



UNIVERSITEIT VAN PRETORIA  
UNIVERSITY OF PRETORIA  
YUNIBESITHI YA PRETORIA

# **The influence of Teaching Handwriting, Reading and Spelling Skills on the accuracy of word level reading**

**Robert John Alexander Stark**

**2009**



**The influence of Teaching Handwriting,  
Reading and Spelling Skills on the accuracy  
of word level reading**

by

**Robert John Alexander Stark**

Submitted in partial fulfilment of the requirements for the degree

**MAGISTER EDUCATIONIS  
(Educational Psychology)**

in the

Department of Educational Psychology  
Faculty of Education  
UNIVERSITY OF PRETORIA

**SUPERVISOR:**  
Dr S Bester

PRETORIA  
August 2009

## Acknowledgements

---

My wife, Shireen. For your unwavering love, encouragement and support, for pushing me at times when I was resisting fiercely – you are a true inspiration.

My parents, Dave and Ros. Thank you for your support and your quiet confidence in me.

---oOo---



## Declarations

---

I declare that: *The influence of Teaching Handwriting, Reading and Spelling Skills on the accuracy of word level reading* is my own work. All the sources used or quoted have been indicated and acknowledged by means of complete references. This dissertation was not previously submitted by me for a degree at another university.

---

Robert John Alexander Stark  
28 August 2009

---oOo---

This serves to certify that Peter John Hulley (I.D. 581216 5147 08 6) edited a thesis for Robert John Alexander Stark, a student at the Faculty of Education, University of Pretoria, entitled: *The influence of Teaching Handwriting, Reading and Spelling Skills on the accuracy of word level reading*.

P. J. Hulley  
(Writing & Editing)  
22 August 2009

---oOo---



The purpose of the study was to investigate the influence of THRASS (Teaching Handwriting, Reading and Spelling Skills) on the word level accuracy skills of a group of grade 2 learners. Word level accuracy is one sub skill in learning to read and is an indicator of the word recognition abilities of the child. THRASS is a program that has been designed to systematically teach phonics and, thus, teaches the basic building blocks of word sounds and structure so as to improve the child's decoding ability and word recognition ability. The research took place within the positivist paradigm and the methodology is quantitative in nature. The data collection method took the form of a one group pretest-posttest design, where a standardised reading test was administered prior to exposing the participants to the THRASS Program and then readministered one year later on the same group of learners. Data analysis took the form of statistical analysis to investigate any statistical significant difference in the word level accuracy skills of those Grade 2 learners. The result showed that over the period of a year the average reading accuracy age for the target population increased by four months. However, after statistical analysis the difference was not statistically significant. The Null Hypothesis that; exposing a group of Grade 2 learners to the THRASS Program for a period of one year will have no statistically significant influence on their word level accuracy skills cannot be rejected . However, the changes both in average reading accuracy as well as error patterns have inspired recommendations for further research.

**Key words:**

Alphabetic principle	Phonics
One-group pretest-posttest design	Phoneme awareness
THRASS Program	Phonological awareness
Word recognition models	Pre-experimental design
Word level reading accuracy	Positivist

---oOo---



## Table of Contents

	Page
<b>Chapter 1</b>	
<b>Introduction and Rationale of Study</b>	
<b>1.1 INTRODUCTION</b>	1
<b>1.2 CONCEPTUAL FRAMEWORK</b>	2
<b>1.2.1 THE ALPHABETIC PRINCIPLE</b>	2
<b>1.2.2 PHONEME AWARENESS</b>	3
<b>1.2.3 PHONICS TEACHING</b>	3
<b>1.2.4 WORD RECOGNITION</b>	3
<b>1.2.5 WORD LEVEL READING ACCURACY</b>	3
<b>1.2.6 THRASS PROGRAM (TEACHING HANDWRITING, READING AND SPELLING SKILLS)</b>	4
<b>1.2.7 INTERRELATEDNESS OF THE KEY CONCEPTS</b>	4
<b>1.3 PRELIMINARY LITERATURE REVIEW</b>	4
<b>1.3.1 OVERVIEW</b>	4
<b>1.3.2 EARLY READING DEVELOPMENT</b>	4
<b>1.3.3 TEACHING HANDWRITING, READING AND SPELLING SKILLS (THRASS)</b>	6
<b>1.3.4 ASSESSMENT OF READING</b>	6
<b>1.3.5 CONCLUSION</b>	7
<b>1.4 STATEMENT OF PURPOSE</b>	7
<b>1.5 RESEARCH HYPOTHESIS</b>	7
<b>1.6 RATIONALE AND GENERAL ORIENTATION OF THE STATEMENT OF PURPOSE</b>	8
<b>1.7 RESEARCH PARADIGM</b>	8
<b>1.8 THE RESEARCH PROCESS</b>	9
<b>1.8.1 RESEARCH DESIGN</b>	9
<b>1.8.2 METHODOLOGICAL PARADIGM</b>	9
<b>1.8.3 SAMPLING</b>	10
<b>1.8.4 DATA COLLECTION</b>	10



1.8.5	DATA ANALYSIS	11
1.8.5.1	Data reduction	11
1.8.5.2	Data display	11
1.8.5.3	Statistical data display	11
1.8.6	VALIDITY AND RELIABILITY CONCERNS	12
1.8.6.1	Validity	12
1.8.6.2	Reliability	12
1.8.7	GRAPHIC SUMMARY OF THE RESEARCH PROCESS	13
1.9	ETHICAL CONSIDERATIONS	13
1.10	LIMITATIONS OF THE RESEARCH	14
1.11	POSSIBLE CONTRIBUTIONS OF THE STUDY	15
1.12	SUMMARY	15

