

Roth, M., McCaul, E. & Barnes, K. (1993). Who becomes an "at-risk" student? The predictive value of a kindergarten screening battery. *Exceptional Children*, 59 (4), 348–358.

Rothenberger, A. (2005). Developmental risks and prevention. *Journal of Child Psychology and Psychiatry*, 46, 805-916.

Rubagumya, C. (2000). Language as a determinant of quality. In J. C. J. Galabawa, F. E. M. K. Senkoro and A. F. Lwaitama (Eds.), *The quality of education in Tanzania. Issues and experiences*. Institute of Kiswahili Research, University of Dar es Salaam.

Salvia, J. A. & Ysseldyke, J. E. (2001). *Assessment.* Boston: Houston Mifflin Company.

Sambo, W. A. L. (2001). The role of private tuition in secondary education in Tanzania. *Journal of the Faculty of Education*, 21, 96-110.

Sandberg, A. & Norling, M. (2006). Early intervention for young children in pre-school: General and specific support. Paper presented at NFPFS.

Satz, P. & Fletcher, J. M. (1988). Early identification of learning disabled children. An old problem revisited. *Journal of Consulting and Clinical Psychology*, 56, 824-829.

Savolainen, H. (2006) Education for all: Where are children with disabilities? In H. Savolainen, M. Matero & H. Kokkala (Eds.), *When all means all. Experiences in three African countries with EFA and children with disabilities* (pp. 172–184). Helsinki: Ministry for Foreign Affairs of Finland.

Scarborough, H. S. (1990). Very early language deficits in dyslexic children. *Child development*, 6, 1728-1743.

Scarborough, H. S. (1998). Early identification of children at risk for reading disabilities: Phonological awareness and some other promising predictors. In B. K. Shapiro, P. J. Accardo & A. J. Capute (Eds.), *Specific reading disability: A view of the spectrum* (pp. 75-119). Timonium, MD: York Press.

Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. & Foorman, B. R. (2004). Kindergartern prediction of reading skills: A longitudinal comparative analysis. *Journal of Educational psychology*, 96, 265-282.

Schatschneider, C. (2006). *Reading difficulties: Classification and issues of prediction*. Paper presented at the Pacific Coast Research Conference, San Diego, CA.

ScorePak (2005). *Item analysis report.* Retrieved on 8/3/2005 from http://www.washington.edu/oea/score1.html Seymour, P. H. K. & McGregor, C. J. (1984). 'Developmental dyslexia: A cognitive experimental analysis of phonological, morphemic and visual impairments.' *Cognitive Neuropsychology*, 1, 43-82.

Share, D. L. (2008). On the Anglocenticities of current reading research and practice: the perils of overreliance on outlier orthography. *Psychological Bulletin*, 134(4), 584-615.

Shepard, L. A. (1997). Children not ready to learn? The invalidity of school readiness testing. *Psychology in the School,* 2, 85-97.

Shihabuddin, C. & Mnyampala, M. (1977). *Historia ya Kiswahili*. Nairobi: Oxford University Press.

Siegel, L. (1989). IQ is irrelevant to the definition of learning disabilities. *Journal of Learning Disabilities*, 22, 469-478.

Siegel, L. S. & Ryan, E. B. (1989). Subtypes of developmental dyslexia: The influence of definitional variables. *Reading and writing: An Interdisciplinary Journal*, 1, 257-288.

Siegel L. S. & Lipka, O. (2008) The Definition of Learning Disabilities: Who is the Individual with Learning Disabilities? In G. Reid, A. Fawcett, F. Manis, L. Siegel (Eds.), *The Sage handbook of Dyslexia*. Sage Publications.

Simmons, D., Coyne, M., Kwok, O., McDonagh, S, Harn, B. & Kame'enui, E. (2008). Indexing response to intervention: A longitudinal study of reading risk from kindergarten through third grade. *Journal of Learning Disabilities*, 41, 143-158.

Slavin, R. & Madden, N. A. (1989). What works for students at risk. A research synthesis. *Educational Leadership*, 5, 4-13.

Snow, C. E., Burns, S. & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children.* Washington, D.C: National Academy Press.

Snowling, M. J. (2000). *Dyslexia*. Oxford, England: Blackwell.

Sousa, D. A. (2005). *How the brain learns to read.* Thousand Oaks, California, Corwin Press.

Spache, G. D. (1976). *Diagnosis and correcting reading disabilities*. Boston: Allyn and Bacon, Inc.

Sparks, R. L. & Lovett, B. J. (2009). Objective criteria for classification of postsecondary students as learning disabled. Effects on prevalence rates and group characteristics. *Journal of Learning disabilities*, 42 (3), 230-239.

Speece, D. L., & Case, L. P. (2001). Classification in context: An alternative approach to identifying early reading disability. *Journal of Educational Psychology, 93,* 735-749.

Speece, D., Mills, C., Ritchey, K. & Hillman, E. (2003) Initial evidence that letter fluency tasks are valid indicators of early reading skill. *Journal of Special Education, 36,* 223-233.

Springer, D. W., Abell, N. & Hudson, W. W. (2002). Creating and validating rapid assessment instruments for practice and research: Part 1. Retrieved on 4th May, 2008, from Sage Publications, http://rsw.sagepub.com

Sprugevica, I. & Hoien, T. (2003). Early phonological skills as a predictor of reading acquisition: A follow-up study from kindergarten to the middle of grade 2. *Scandinavian Journal of Psychology, 44*, 119-124.

SPSS Inc., (2001). SPSS for Windows, Rel. 11.0.1. Chicago: SPSS Inc.

Stanovich, K. E. (1986). Mathew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21, 360-407.

Stanovich, K. E. & Siegel, L. S. (1994). The phenotypic performance profile of reading-disabled children: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology*, 86, 24-53.

Stanovich, K.E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York: Guilford Publications, Inc.

StatSoft. (2008). *Reliability and item analysis*. Retrieved on 26 April 2009 from http://www.statsoft.com/textbook/reliability-and-item-analysis/

Stauffer, R. G. (1969). *Teaching reading as a thinking process.* New York: Harper & Row Publishers.

Stuart, M. & Coltheart, M. (1988). Does reading develop in a sequence of stages? *Cognition*, 30, 139-181.

Sumra, S. & Rajani, R. (2006). Secondary education in Tanzania: Key policy challenges. Hakielimu Working Paper. info@hakielimu.org and www.hakielimu.org.

Sylva, K., Siraj-Blatchford, I. & Taggart, B**.** (2003). Assessing quality in the early years. Trentham Books.

Taggart, B., Sammons, P., Smees, R., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Elliot, K. & Lunt, I. (2006). Early identification of special education needs and the definition of at risk: The early years transition and special educational needs (EYTSEN) project. *British Journal of Special Education*, 1, 40 - 45.

Tanzania Institute of Education (TIE) (2000). *Kusoma kitabu cha kwanza (Darasa la kwanza)*. (6th ed.). Dar es Salaam: Tanzania Publishing House.

Thorndike, E. L. (1917). Reading as reasoning. *Journal of Educational Psychology,* 8, 323-332.

Tinker, M. A. & McCullough, C. M. (1975). *Teaching elementary reading*. (4th ed.). Englewood Cliffs, NJ: Prentice Hall. (Original work published 1952)

Tonini, D. C. (2009). "The Secondary Education Development Plan in Tanzania: International influences, policy choices and impacts" *Paper presented at the annual meeting of the 53rd Annual Conference of the Comparative and International Education Society, Francis Marion Hotel, Charleston, South Carolina.*

Torgesen, J. K. (2000). Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. *Learning Disabilities Research & Practice*, 15, 55-64.

Torgesen, J. K. (2002). The prevention of reading difficulties. *Journal of School Psychology*, 40, 7-26.

Torgesen, J. K., Morgan, S.T. & Davis, C. (1992). Effects of phonological awareness training on word learning in kindergarten children. *Journal of Educational Psychology*, 84, 364-370.

Torgesen, J. K. & Davis, C. (1996). Individual difference variables that predict response to training in phonological awareness. *Journal of Child Experimental Psychology*, 1, 1-21.

Torgesen, J. K., Wagner, R. K. & Rashotte, C. A. (1999). *Test of word reading efficiency*. Austin, TX: Pro-Ed.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Rose, E., Lindamood, P., Conway, T. & Garvin, C. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology*, 91, 1-15.

Tzivinikou, S. (2004). Strengths and weaknesses of screening reading difficulties, *Electronic Journal of Research in Educational Psychology*, 2, 59-74.

UNICEF, (2009). *Child-friendly schools: Emerging practices in Eastern and Southern Africa*. A human rights-based approach. Retrieved on 2/04/2011 from www.unicef.org

United Republic of Tanzania URT (2008).*The development of education*. National Report of the United Republic of Tanzania. Ministry of Education and Vocational Training Mainland and Ministry of Education and Vocational Training Zanzibar.

United States Department of Education U.S.D.E. (2001). *How to help every child become a reader*. Indianapolis, IN: Your Domain Publishing.

Urevbu, A. (1991). *Curriculum studies*. Singapore: Longman.

Uwezo, (2010). Are our children learning? Annual learning assessment report Tanzania 2010.Uwezo Tanzania (www.uwezo.net).

Valenti, J. E. & Tracey, D. H. (2009). Full-day, half-day, and no preschool effects on urban children's first-grade reading achievement. *Education and Urban Society* (online first: doi: 10.1177/0013124509336060).

Vaughn, S., Fletcher, J. M., Francis, D. J., Denton, C. A., Wanzek, J., Wexler, J., Cirino, P. T., Barth, A. E. & Romain, M. A. (2008). Response to intervention with older students with reading difficulties. *Learning and Individual Differences*, 18, 338-345.

Vellutino, F. R., Scanlon, D. M. & Lyon, G. R. (2000). Differentiating between difficult to remediate and remediated poor readers: More evidence against the IQ-achievement discrepancy definition of reading disability. *Journal of Learning Disabilities*, 33, 223-239.

Vellutino, F. R., Fletcher, J. M., Snowling, M. J. & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades. *Journal of Child psychology and Psychiatry*, 45, 2-40.

Vellutino, F. R., Scanlon, D. M., Small, S. & Fanuele, D. P. (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: Evidence for the role of kindergarten and first grade interventions. *Journal of Learning Disabilities*, 39, 157-169.

Vervaeke, S. L., McNamara, J. K. & Scissons, M. (2007). Kindergarten screening for learning disabilities. *Journal of Applied Research on Learning*, 1(1), 1-13.

Wagner, R., Torgesen, J. & Rashotte, C. (1994). The development of reading-related phonological processing abilities: New evidence of bi-directional causality from a latent variable longitudinal study. *Development Psychology*, 30, 73-87.

Wagner, R. K., Torgesen, J. K. & Rashotte, C. A. (1999). *CTOPP: Comprehensive test of phonological*

Walberg, H. J. & Tsai, S. L. (1983). Mathew effects in education. *American Educational Research Journal*, 20, 359-373.

Wald, B. (1987). "Swahili and the Bantu languages". In B. Comrie, (Ed), *The World´s Major Languages*, (pp. 991-1014). New York: Oxford University Press.

Wenner, G. (1995). Kindergarten screens as tools for the early identification of children at risk for remediation or grade retention. *Psychology in the Schools*, 32, 249-254.

Whitehurst, G. J. (2003). The NCLD Get ready to read! Screening tool technical report. Retrieved on 14th May, 2008, from National Centre for Learning Disabilities. http://www.kb.se/

Whiteley, W. (1969). *Swahili: The rise of a national language*. London: Methuen.

Wimmer, H. & Goswami, U. (1994). The influence of orthographic consistency on reading development: Word recognition in English and German children. *Cognition,* 51, 91-103.

Wimmer, H., Mayringer, H. & Landerl, K. (1998) Poor reading: A deficit in skill-automation or a phonological deficit? *Scientific Studies of Reading*, 4, 321-340.

Winskel, H. & Widjaja, V. (2007). Phonological awareness, letter knowledge and literacy development in Indonesian beginner readers and spellers. *Applied Psycholinguistics,* 28(1), 21-43.

Wolf, M. & Bowers, P. G. (2000). Naming-Speed processes and developmental reading disabilities: An introduction to the special issue on the double-deficit hypothesis. *Journal of Learning Disabilities,* 4, 322-

Woodcock, R.W., Mather, N. & Schrank, F. A. (2004). Woodcock-Johnson III Diagnostic Reading Battery. Itasca, IL: Riverside Publishing.

World Federation of Neurology (1968). Definition of dyslexia. WFN.

Wren, S. (2004). Descriptions of early reading assessments. Southwest Educational Development Laboratory. Retrieved on November 16th, 2004, from http://www.balancedreading.com/ assessment/assessment.pdf

Yopp, H. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly*, 23, 159-177.