---oOo---

## **Chapter 2**

### **THRASS, Early Reading Development, Instruction and Assessment**

<b>2.1</b>	<b>CHAPTER OVERVIEW</b>	16
<b>2.2</b>	<b>INTRODUCTION</b>	16
<b>2.3</b>	<b>THRASS IN THE SOUTH AFRICAN CONTEXT</b>	17
<b>2.4</b>	<b>THE THRASS PROGRAM</b>	18
2.4.1	INTRODUCTION	18
2.4.2	THE THRASS STAGES	20
<b>2.5</b>	<b>PHONEME AWARENESS</b>	22
<b>2.6</b>	<b>WORD RECOGNITION MODELS</b>	23
2.6.1	INTRODUCTION	23
2.6.2	THE DUAL-ROUTE MODEL	24
2.6.3	THE ANALOGY MODEL	25
2.6.4	THE CONNECTIONIST MODEL	25
2.6.5	WORD RECOGNITION MODELS – CONCLUSION	26
<b>2.7</b>	<b>FORMAL READING INSTRUCTION THEORY</b>	27
2.7.1	INTRODUCTION	27
2.7.2	THE BOTTOM-UP APPROACH TO READING INSTRUCTION	27
2.7.3	THE TOP-DOWN APPROACH TO READING INSTRUCTION	28
2.7.4	FORMAL READING INSTRUCTION – CONCLUSION	28
<b>2.8</b>	<b>ASSESSMENT OF READING</b>	29
2.8.1	INTRODUCTION	29
2.8.2	THE PURPOSE OF ASSESSMENT	29
2.8.3	EXPERTISE NECESSARY TO CONDUCT READING ASSESSMENTS AND ANALYSE RESULTS	30
2.8.4	INSTRUMENTS FOR ASSESSMENT	31
2.8.5	THE NEALE ANALYSIS OF READING ABILITY-REVISED (NARA)	32
2.8.6	ASSESSMENT OF READING ACCURACY	33
<b>2.9</b>	<b>CONCLUSION</b>	34



## Chapter 3 Research Methodology

<b>3.1</b>	<b>CHAPTER OVERVIEW</b>	35
<b>3.2</b>	<b>RESEARCH PARADIGM</b>	35
<b>3.3</b>	<b>RESEARCH DESIGN</b>	36
<b>3.4</b>	<b>METHODOLOGICAL PARADIGM</b>	38
3.4.1	QUANTITATIVE MEASUREMENT VALIDITY	38
<b>3.5</b>	<b>THE VALIDITY AND RELIABILITY OF THE NARA</b>	40
<b>3.6</b>	<b>SAMPLING</b>	42
<b>3.7</b>	<b>PLANNING FOR DATA ANALYSIS</b>	42
3.7.1	DATA REDUCTION	43
3.7.2	DATA DISPLAY	43
3.7.3	DATA ANALYSIS	43
3.7.4	CONCLUSION AND VERIFICATION OF THE DATA	45
3.7.4.1	Validity	46
3.7.4.2	Reliability	48
<b>3.8</b>	<b>CONCLUSION</b>	49

---oOo---



## Chapter 4 Quantitative Analysis

<b>4.1</b>	<b>INTRODUCTION</b>	50
<b>4.2</b>	<b>DESCRIPTIVE ANALYSIS</b>	50
4.2.1	BAR CHARTS	50
4.2.1.1	Stanines	52
4.2.2	FREQUENCY DISTRIBUTIONS OF ANALYSIS OF READING ERRORS	53
<b>4.3</b>	<b>INFERENTIAL STATISTICAL ANALYSIS</b>	56
4.3.1	TEACHER EFFECT	56
4.3.2	COMPARISON OF THE PRETEST AND POSTTEST REGARDING AVERAGE ACCURACY	57
4.3.3	INVESTIGATING PROGRESS OF LEARNERS BETWEEN PRETEST AND POSTTEST RESULTS	58
4.3.4	COMPARISON OF THE AVERAGE STANINE SCORES ON THE PRETEST AND THE POSTTEST	60
4.3.5	COMPARISON OF READING ERRORS ON THE PRETEST AND POSTTEST	60
4.3.5.1	Number of words read	61
4.3.5.2	Types of errors made on the pretest and the posttest	61
<b>4.4</b>	<b>SUMMARY OF QUANTATIVE ANALYSIS</b>	62
<b>4.5</b>	<b>CHAPTER SUMMARY</b>	63