Zabrucky, K. & Ratner, H. H. (1992). Effect of passage type on comprehensive monitoring and recall in good and poor readers. *Journal of Reading Behavior*, 24, 373-391.

Zeffiro, T. & Eden, G. (2000). The neural basis of developmental dyslexia. *Annals of Dyslexia*, 50, 3-30.

Ziegler, J. C. & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, 131, 2-29.

Ziolkawska, R. (2007). Early intervention for students with reading and writing difficulties. *Reading Improvement,* 44(2), 76-86.

Appendix 1

KISWAHILI PHONESET

Vowels

Grapheme	Phoneme	As in
a	/a/	b**a**ba (father); between c***a***t and c***a***r
e	/e/	p**e**mbe (horn); p***e***n
i	/i/; /I/	mt**i** (tree); End of happ***y***
o	/o/	tan**o** (five); b***o***x
u	/u/	nd**u**gu/d**u**gu (family relative); fl***u***me

Consonants

Grapheme	Phoneme	As in
b	/b/	**b**ata (duck); ***b***ox but implosive
ch	/ʧ/	**ch**umba (room); ***ch***eese
d	/d/	**d**eni (debt); ***d***og but implosive
dh	/ð/ dialect /z/	**dh**oruba (storm); ***th***is
f	/f/	**f**imbo (cane); ***f***an
g	/g/	**g**oti (knee); ***g***ate but implosive
gh borrowed	/ɤ/	lu**gh**a (language); voiced velar fricative (lo***ch*** but voiced)
h	/h/	**h**ewa (air); ***h***at
j	/j/; /dj/	**j**engo (building); ***j***ungle
k	/k/	**k**azi (work); kick but implosive
kh borrowed	/x/	nus**kh**a (duplicate)
l	/l/	a**l**ama (mark); ***l***eg (not dark L)
m	/m/	**m**ama (mother); ***m***at
n	/n/	**n**azi (coconut); ***n***ap
ng	/ŋg/	kupa**ng**a (arrange); si***ng***
ng’	/ŋ/	**ng**’ombe (cow); hu***ng***er
ny	/ɲ/	**ny**umba (house)
p	/p/	**p**aka (cat); ***p***at
r	/r/	**r**adi (thunder/lightning); velar flap (***r***adio)
s	/s/	**s**amaki (fish); ***s***ip
sh	/ʃ/	**sh**ule (school); ***sh***ip
t	/t/	**t**angu (since); ***t***op
th	/ɵ/ dialect /s/	**th**umni (fifty cents); ***th***in
v	/v/	**v**azi (dress); ***v***ery
w semi-vowel	/w/	**w**ali (rice); ***w***et
y semi-vowel	/j/	**y**ai (egg); ***y***ellow
z	/z/	**z**awadi (present/reward); zebra

Note: Kiswahili does not use the letters c, x and q. Consonants gh and kh are borrowed from Arabic and Persian.

Source: Compiled from (Alcock & Ngorosho, 2000; Gakuru et al., 2004).

Appendix 2

LIST OF HIGH FREEQUENCY WORDS

kichwa mboga mchwa mfupa mtoto sufuria kwa leo ya redio sahani saba stempu bustani watoto chai chao choo chui ng'ombe mgonjwa ghorofa na shingo simba keki papai paka samaki ana nanasi kima mwanafunzi pweza pwani wa thelathini themanini chakula mkate jaji joho jogoo kaka kuku kofia ni gari ua au oa baba bibi babu dada debe dafu gogo dogo fagio mimi mama watu sukari taa tisa viatu viti vijiko yai yuko zeze fua gauni gunia galoni maji hii hili dereva viatu mayai tandu ndoo kifaranga tembo ugali nyanya nyuki nyoka dhambi fedha njiwa njuga njegere kanzu panzi nyuki shati kwato gwaride kwa (Three and above appearances)

LIST OF LOW FREEQUENCY WORDS

rudi furahi bakora karai huruma barabara hereni raha barua birika rafiki harufu koroga gurudumu ninakula amenunua kuna kokoliko huku kule hama miiba makali hamu mahali mema chuma kamua amua mali hima duma kama nini nuna nina nunu nono nona mimina birika pera gari suka sikio sabuni sokoni salama kasuku usiku saidia pasi kanisa sukuma papasi simama kabisa sikukuu soma sala samaki safari kisima kamasi hesabu kasirika kanisa pasi usile anasoma kitabu sokoni (One or two appearances)

(The words are a compilation based on manual word count of three authorized grade one reading text books refered below).

The Kiswahili alphabet:

a b ch d e f g h i j k l m n o p r s t u v w y z

Source: Kihampa, M. G. (1997); Mkinga, M.G. (2000); Tanzania Institute of Education (TIE) (2000).

Appendix 3

IDENTIFYING FIRST GRADERS AT RISK OF READING AND WRITING DIFFICULTIES

GROUP-BASED SCREENING TEST BATTERY IN KISWAHILI

LONG VERSION

The subscales

(i) Picture-Word
(ii) Word-Picture
(iii) Picture-Letter
(iv) Letter-Picture
(v) Initial Sound
(vi) Word-Chain
(vii) Spelling

The test battery measures phonological awareness, letter identification, word identification and spelling. The tests should be administered halfway in grade one. Before administering the tests, clear and detailed instruction should be given to the pupils. The pupils should be practically engaged in the examples given by the examiners on how to respond to each test.