---oOo---



## Chapter 5 Findings and Conclusions

<b>5.1</b>	<b>INTRODUCTION</b>	64
<b>5.2</b>	<b>DISCUSSION OF RESEARCH RESULTS</b>	65
<b>5.3</b>	<b>LIMITATIONS OF THE STUDY</b>	66
5.3.1	ABSENCE OF CONTROL GROUP	66
5.3.2	TIME ALLOWED FOR THRASS IMPLEMENTATION	66
5.3.3	MATURATION	67
5.3.4	INDIVIDUAL DIFFERENCES	67
5.3.5	EXPOSURE TO OTHER LEARNING OPPORTUNITIES	67
5.3.6	SUBJECT REACTION TOWARD ASSESSMENT	68
5.3.7	NARA IN SOUTH AFRICAN CONTEXT	68
5.3.8	THRASS IN SOUTH AFRICAN CONTEXT	68
<b>5.4</b>	<b>CONTRIBUTIONS OF THE STUDY</b>	
<b>5.5</b>	<b>RECOMMENDATIONS FOR FURTHER RESEARCH</b>	70
5.5.1	LONGITUDINAL STUDY	71
5.5.2	RESEARCH USING CONTROL GROUPS	71
5.5.3	ERROR ANALYSIS	71
5.5.4	ASSESSING THE DEVELOPMENT IN READING ACCURACY IN DIFFERENT CONTEXTS	71
5.5.5	INTENSITY OF EXPOSURE TO THE THRASS PROGRAM	71
5.5.6	ASSESSING DIFFERENT ASPECTS OF READING DEVELOPMENT FOLLOWING EXPOSURE TO THRASS	72
<b>5.6</b>	<b>DEVELOPMENTS IMPACTING ON FUTURE THRASS RELATED STUDIES</b>	72
5.6.1	DEVELOPMENT OF TEACHING MEDIUMS	73
5.6.2	CONFIDENCE IN THE THRASS PROGRAM	73
<b>5.7</b>	<b>CONCLUSION</b>	73
<b>REFERENCES</b>		74

## List of Figures

---

	Page
Figure 2.1: The THRASS Stages .....	21
Figure 4.1: Average chronological age versus average accuracy age-pretest and posttest .....	51
Figure 4.2a: Normal distribution curve and stanine illustration .....	52
Figure 4.2b: Bar chart of average stanine for pretest versus posttest .....	53
Figure 4.3: Comparison of total number of words read: pretest versus posttest .....	54
Figure 4.4: Percentage reading errors for the pretest versus posttest .....	54
Figure 4.5: Number of words read accurately on pretest and posttest .....	55
Figure 4.6: Percentage reading error in total number of words read for pretest and posttest by reading error category .....	55

## List of Tables

---

	Page
Table 4.1: Central tendencies and measures of spread for the pretest and posttest ages .....	51
Table 4.2: Kruskal-Wallis test to detect teacher effect in the pretest scores .....	57
Table 4.3: Kruskal-Wallis test to detect teacher effect in the posttest scores .....	57
Table 4.4: Friedman test on the accuracy between the pretest and posttest (n=36) .....	58
Table 4.5: Contingency table and frequencies of posttest and pretest categories for Class 1 .....	59
Table 4.6: Contingency table and frequencies of posttest and pretest categories for Class 2 .....	59
Table 4.7: Statistical analysis of pretest stanines versus posttest stanines(n=36) .....	60
Table 4.8: Comparison between the number of words read during the pretes and the posttest (n=36) .....	61
Table 4.9: Comparison between the different types of reading errors made during the pretest versus the posttest (n=36) .....	62



## **1.1 INTRODUCTION**

Since 2002 I have had the privilege of working with diverse groups of learners<sup>1</sup> and have, therefore, been exposed to the development of many pupils in terms of learning to read. I have also been exposed, to some degree, to the differing levels of competence with which children read. It was the varying reading strategies learners apply as well as their varying reading abilities that developed into an area of interest for me.

I was first introduced to the THRASS Program in my capacity as an educator at a public school that works with children with special needs and I became intrigued by the enjoyable and simple approach to teaching reading it presented. In 2007 a decision was made at the private school on the East Rand where I am currently employed as school counsellor to implement that programme. And, after attending a THRASS teaching course in 2006, I was subsequently involved in the implementation phase. During that period I became interested in researching the programme. As the THRASS Program is focused on teaching word level accuracy skills, it seemed natural to research the influence the programme had on the word accuracy levels of the learners who were exposed to it.

The broad focus of this study is on early reading development and, more specifically, phoneme awareness and word level reading. The philosophy that supports the THRASS system is that by systematically exposing learners to the 44 phonemes (24 consonants and 20 vowels) and the 120 key graphemes (66 consonants and 60 vowels) of the English language it enables the learner opportunity to become phoneme and grapheme rich (Davies, 2006).

THRASS is recommended by the National Literacy Strategy Framework (NLS) in Britain as an early phonics instructional system. Research done by the NLS indicated that there were significant progress rates (Rose, 2006:24) in the accuracy of reading, pronouncing and spelling of words of children who had gone through the four-year THRASS Program. The NLS framework is based on the same principles as the South African revised National Curriculum namely:

---

<sup>1</sup> For this study the term learner refers to the norm population used in this study, namely a group of Grade 2 learners.

*Through phonics we should be teaching the children to identify sounds in spoken words [phonological awareness]; recognise the common spellings for each phoneme [phoneme grapheme correspondence]; blend phonemes into words for reading, segment words into phonemes for spelling (Rose, 2006:10).*

The South African Department of Education (DoE) approaches the teaching of reading in their revised National Curriculum with the focus being on children's emergent literacy i.e. reading 'real' books, writing for genuine purposes and attention to phonics (RNCS<sup>2</sup>, 2005:22). The learning of phonics is addressed in the RNCS in Learning Outcome 3, Reading and Viewing which is achieved when the learner begins to develop phonic awareness, recognises initial consonant and short vowel sounds (RNCS, 2005:37).