Ephraim S. Kalanje

Appendix 4

QUESTIONNAIRE

BASIC INFORMATION RELATING TO BASIC READING AND WRITING SKILLS ACQUISITION

A. Demographic information (Fill in or put a tick where appropriate)

Name……School……District……Region……
Male……..Female……Age (years)…….
Birth: Normal……..Premature……
Disability: Name…….Chronic illness: Name…….
General health: Good……Fair…….Poor……
Adequate fine motor coordination: Yes……No……..
Persons at home: Number……
Family status: Father and Mother…….Single……...Guardian……
Attending nursery school: Yes……No……School attendance: Days……

B. Home environment (Circle the appropriate response)

1	Father, Mother, Guardian reading and writing ability	Good	Fair	None
2	Do you get all your school needs?	Yes	Sometimes	No
3	Do you get enough food to eat at home?	Yes, Always	Sometimes	No
4	Does your father, mother, guardian tell you stories?	Yes	Sometimes	No
5	Do you have many books to read at home?	Many	A Few	None
6	Do you use Kiswahili at home for communication?	Yes	Sometimes	No
7	Do you use another language at home for communication?	Yes	Sometimes	No
8	Are you overburdened with house-chores at home?	Yes	Sometimes	No
9	Are you engaged in petty-business at home?	Yes	Sometimes	No
10	Dou you have some time to do some school work at home?	Yes	Sometimes	No
11	Do you receive support with your school work at home?	Yes	Sometimes	No
12	Do you like and enjoy reading books?	Yes	A little	No

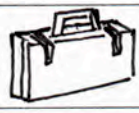
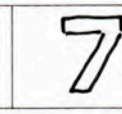
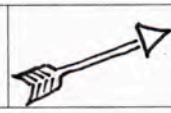
C. School environment (Circle the appropriate response)

1	Does the child speak well with proper articulation?	Yes	Sometimes	No
2	Does the child use vocabulary corresponding to age?	Yes	Sometimes	No
3	Does the child remember names, objects easily?	Yes	Sometimes	No
4	Does the child understand normal conversation?	Yes	Sometimes	No
5	Does the child tell stories?	Yes	Sometimes	No
6	Does the child play with peers?	Yes	Sometimes	No
7	Can the child work independently?	Yes	Sometimes	No
8	Does the child take responsibility for given tasks?	Yes	Sometimes	No
9	What is the child's performance in other school subjects?	Very Good	Good	Poor
10	Is the child willing to compare his/her work with others?	Yes	Sometimes	No
11	Does the child like going to school?	Yes, always	Sometimes	No
12	Is the child active and alert at school?	Yes, Always	Sometimes	No
13	Does the child go to the library or borrow books?	Yes, Often	Sometimes	No

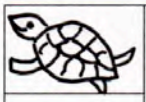
(i) PICHA – NENO (PICTURE – WORD)

	shati	mtoto	nyanya	vijiko
	redio	nyoka	ua	mayai
	nanasi	mboga	kichwa	gauni
	mfupa	zeze	galoni	njiwa
	sukari	nyuki	paka	dereva
	ugali	njegere	kuku	gari
	fagio	pweza	tisa	dafu
	simba	samaki	sahani	kichwa
	shingo	njuga	ndege	fedha
	panzi	ng'ombe	papai	keki

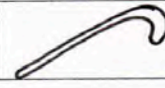
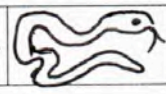
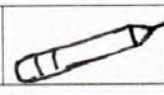
(ii) NENO – PICHA (WORD – PICTURE)

simba				
pweza				
sufuria				
paka				
gari				
tembo				
nyoka				
saba				
shati				
ghorofa				

(iii) PICHA – HERUFI (PICTURE – LETTER)

	k	p	f	i
	z	n	a	m
	u	ch	h	b
	w	l	t	d
	b	d	j	g
	j	g	v	ch
	l	d	r	z
	v	m	f	a
	a	s	h	b
	g	y	p	m
	u	f	n	o
	p	e	n	a

(iv) HERUFI – PICHA (LETTER – PICTURE)

z					
u					
h					
l					
w					
b					
i					
d					
f					
n					
p					
a					

SAUTI SAWA MWANZONI (INITIAL SOUND)

(vi) MKUFU WA MANENO (WORD - CHAIN)

5 (10 words)

auoa	anayai	leomimi	aukwa	ugalitaa

10 (20 words)

tisavijiko	mamatembo	jogoochui	njugachai	sufuriasaba

15 (30 words)

kwatopwani	watuchoo	nanasindogo	sukaripapai	samakisaa

20 (45 words)

viatugogosaba	kofiabibifagio	majindooredio	simbashingogari	debenyanyaviti

shatikanzunjiwa	ng'ombegwaridedafu	chakulafuanyuki	kichwamkatenyoka

25 (60 words)

njegeresahanimboga				

(vii) KUANDIKA MANENO (SPELLING)

1. au
2. na
3. ya
4. leo
5. kuku
6. watu
7. maji
8. paka
9. chai
10. viatu
11. gunia
12. sahani
13. gwaride
14. mgonjwa
15. dhambi
16. themanini
17. ghorofa
18. kifaranga
19. mchwa
20. shingo
21. mwanafunzi
22. ndoo
23. bustani
24. ng'ombe
25. nyanya

Appendix 5

IDENTIFYING FIRST GRADERS AT RISK OF READING AND WRITING DIFFICULTIES

INDIVIDUAL SCREENING TEST BATTERY IN KISWAHILI

The subscales

(i) Pseudo-Word
(ii) Reading Words Aloud
(iii) Actual Text Reading
(iv) One Minute Reading
(v) Writing

The test battery measures reading fluency and writing. The tests should be administered halfway in grade one, a month later after administering the group test. Before administering the tests, clear and detailed instruction should be given to the pupils. The pupils should be practically engaged in the examples given by the examiners on how to respond to each test.

Ephraim S. Kalanje

(i) MANENO BANDIA (PSEUDO – WORDS)

ui	oo	eo	ia	oi
se	no	de	pu	za
vae	uda	ale	pio	heo
kedu	memu	siso	zimu	nuki
mulido	paniso	hiroki	ahuke	zamidi
mdote	sinda	twisa	kwedu	gwido
kwendye	destami	shindu	fyoka	kinyame
sozweka	mchwela	king'oda	sidhambe	kithema
kingwendu	mwenaleenge	kifwefwe	vinywendebi	kwigende
wangechenjwa	horookwenda	anikwimboza	matembedwe	fokozweka

(ii) KUSOMA KWA SAUTI (READING WORDS ALOUD)

oa	leo	saa	bubu	nuna
baba	kikao	yai	bata	kaka
birika	sabuni	kitabu	papasi	sukuma
sukukuu	kasirika	gurudumu	amenuna	barabara
njegere	kifaranga	thelathini	mgonjwa	sungura
mwanafunzi	nyanya	dhambi	ng'ombe	aiskrimu

Paka

Paulo anafuga paka.
Paka huyo anaitwa Pupa.
Pupa anapenda sana kunywa maziwa.
Paulo anampenda sana paka wake.
Kila siku jioni Paulo anampa Pupa maziwa.
Mchana Pupa hukaa nje karibu na pipa.
Pipa hilo liko nje ya jiko.
Usiku Pupa hukamata panya.
Panya wanaposikia sauti ya Pupa wanakimbia.
Pupa analia nyau, nyau, nyau.

Sikukuu

Jana kulikuwa na sikukuu nyumbani kwetu.
Tuliwakaribisha majirani na marafiki zetu.
Baba alichinja mbuzi mmoja.
Alinunua masanduku matano ya soda.
Babu alileta kuku wawili.
Shangazi alileta samaki watatu wakubwa.
Mama na marafiki zake walipika sufuria nne za vyakula.
Tulikula wali, ndizi, nyama na samaki.
Tulikunywa soda mbili kila mmoja.
Kila mtoto alipewa biskuti sita na peremende saba.
Baada ya kula tulipumzika.
Tulicheza ngoma tangu saa nane mpaka saa tisa.
Saa kumi wageni waliondoka.
Sote tulifurahi sana.

KUANDIKA (WRITING)

1. Jina……..(Name)
2. Kuandika herufi zote (Writing all letters)

a	b	c	d
e	f	g	h
i	j	k	l
m	n	o	p
r	s	t	u
v	w	y	z

3. Kuandika maneno (Writing words)

(a) ……oa…… (b)……au……..
(c)……ya……. (d)……na……..

4. (a) ……babu…… (b)……mimi……

(c)……kima…… (d)……tena……

5. Kuandika sentensi (Writing sentences)

(a) anasoma (b) wanakula

(c) Chakula kimekwisha (d) Juma amelala ndani

Appendix 6

IDENTIFYING FIRST GRADERS AT RISK OF READING AND WRITING DIFFICULTIES

FOLLOW UP SCREENING TEST BATTERY IN KISWAHILI

The subscales

(i) Picture-Word
(ii) Initial Sound
(iii) Word-Chain
(iv) Spelling

The test battery measures phonological awareness, letter identification, word identification and spelling. The tests should be administered halfway in grade two. Before administering the tests, clear and detailed instruction should be given to the pupils. The pupils should be practically engaged in the examples given by the examiners on how to respond to each test.

Ephraim S. Kalanje

(i) PICHA – NENO (PICTURE - WORD)

	meza	kopo	saa	bata
	kitabu	birika	kisu	pua
	sikio	kipepeo	duara	miwani
	kitanda	ndoo	mlango	utando
	bendera	ufunguo	sanduku	filimbi
	mdomo	mswaki	mlima	mawingu
	dirisha	chupa	mkasi	mpira
	ngazi	picha	gurudumu	mbwa
	mrembo	konokono	chanuo	nyota
	kikombe	mwavuli	panya	chungwa

(ii) SAUTI SAWA MWANZONI (INITIAL SOUND)

(iii) MKUFU WA MANENO (WORD CHAIN)

5 (10 words)

pipipeni	ogamaji	puabao	bogamama	batakuku

10 (20 words)

papakuni	simbanyani	viatundoo	kobeyai	mpiratatu

15 (30 words)

bibikima	jichoembe	ungajani	mbwambu	mbogachai

20 (45 words)

shatisketimswaki	mtotomchwapunda	shikashonakamba	kwatokwenuimba	kondakasemaleo

dhahabuthamaning'oa	stakishitaking'amua	kwapuakimbiakomoa	mswakipwanibweni

25 (60 words)

pwezabwatabweka			

(iv) KUANDIKA MANENO (SPELLING)

1. mama
2. bibi
3. kata
4. kuni
5. gauni
6. chakula
7. vichwa
8. baiskeli
9. zaka
10. kikapu
11. simba
12. mbwa
13. mbu
14. soksi
15. maandazi
16. shambani
17. ghala
18. sukari
19. kikaango
20. mwanahabari
21. mwenyekiti
22. nyang'au
23. thamani
24. mwanachuo
25. dhahabu

Appendix 7

Descriptive statistics and reliability of the group test items

Analysis results of picture-word subscale and the including items are presented in Table 1. The mean is 59.7, SD=34.0 and reliability, r=.92. Item difficulty index is between 0.21 and 0.90. In this case the test had good enough items (p=0.21 < > 0.82) to be retained except for item 1 (p=0.90) which was easy for the children. Item discrimination index ranged from .51 to .82 indicating that all the items discriminated well.

Table 1. Descriptive statistics for group test picture-word items (N=337)

Item	Mean (%)	sd	*rpbis*	α
1	90	30.6	.51	.93
2	82	38.6	.67	.92
3	80	40.4	.70	.92
4	74	44.2	.77	.91
5	67	47.2	.80	.91
6	59	49.3	.82	.91
7	50	50.1	.82	.91
8	42	49.4	.77	.91
9	34	47.4	.70	.92
10	21	40.8	.53	.93
Total	**59.7**	**34.0**		**.92**

Note: Note: Mean and alpha if item deleted.

Results for the word-picture subscale (Table 2) are more or less similar to the picture word scale. The mean is 59.0, SD 34.8, and r = .93 for reliability. Item difficult index ranged from 0.26 to 0.88. Nine items between 9 (p=0.36) and 2 (p=0.82) were good enough to be retained. Item 1 (p=0.88) was easy for most children. Item discrimination index ranged from .55 to .83 indicating that all items discriminated well.

Table 2. Descriptive statistics for group test word-picture items (N=337)

Item	Mean (%)	SD	r *pbis*	α
1	88	32.4	.55	.93
2	82	38.8	.69	.92
3	77	41.9	.74	.92
4	73	44.6	.80	.91
5	66	47.4	.83	.91
6	55	49.8	.79	.92
7	48	50.0	.76	.92
8	39	49.0	.75	.92
9	36	48.0	.70	.92
10	26	43.7	.56	.93
Total	**59.0**	**34.8**		**.93**

Note: Note: Mean and alpha if item deleted

Picture-letter and letter-picture subscales results also share a similar pattern of results. Their means, SDs and the reliability alpha coefficient are almost identical (Tables 3 and 4). Both the tests have a good item difficult index range and all most all items discriminate well. Only one item 12 (p=0.10), in the letter picture scale, was difficult.