The THRASS Program was developed for learners in the United Kingdom and, although reports of its success there can be accessed, we do not know what the effectiveness thereof would be in the South African context. It is, therefore, of utmost importance to start a research process to determine the influence of that programme and its possible value in addressing the word level accuracy skills of learners in South Africa.

## 1.2 CONCEPTUAL FRAMEWORK

In this section working definitions of the key concepts as they are relevant to this study are provided; which concepts are extrapolated upon in the literature review presented in Chapter 2.

### 1.2.1 THE ALPHABETIC PRINCIPLE

The alphabetic principle is the awareness that there is a systematic relationship between speech sounds (phoneme awareness) and words in print (orthographic awareness) (Adam, 1990). The alphabetic principle is often referred to as the code of reading and once a child has become aware that there is a relationship between the spoken and written word, deciphering this code facilitates the ability to read, write and spell – and then discover an ever-expanding world of meaning. The deciphering of the code is where formal reading teaching commences and, as Schumm (2006) describes, that formal teaching of reading begins with the teaching of phonics.

---

<sup>2</sup> The RNCS refers to the Revised National Curriculum Statement that outlines the curriculum for the learning areas for children from Grades R-12 in South Africa.

### 1.2.2 PHONEME AWARENESS

One of the primary associations children need to make is that sounds make up words and that they can be broken down into small sound units. This association is discussed as phonological awareness. Phoneme awareness forms a component of phonological awareness. Phoneme awareness is defined as the awareness of the small units of sound that influences the meaning of the word (Gillon, 2004).

### 1.2.3 PHONICS TEACHING

Phonics teaching helps children clearly see the relationship between letters and sounds and it clarifies that letters actually mimic the phonemic sequences of words. According to Schumm (2006), the primary approach to phonics teaching is associated to the bottom-up<sup>3</sup> perspective and is called synthetic phonics. This perspective emphasises part-to-whole instruction, as children are taught sounds in isolation and are then asked to blend the sounds to form words. Analytic phonics is associated with the top-down perspective, which advocates that reading is primarily meaning making, fuelled by the readers growing knowledge of the spoken word. According to Schumm (2006), analytic phonics emphasises whole-to-part-to-whole instruction, which translates into children being given whole words with similar patterns and the child is then expected to make generalisations about the parts of words that are similar.

### 1.2.4 WORD RECOGNITION

Word recognition is the association of a printed word with its spoken equivalent (Share & Stanovich, 1995). Two main models of word recognition, the dual-route model and the connectionist model, dominate word recognition theory; however, there are related models such as the modified dual-route model as well as the analogy model (Gillon, 2004). This was discussed in Chapter 2.

### 1.2.5 WORD LEVEL READING ACCURACY

Word level reading accuracy refers to the ability to recognize or decode words correctly. Thorough understanding of the alphabetic principle, the ability to blend sounds together (Ehri & McCormick, 1998, Hudson, Lane & Pullen, 2005), and knowledge of a large bank of high-frequency words is required for word-reading accuracy (Byrne & Fielding-Barnsley, 1989).

---

<sup>3</sup> The bottom-up approach to reading instruction and synthetic phonics is discussed in Chapter 2, Section 2.7.2.

## 1.2.6 THRASS PROGRAM (TEACHING HANDWRITING, READING AND SPELLING SKILLS)

THRASS is a programme designed for word level teaching, that is, the phonics aspect of teaching beginner reading skills. THRASS breaks away from the traditional synthetic phonics approach where children are taught that 'one letter makes one sound' to an approach called the phoneme-grapheme (phonographic) method, which is phonics teaching that exposes children from the very beginning of the reading instruction to all the 44 phonemes of spoken English and the graphs (one letter graphemes), digraphs (two letter graphemes), trigraphs (three-letter graphemes) and quadgraphs (four letter graphemes) of written English (Davies 2006).

## 1.2.7 INTERRELATEDNESS OF THE KEY CONCEPTS

The key concepts defined above are interrelated with each other as they integrate to form the foundations of word level reading. They are progressive i.e. one stage is completed before focusing on the next; yet, at the same time, each stage is reliant on the others. A child needs to understand the alphabetic principle and this through the teaching of phonics that understanding is facilitated. However phoneme awareness also influences a child's understanding of the alphabetic principle and phonic teaching plays a role in phoneme awareness. Once the alphabetic principle has been grasped and phonics instruction takes place a child's word recognition should improve and therefore his accuracy of reading. THRASS teaches phonographic phonics to improve a child's word level reading.

## 1.3 PRELIMINARY LITERATURE REVIEW

### 1.3.1 OVERVIEW

The primary focus of the literature reviewed for this study is on the developmental processes of early reading development and, more specifically, on the foundations necessary to facilitate learning to read. The THRASS program was discussed in relation to the key concepts mentioned previously as well as the assessment of reading progress.

### 1.3.2 EARLY READING DEVELOPMENT

The process of learning to read proficiently extends over several years and involves the acquisition and co-ordination of a range of sub-skills. At the basic level, reading is an alphabetic system that involves building both a sight word vocabulary and developing decoding skills (i.e. word recognition ability). The goal of reading, however, is not merely word level decoding or word recognition but comprehension, which requires the integration of

meaning across sentences and within texts, making appropriate use of contextual cues and inferences based on general knowledge (Snowling, 1998). The focus of this study is limited to word recognition and word level decoding; it excludes comprehension - although it is acknowledged that the focal areas selected do play a vital role in the comprehensive goal of reading.