Table 3. Descriptive statistics for group test picture-letter items (N=337)

Item	Mean (%)	SD	r *pbis*	α
1	85	35.3	.61	.93
2	82	38.1	.66	.93
3	76	42.6	.73	.93
4	72	45.2	.76	.93
5	66	47.6	.85	.92
6	62	48.6	.84	.92
7	60	49.1	.84	.93
8	50	50.1	.81	.93
9	42	49.5	.74	.93
10	35	47.9	.68	.93
11	25	43.5	.58	.93
12	15	35.6	.42	.93
Total	**67.1**	**41.0**		**.93**

Note: Mean and alpha if item deleted.The general mean for the letter-picture subscale items was 67 (SD 40), (Table 17). Reliability was .94. Item difficult index ranged from 0.10 to 0.85, therefore 11 items between 11 (p=0.24) and 1 (p=0.85) qualified the criteria to be retained as good items. One item, 12 (p=0.10), was difficult for most children. Item discrimination ranged from .38 to .87 indicating that all the items discriminated well.

Table 4. Descriptive statistics for group test letter-picture items (N=337)

Item	Mean (%)	SD	r *pbis*	α
1	85	35.9	63	.93
2	85	36.2	.62	.93
3	79	41.0	.72	.93
4	74	43.8	.78	.93
5	66	47.5	.84	.93
6	61	48.8	.86	.92
7	61	49.0	.87	.92
8	52	50.0	.83	.93
9	40	49.1	.75	.93
10	32	46.7	.67	.93
11	24	43.0	.58	.94
12	10	30.6	.38	.94
Total	**67**	**40.3**		**.93**

Note: Note: Mean and alpha if item deleted.

Table 5 indicates that the general mean for the initial sound subscale items was 48 (SD 33). Reliability was .91. Item difficult index ranged from 0.27 to 0.85. All the 10 items met the criteria to be retained as good items. Item discrimination index ranged from .53 to .81, indicating that all the items discriminated well. The strategy of random organization of the task items in this scale may have created the order of increasing difficulty in this scale. See explanation in the method chapter.

Table 5. Descriptive statistics for group test initial sound items (N=337)

Item	Mean (%)	SD	r *pbis*	α
1	85	36.2	.55	.91
2	79	40.6	.64	.90
3	72	44.8	.71	.90
4	62	48.7	.78	.89
5	54	49.9	.81	.89
6	35	47.8	.68	.90
7	22	41.5	.60	.90
8	16	37.0	.53	.91
9	31	46.5	.71	.90
10	27	44.5	.66	.90
Total	**48.4**	**32.5**		**.91**

Note: Note: Mean and alpha if item deleted.

Table 6 indicates that the general mean for the word chain subscale items was 12.39 (SD 8.28). Reliability was .97. Item index difficulty ranged from 0.07 to 0.91. The 20 items between 22 (p=0.20) and 3 (p=0.83) were good enough to be retained. Five items did not meet the criteria: items 1(p=0.91) and 2 (p=0.86) were easy. On the other hand, items 23 (p=0.12), 24 (p=11) and 25 (p=0.07) were difficult for the children. Item discrimination index ranged from .39 to .87, indicating that all items discriminated well.

Table 6. Descriptive statistics for group test word chain items (N=337)

Item	Mean (%)	SD	r *pbis*	α
1	91	28.6	.44	.97
2	86	35.0	.56	.97
3	83	37.6	.61	.97
4	80	40.5	.67	.97
5	76	42.7	.71	.97
6	72	45.1	.76	.97
7	69	46.4	.79	.97
8	66	47.6	.82	.97
9	63	48.4	.83	.97
10	60	49.0	.84	.97
11	55	49.8	.87	.97
12	53	50.0	.87	.97
13	51	50.1	.87	.97
14	47	50.0	.87	.97
15	45	49.8	.86	.97
16	41	49.2	.84	.97
17	38	48.6	.83	.97
18	33	47.1	.79	.97
19	29	45.2	.76	.97
20	28	45.0	.75	.97
21	23	42.1	.69	.97
22	20	39.8	.64	.97
23	12	32.1	.51	.97
24	12	31.7	.50	.97
25	07	25.3	.39	.97
Total	**12.39**	**8.28**		**.97**

Note: Mean and alpha if item deleted.

Table 7 indicates that the general mean for the spelling subscale items was 11.66 (SD 8.57). Reliability was .97. Item difficulty index ranged from 0.10 to 0.88.Twenty two items between 23 (p=0.20) and 2 (p=0.85) were good enough to be retained. Three items did not meet the criteria: Item 1 (p=0.88) was easy. Items 24 (p=0.12) and 25 (p=0.10) were difficulty. Item discrimination index ranged from .472 to .864 indicating that all the items discriminated well.

Table 7. Descriptive statistics for group test spelling items (N=337)

Item	Mean (%)	SD	r *pbis*	α
1	88	32.7	.47	.97
2	85	35.3	.52	.97
3	82	38.8	.58	.97
4	77	41.9	.64	.97
5	72	45.1	.71	.97
6	68	46.8	.75	.97
7	67	47.1	.76	.97
8	64	48.0	.77	.97
9	57	49.6	.82	.97
10	55	49.8	.82	.97
11	50	50.1	.84	.97
12	45	49.9	.85	.97
13	41	49.3	.86	.97
14	37	48.4	.86	.97
15	36	48.0	.86	.97
16	34	47.5	.86	.97
17	34	47.3	.86	.97
18	32	46.6	.85	.97
19	28	45.1	.82	.97
20	26	43.8	.79	.97
21	25	43.1	.78	.97
22	25	71.8	.49	.97
23	20	39.7	.70	.97
24	12	33.1	.56	.97
25	10	29.8	.49	.97
Total	**11.66**	**8.57**		**.97**

Note: Mean and alpha if item deleted.

Appendix 8

Descriptive statistics and reliability of the follow-up test items

Table 8 shows that the mean for the Follow-up Spelling subscale items was 13.38 (SD 6.29). The reliability was .95. Item difficulty index ranged from 0.02 to 0.95. Fourteen items between nr 19 (p=0.21) and nr 6 (p=0.85) were good enough to be retained. Eleven items did not meet the criteria: items 1, 2 and 3 (p=0.95) were easy. Item 21 (p=0.08), items 22-24 (p=0.04) and item 25 (p=0.02) were difficult. Item discrimination index ranged from .38 to .83, indicating that all the items discriminated well. The items in this scale were arranged in order of increasing difficulty.