Early reading skills develop through exposure to environmental print, building book and print awareness, along with exposure to the sounds of the language (phonetic awareness) which, according to Sulzby and Teal (1991), precede formal reading instruction. When children have been sufficiently exposed to print in its various forms and have mastered this emergent step in the literary process, Adams (1990) suggests that children will then be ready to master the idea of the alphabetic principle; which is the awareness that there is a systematic relationship between speech sounds (phonemic awareness) and words in print (orthographic awareness).

There are yet other elements that have to be mastered when learning to read. Bouwer (2000) and Pike, Compain and Mumper (1997), state that readers make use of three sources of information, namely, semantic, syntactic and graphophonic information. Graphophonic information underpins the decoding process, affording the reader access to each word in the text (Bouwer, 2000). It is this graphophonic information that is of primary interest in this study. Graphophonic information includes the facilitation and instruction of phonics. Phonics instruction teaches sound-symbol relationships and patterns, with the aim of reinforcing the alphabetic principle.

Phonics teaching is thus closely associated with both learning to read as well as learning to spell (Manyak, 2007). Dechant (1993) stresses that phonics teaching is critical as it helps children to clearly see the relationship between letters and sounds and it clarifies that letters actually mimic the phonemic sequences of words. Dechant (1993) elaborates further stating that phonics instruction is a step in the process of learning to read that involves automatic semantic activation, complete understanding of the alphabetic principle, proficient word identification, better reading and spelling, and, finally, to the ultimate goal of reading - comprehension.

The alphabetic principle is often referred to as the code of reading and, once a child has become aware that there is a relationship between the spoken and written word, deciphering this code facilitates the ability to read, write and spell and then discover an ever expanding world of meaning. The deciphering of the code is where formal reading instruction begins and, as Schumm (2006) describes, this formal teaching of reading begins with the teaching

of phonics. The THRASS Program advocates the use of both synthetic phonics and analytic phonics (Davies, 2006).

### **1.3.3 TEACHING HANDWRITING, READING AND SPELLING SKILLS (THRASS)**

The THRASS Program postulates that synthetic and analytic phonics learning should take place through phoneme-grapheme or new phonics teaching that exposes children to the 44 phonemes of spoken English and the graphs, digraphs, trigraphs and quadgraphs of written English. This begins by exposing THRASS learners to 120 keywords that each contain a key grapheme, representing a specific consonant phoneme or vowel phoneme that are explained, discussed, understood and assessed at the appropriate sub-stage or stage of the THRASS Program (Davies, 2006).

Children, therefore, do not only identify letters or decode words by the letter sound; they use their understanding of graphs, digraphs, trigraphs and quadgraphs to decode many words of the English language forming the core of new phonics teaching (Davies, 2006). What is unclear, however, is the effectiveness of such teaching. It appears to make logical sense to expose children to all the phonemes and graphemes of the English language through this program. However, it is not clear as to whether children fully comprehend this approach and if indeed it positively affects their understanding of the alphabetic principle and, thus, their word level reading. While some success has been documented in the United Kingdom, evaluation of the effectiveness of that programme in South Africa has yet to be established.

### **1.3.4 ASSESSMENT OF READING**

Children's reading ability is measured to assess the development and progress of reading and the effectiveness of the teaching method being used. Robeck and Wallace (1990) purport that reading progress is measured and difficulties assessed in order to take the most logical steps when instructing students. Tests help professionals *construct reasonable hypotheses and explanations of a student's progress and go some ways as to provide valuable information for determining factors that may be causing a particular reading problem for a pupil* (Robeck & Wallace, 1990:294). Schumm and Arguelles (2006) list the various purposes of assessment and, whilst they tend to agree with Robeck and Wallace (1990), they include the evaluation of the strengths and weaknesses of instructional programmes as a purpose of assessment.

The assessment measure that was used for this study is the Neale Analysis of Reading Ability (NARA). The NARA is a norm-referenced test that consists of a set of graded

passages for testing the accuracy, comprehension and rate of oral reading, and a set of supplementary tests for diagnostic assessment. *It is both an attainment test and a diagnostic test and can be used to assess reading progress objectively in a school setting, for example, or can be used to obtain structured diagnostic or clinical observations of an individual's reading behaviour* (Neale, 1997:6). Schumm (2006) suggests that norm referenced tests are typically recommended for assessing reading instruction programmes and as screening tools. The emphasis of norm-referenced tests is on the relative standing of an individual among students of a norm group of the same age or grade level, or the performance of a group of students as a whole at the same age or grade level (Foxcroft & Roodt, 2001).

It is recognized that the NARA assesses three components of reading, namely, accuracy, comprehension and rate. For the purpose of this study only the accuracy score will be scrutinized, as the THRASS Program emphasizes word level decoding skills and phonetic awareness which are involved in recognition and pronunciation of words i.e. word recognition and reading accuracy.

### **1.3.5 CONCLUSION**

Early reading development has been introduced as the broad subject informing this study. Phonics instruction was given specific attention as being a critical skill in learning to decode the code (alphabetic principle) of words and its relationship to the THRASS Program. The THRASS Program uses a combination of synthetic and analytic phonics, teaching children about the 44 phonemes and 120 graphemes of the English language from the outset, with the aim of increasing the efficiency at which children grasp the alphabetic principle. The NARA *aims to identify any pattern that may be occurring at the grapho-phonetic, the syntactic, or the semantic level of reading behaviour* (Neale, 1997:49). It is the identification of any pattern at the grapho-phonetic level (reading accuracy) that may be occurring that is of particular interest for this study.

### **1.4 STATEMENT OF PURPOSE**

The focus of this study is to investigate the word level accuracy scores of a group of Grade 2 learners prior to and after exposing them to the THRASS Program for one year.

### **1.5 RESEARCH HYPOTHESIS**

The hypothesis stated for this study is:

- Ho: Exposing a group of Grade 2 learners to the THRASS Program for a period of one year will have no statistically significant influence on their word level accuracy skills.
- Ha: Exposing a group of Grade 2 learners to the THRASS Program for a period of one year will have a statistically significant influence on their word level accuracy skills.

## **1.6 RATIONALE AND GENERAL ORIENTATION OF THE STATEMENT OF PURPOSE**

This research was initiated from the need for it being expressed by the management of the school at which I am employed, who wanted to determine whether the implementation of the THRASS Program had any positive influence on the development of the reading abilities of learners. Being a counsellor at that school I seized the opportunity to investigate the learner reading development by means of a scientific approach, which study has culminated in this report. However, even before that request was delivered, it did interest me to conduct research on the THRASS Program within the South African context. My interest in this field will be supported by recommendations for future research in Chapter 5.

## **1.7 RESEARCH PARADIGM**

It is acknowledged by the researcher that the research perspective or paradigm within which research is conducted usually informs the methodology of the research. However in this specific study the research was conducted ex post facto, after the assessments results were produced.

The research paradigm of this study was thus informed by the research methodology which takes the form of quantitative analysis of the results obtained from the pretest and posttest administered to the target population. The data collection method is positivist in nature, as it is an objective measure using numbers (word level accuracy scores) and subjecting those numbers to various statistical procedures, with the results being objectively observed (Scott & Usher, 2001). Thus this study is based within the positivist paradigm<sup>4</sup>.

Although ontologically, epistemologically and methodologically it is the aim through this study to view reality from a positivist perspective i.e. as orderly, predictable and controllable (Scott & Usher, 2001), it is acknowledged that reading and the acquiring of word level accuracy may be informed and influenced by a myriad of variables. Those variables are, as far as is possible, acknowledged in this study and referred to later. It is, however, the primary

---

<sup>4</sup> The positivist paradigm is discussed in more detail in Chapter 3, Section 3.2.

objective through this research study to measure and observe the data and present the findings of the statistical evaluation objectively.

According to Mcmillan and Schumacher (2001:329), *there are certain cases in pre-experimental design in which threats to internal validity can be ruled out on the basis of accepted theory, common sense, and other data*; and it is with this goal in mind that I interpreted the results from the pretest-posttest design.

## 1.8 THE RESEARCH PROCESS

### 1.8.1 RESEARCH DESIGN

This study was a pre-experimental one-group pretest-posttest design<sup>5</sup>. The participants were specifically and purposefully selected and the control of extraneous variables in the design was limited. According to McMillan and Schumacher (2001:329) *pre-experimental designs are often used as a way of generating ideas*, which can then be researched in more detail at a later stage.

The aim of this study was to start a process of research i.e. to start generating ideas on how the THRASS Program can be researched in the South African context and arrive at conclusions on the possible efficacy of the THRASS Program at the private school where I am employed.

With that aim in mind, I felt it was adequate and appropriate to start the research process with a pre-experimental design, albeit that there are certain internal validity implications, which are discussed in Chapter 3.

### 1.8.2 METHODOLOGICAL PARADIGM

The research methodology is quantitative<sup>6</sup> in nature. There is a strong tradition in educational research to use numbers and measurement. The approach emphasizes *a priori* categories to collect data in the form of numbers. The goal is to collect data to provide statistical descriptions, relationships and explanations. Quantitative techniques are used with experimental and descriptive designs as a way to summarise a large number of observations and to indicate numerically the amount of error in collecting and reporting the data (McMillan & Schumacher, 2001).

---

<sup>5</sup> The pre-experimental one group pretest-posttest design was discussed in more detail in Chapter 3, Section 3.3.

<sup>6</sup> The Quantitative methodological paradigm was discussed in more detail in Chapter 3, Section 3.4.

The data for this study has been collected in the form of numbers or accuracy scores that have been calculated from the answer sheets of the standardized test used - namely the NARA. These results have then been subjected to various statistical methods of analysis in order to make comparisons and draw conclusions.

### 1.8.3 SAMPLING<sup>7</sup>

The participants involved in the research process will be taken from both Grade 2 classes at the school – which is the entire population of Grade 2 learners. The results of this research study will thus be representative of this population of Grade 2 learners only. It was never the intention to generalize the findings of this study, since the main focus was to determine the reading development/progress of a specific group of learners and thus provide insight to the teaching staff at the specific private school with regard to the efficacy of the THRASS Program, which was implemented at the school in 2007.

Demographically the participants are of mixed race and gender. Their ages range from 7 years 3 months to 8 years 7 months on the pretest and 8 years 3 months and 9 years 7 months on the posttest. Thus convenience as well as purposive sampling was applied to this study.