Table 8. Descriptive statistics for follow-up spelling scale items (N=330)

Item	Mean (%)	SD	r *pbis*	α
1	95	20.9	.43	.95
2	95	21.5	.44	.95
3	95	21.5	.44	.95
4	88	32.0	.62	.95
5	88	32.3	.63	.95
6	85	35.6	.68	.95
7	82	38.9	.73	.95
8	81	39.4	.74	.95
9	74	44.1	.80	.94
10	72	45.2	.81	.94
11	68	46.5	.83	.94
12	67	47.2	.83	.94
12	63	48.3	.83	.94
14	59	49.2	.82	.94
15	52	50.0	.80	.94
16	44	49.7	.77	.95
17	40	49.1	.77	.95
18	33	47.1	.71	.95
19	21	40.7	.62	.95
20	14	35.0	.56	.95
21	08	26.5	.49	.95
22	04	20.2	.45	.95
23	04	20.2	.45	.95
24	04	18.7	.43	.95
25	02	13.4	.38	.95
Total	**13.38**	**6.29**		**.95**

Note: Note: Mean and alpha if item deleted.

Table 9 indicates that the general mean for the picture-word subscale items was 5.66 (SD 2.48). The reliability was .86. Item difficulty index ranged from 0.02 to 0.95. Three items were very easy (nr 1, p=0.95 nr 2, p=0.91 and nr 3, p=0.86) and two were very difficult (nr 9, p=0.10 and nr 10, p=0.02). Item discrimination index ranged from .20 to .78, which indicates that all items discriminate well except for item nr 10 (D=.20). It should be noticed that the very low scores on this scale may have been resulted from some children not completing the task towards the end for reasons explained in the methods chapter.

Table 9. Descriptive statistics for follow-up picture-word items (N=330)

Item	Mean (%)	SD	r *pbis*	α
1	95	20.9	.43	.86
2	91	28.8	.56	.85
3	86	34.4	.67	.84
4	79	40.9	.76	.83
5	72	44.9	.78	.83
6	61	48.9	.76	.83
7	48	50.0	.68	.84
8	21	40.7	.48	.86
9	10	30.0	.37	.86
10	02	14.4	.20	.87
Total	**5.66**	**2.48**		**.86**

Note: Mean and alpha if item deleted.

Table 10 indicates that the general mean for the follow-up initial sound subscale items was 4.57 (SD 2.30). The reliability was .84. Item difficulty index ranged from 0.01 to 0.95. Items nr 1 (p=0.95) and nr 2 (p=0.88) were easy. Three items were difficult (nr 8, p=0.09, nr 9, p=0.02 and nr 10, p=0.01). Item discrimination index ranged from .19 to .75. This indicates that all items discriminate well, except items nr 9 (D=.21) and 10 (D=.19). The decreasing of order of difficulty of items in this scale and the reasons for very low scores may have been caused by the random arrangement of the scale items to create a proactive and retroactive interference strategy and some children not completing the tasks, as explained in the method chapter.

Table 10. Descriptive statistics for follow-up initial sound items (N=330)

Item	Mean (%)	SD	r *pbis*	α
1	95	22.1	.38	.84
2	88	33.0	.56	.82
3	78	41.8	.70	.81
4	69	46.3	.75	.80
5	57	49.6	.75	.80
6	39	48.9	.68	.81
7	20	40.1	.54	.82
8	09	28.4	.41	.83
9	02	14.4	.21	.84
10	01	09.5	.19	.85
Total	**4.57**	**2.30**		**.84**

Note: Mean and alpha if item deleted.

Table 11 indicates that the general mean for the follow-up word chain subscale items was 12.06 (SD 5.80). The reliability was .94. Item difficulty index ranged from 0.00 to 0.94. Thirteen items between 18 (p=0.16) and 6 (p=0.83) were good enough to be retained. Five items 1 to 5, (p=.89 – 94) were easy, while seven items, 19 to 25 (p=.00 - .11) were difficult. Items also discriminate well, except for the items nr 22 to 25 (D=.11 - .21). The items in this scale were arranged in order of increasing difficulty. This is explained in the method chapter.

Table 11. Descriptive statistics for follow-up word chain (N=330)

Item	Mean (%)	SD	r *pbis*	α
1	94	24.0	.49	.94
2	91	28.0	.58	.94
3	89	30.9	.64	.94
4	89	30.9	.64	.94
5	89	31.7	.65	.94
6	83	37.7	.73	.94
7	83	37.4	.73	.94
8	79	41.0	.74	.94
9	75	43.5	.78	.94
10	68	46.8	.79	.94
11	63	48.3	.80	.93
12	57	49.6	.81	.93
12	52	50.0	.80	.93
14	49	50.1	.79	.93
15	43	49.6	.75	.94
16	32	46.7	.68	.94
17	30	46.0	.66	.94
18	16	36.9	.51	.94
19	11	31.7	.43	.94
20	05	20.9	.33	.94
21	05	21.6	.33	.94
22	02	12.3	.21	.94
23	00	05.5	.11	.94
24	00	05.5	.11	.94
25	00	05.5	.11	.94
Total	**12.06**	**5.80**		**.94**

Note: Mean and alpha if item deleted.

Appendix 9

Short version of the group test

The short version of the group test comprises three subscales. The subscales include picture letter measuring, letter knowledge, spelling, and word chain measuring word knowledge. These subscales have been selected because they have been shown to have good reliability and they are easy to administer to a group of children. They are also cost-effective and require comparatively little time to administer. The instructions and demonstrations for the chosen scales are also easily understood by the children. The subscale picture-word is not included in the short version, although it was the strongest single predictor for the follow-up scores. However, according to the experience of the teachers, it was suggested that the chosen subscales be used in the classroom praxis.

The short version has a good technical adequacy. The three subscales significantly and strongly intercorrelate (r=.82 -.88).

The subscales of the short version also correlate significantly with the corresponding subscales in the original version. The correlation coefficients are as follows:

Short version picture letter - original picture-letter, r=.982
Short version spelling - original spelling, r=.988
Short version word chain - original word chain, r=.941

Items analysis reveals that the reliability of the three scales is good (alpha .87 - .93) and that the items discriminate well. Table 1 presents the items selected from the original test for each subscale in the short version. The means, SDs and alpha coefficients are also presented. The original test consisted of seven scales with a total of 104 items. The short version has three scales, comprising 30 items. This is 48% of 62 items corresponding to those in the three original subscales. The table also indicates the specific items that were selected from the original long version.

Table1: Number of items, items selected means, SDs and alpha coefficients.