### 1.8.4 DATA COLLECTION

The advantage of applying a quantitative research method is the *demonstration that accurate and observable progress could be statistically demonstrated by averages of comparison* (Neuman, 1997:56). Quantitative data will be collected through applying a norm-referenced test. Using the NARA will assess the individual reading abilities of the target population. The pretest results and the posttest results will be statistically compared, to indicate and validate an increase/decrease in the specific construct of reading accuracy as measured by the NARA.

The results were originally recorded on the answer sheet of the NARA. An accuracy score was then calculated and tallied for each individual participant. Thereafter those scores were grouped together and placed on a spreadsheet where averages for the chronological ages and test ages were documented. This was done for both the pretest and posttest scores. The error analysis comprised recording each individual's error record on the NARA answer sheet for both the pretest and posttest, which was recorded on a spreadsheet. That allowed for the average errors to be tallied as well as any error patterns to be observed. The total number of

---

<sup>7</sup> Sampling is discussed further in Chapter 3, Section 3.6.

errors and the total number of words was recorded. Those results were then subjected to various statistical tests to establish any significance of variables.

### 1.8.5 DATA ANALYSIS

De Vos (2000:48) describes the data analysis stage as the '*ponder-and-check stage*', as *efforts are made to draw evidence that relates to the concept together*. The data-analysis method proposed by Huberman and Miles (1997), De Vos (2000) and Leong (2006) indicate three associated sub-processes that may occur before, during and after data collection. The main steps that can be followed during analysis and interpreting qualitative data are outlined below.

#### 1.8.5.1 Data reduction

The researcher *selects a conceptual framework, research questions, cases and psychometric assessment instruments to discern the focal point and differentiate the aim of the particular field of study from the broad data available* (De Vos, 2000:56). Data were continuously reduced by averages of various data collection methods: data from initial pretest results from standardized media versus the data from standardized media received in the posttest. The error analysis was taken from the NARA. The categories of errors are stipulated in NARA as Mispronunciations, Substitutions, Refusals, Additions, Omissions and Reversals.

#### 1.8.5.2 Data display

This step involved thorough organisation of concise data (namely: norm scores, tabulating averages, and thorough statistical analysis of standardised results). The display facilitated the creation of a foundation that stimulated further examination and contemplation that assisted in simplifying the final procedure of having to draw conclusions from the data.

#### 1.8.5.3 Statistical data display

Various statistical measures have been applied to test the differences observed through the results. Some of the tests used are the Friedman Test, Kruskal-Wallace Test and the T-test<sup>8</sup>. The aim of applying the Kruskal-Wallace Test was to accommodate the statistical analysis of two independent samples, namely, the two Grade 2 classes, allowing the comparison between the two samples. The Friedman Test accommodated the use of two sample groups that were associated to each other, namely the pretest and posttest groups, which thus

---

<sup>8</sup> Those tests are described in more detail in Chapter 3, Section 3.7.3.

enabled the statistical comparison or averages of both pretest and posttest groups and tested the hypothesis related to the averages observed (Steyn, Smit, Du Toit & Strasheim 1994). The T-test was conducted in order to establish the degree of difference in the averages related to the pretest and posttests, and the probability of the degree of difference being statistically significant.

### **1.8.6 VALIDITY AND RELIABILITY CONCERNS<sup>9</sup>**

#### **1.8.6.1 Validity**

The adequacy of the research must be considered. i.e. The validity of the research must be maximised. There are two main kinds of validity one may consider, namely, internal and external validity, which must be integrated into a research design (Terre Blanche & Durrheim, 1999). For a study to possess internal validity its findings are said to flow directly and unproblematically from its methods, the findings and conclusions are then said to be sustained (De Vos, 2000). A research study is said to have external validity when its findings and conclusions can be generalised beyond the confines of the design and study setting (Neuman, 1997). Generalisability refers to the extent to which the results or findings of a study can be extrapolated to a wider context than that used in the research design; the context refers to the population from which the sample was drawn and operationalisation of the measurement methods of the key variables (Terre Blanche & Durrheim, 1999).

The validity of the data produced in this study is discussed in Chapter 3. It does not fall within the scope of this study to generalise the research result to wider context. Thus the focus is on internal validity. The purpose of this research is to describe the data for this specific population only.