Shorter version					**N=337**
Subscale	**Items**	**Items selected from original test**	**Mean**	**SD**	**α**
shvpletter	5	2,4,6,8, and 12	3.02	1.87	.87
shvspell	10	2,5,10,12,13,14,16,18,21,24	4.39	3.64	.94
shvchword	15	1,2,3,4,5,16,17,18,19,20,21,22,23,24,25	6.52	4.39	.93
Total	**30**				

Note: shvpletter=shorter version picture-letter; shvspell=shorter version spelling and shvchword=shorter version chain word.

Further evidence of the good functioning of the short version is explained from the K-Means cluster analysis results (Table 2) of the three short version subscales. The four clusters are distinct and similar to those from the original clusters. The 89 (26%) *at risk* children classified by the short version are quite close to the 82 (24%) classified by the long version. The slight difference is not a problem from the risk point of view.

In the basic study the children were grouped into four groups based on their scores in the different scales in the test. The statistical method used was cluster analysis (K-Means). The groups were named *at risk, strugglers, readers* and *good readers* according to their mean scores from the scales. The same procedure was repeated with the short version of the test. The results are almost similar, indicating that the short version is appropriate for use in screening. The results are presented in Table 2.

Table 2: Comparing the amount of children classified in different cluster groups in the original and short version of the test.

Test	At-risk		Strugglers		Readers		Good readers	
	N	%	N	%	N	%	N	%
Original version	82	24.3	75	22.3	81	24	99	24.9
Short version	89	26.4	97	28.8	70	20.8	81	24.0

The distribution of children in cluster groups according to gender, age, and school status and school and nursery attendance is presented in Table 3. Some differences should be noticed. The main groups in focus are the *at-risk* group and the *strugglers* (clusters one and two).

There are clearly a larger amount of children from the older age groups, from low performing and peripheral schools, with poor school attendance and with no experience of nursery school in the at risk group and in the strugglers group than in the other groups.

Table 3: The distribution of children in cluster groups according to gender, age, school status, and school and nursery attendance

Factors	Scales	Cluster groups				Total	
		1	2	3	4	N	%
Age groups	9 - 12	42	46	12	0	26	100
	6 - 8	25	27	22	26	311	100
Gender	Girls	24	26	22	28	156	100
	boys	29	31	19	21	181	100
School status	low	41	30	20	9	79	100
	periphery	35	37	7	21	71	100
	good	29	24	23	25	80	100
	High	8	26	29	36	107	100
School attendance	poor	86	14	0	0	43	100
	good	25	42	16	17	184	100
	very good	6	13	36	45	110	100
Attending nursery	no	52	33	14	1	151	100
	yes	6	25	26	43	186	100
Total		26	29	21	24	337	100

Note: Cluster groups: 1=in risk; 2=strugglers; 3=readers; 4=good readers

A summary variable was created and a score was counted based on the three scales (mean=147.7, SD=96.3, maximum score 300, minimum score=0, median =143 and mode=0). The analysis of the distribution reveals that 25% of the children have scores below 70. The cluster analysis ended up with 26% based on the short version, whereas the analysis with the original test resulted in 24%. The suggestion is that children scoring lower than 70 (*at risk* group) need intensive support. Children scoring between 70 and 150 also belong to a group needing attention and observation. They can somehow manage, but are struggling. With extra attention and some support, most of them probably can manage well. When using the test, notice should be taken of the fact that many background factors have an effect on the test results.

Appendix 10

IDENTIFYING FIRST GRADERS AT RISK OF READING AND WRITING DIFFICULTIES

GROUP-BASED SCREENING TEST BATTERY IN KISWAHILI

SHORT VERSION

The subscales

(i) Picture-Word
(ii) Spelling
(iii) Word-Chain

The test battery measures phonological awareness, letter identification, word identification and spelling. The tests should be administered halfway in grade one. Before administering the tests, clear and detailed instruction should be given to the pupils. The pupils should be practically engaged in the examples given by the examiners on how to respond to each test.

Ephraim S. Kalanje

(i) PICHA – HERUFI (PICTURE - LETTER)

	z	n	a	m
	w	l	t	d
	j	g	v	ch
	v	m	f	a
	p	e	n	a

(ii) KUANDIKA MANENO (SPELLING)

1. na (2)
2. kuku (5)
3. viatu (10)
4. sahani (12)
5. gwaride (13)
6. mgonjwa (14)
7. themanini (16)
8. kifaranga (18)
9. mwanafunzi (21)
10. ng'ombe (24)

(iii) MKUFU WA MANENO (WORD - CHAIN)

5 (10 words)

auoa	anayai	leomimi	aukwa	ugalitaa

10(25 words)

viatugogosaba	kofiabibifagio	majindooredio	simbashingogari	debenyanyaviti

shatikanzunjiwa	ng'ombegwaridedafu	chakulafuanyuki	kichwamkatenyoka

15(40 words)

njegeresahanimboga				

Early identification of children at risk of reading and writing difficulties is important for prevention and the provision of proper intervention before failure sets in. This thesis addresses the issue in the Tanzanian context, where many children in primary schools face reading and writing problems. This is evidenced in the high repetition and dropout rates and poor school attendance of children. Many children complete primary education without the skills of reading and writing. The children remain unidentified for many reasons, amongst them a lack of proper screening instruments for identification.

The aim of this study was to create and validate a theoretically founded group-based screening tool in Kiswahili for identifying first graders at risk of reading and writing difficulties. The role of certain home- and school-related factors for the children's reading and writing ability was also analyzed. The created screen was shown to have high reliability.

Of 337 children screened in Morogoro Municipality, Tanzania, 24% were found to be at risk of reading and writing difficulties and in need of special attention. These children obtained very low scores on all the tasks designed to measure letter identification, word identification, phonological awareness and spelling. These components are known to be sensitive predictors of reading in alphabetic languages, including transparent orthographies like Kiswahili. The screen also identified a group of children who are struggling, but who can manage with a little support. Many over-aged school beginners seem to belong to the at-risk group. Of the home factors, the number of books at home explained 38% of the variance in reading and writing ability. Parents´ reading ability and the support children get at home were also strong factors. Of the school-related factors, school attendance was critical, explaining 21% of the variance. Nursery school experience was also important. The study emphasizes that action is needed to support homes and schools in the process of improving school attendance.

The screening instrument is considered effective for identifying children at risk of reading and writing difficulties. In addition to the original version, a short version is also presented. Both versions are easy to administer in a short time and do not require special orientation. The screen can be used by teachers as well as parents.

Åbo Akademi University Press
ISBN 978-951-765-590-3

Ephraim S. Kalanje : Identification of First Graders at Risk of Reading and Writing Difficulties

2011