#### **1.8.6.2 Reliability**

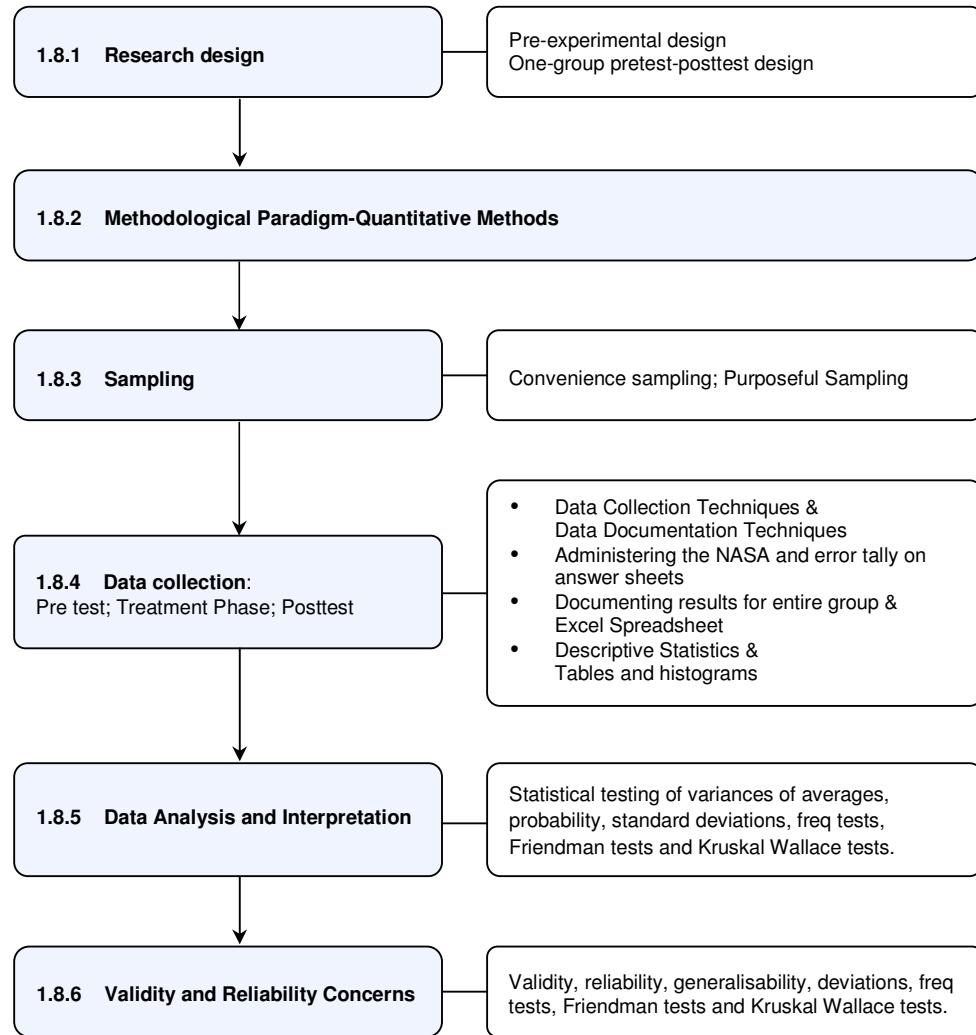
The reliability of a measure and findings is the degree to which the results are repeatable and dependable (Terre Blanche & Durrheim, 1999). The reliability of the data collection method applied for this research, i.e. the NARA, has been reported on to reveal three main types of reliability: stability (parallel forms) reliability, internal consistency, and standard error of measurement of the test. The repeatability of this study will be complex, especially with the same population, as it is not possible for this group of Grade 2 learners to be unexposed to the THRASS Program. However the dependability of the findings can be investigated by subjecting the data to various statistical tests. Internal consistency has been used to establish the reliability of the findings of this study. Internal consistency has been calculated

---

<sup>9</sup> Validity and Reliability was further explored in Chapter 3, section 3.7.4.

to establish reliability by estimating the degree to which each item on the scale correlates to each other (Neuman, 1997).

### 1.8.7 GRAPHIC SUMMARY OF THE RESEARCH PROCESS



### 1.9 ETHICAL CONSIDERATIONS

De Vos (2000) highlights the importance of considering ethical guidelines in the social sciences, and remarks on the distinctive ethical problems that may manifest in the field of social sciences. De Vos (2000) points out a number of obstacles that should be avoided by the researcher and based on the aforementioned the following strategies were followed:

- Informed consent: A clear indication of the goal of the investigation, the procedures, advantages, and disadvantages, the dangers to which the participant may be exposed and the credibility of the researchers was discussed with the

participants and their guardians and a letter of consent was issued outlining the above.

- Informed consent from the institution (i.e. the school and the teacher) was necessary, as the school is responsible for the children's well being during school hours and while the child is on the school property. The indications and goals of the study were outlined in the form of a letter to the school with a discussion.
- The results of the study will be kept confidential, the names and individual results did not form part of the publication of the research study and the accuracy scores and error analysis was published for the group not for the individual child.
- Harm to the participant was minimized as far as possible. De Vos (2000) emphasizes the ethical obligation towards the participant to protect him/her against any form of physical or emotional harm<sup>10</sup>.
- The analysis and interpretation of the results was performed by a statistical services organisation not by me, as the researcher. This contributed to the validity and reliability of the results.
- The results of the study were made known to the participants and their guardians.

## **1.10 LIMITATIONS OF THE RESEARCH**

It is appropriate to consider that we can only realistically gain a limited understanding of the world by referring to the facts (Terre Blanche & Durrheim, 1999). It is thus useful to, firstly, acknowledge this limitation and, secondly, explore the limitation, by taking into account appropriate background knowledge and influences that may have had an impact on the study and results thereof:

- Due to the nature of the sample population, the findings were not suitable to be generalized to populations outside the grade and class under investigation.
- The THRASS Program may or may not be credited with acknowledgement of improved or decreased performance, where the credit may be associated to one or many other aspects of the reading process.
- The results for each individual on the NARA may not be an entirely true reflection of each child's ability with reading, as there are many reasons why a child may not perform to his/her potential on the test. For example: cultural or linguistic deprivation, non English speaking background, sensory impairments, emotional problems or poor motivation (NARA, 1997)<sup>11</sup>.

---

<sup>10</sup> Due to the subjects of this study being familiar to me, it was considered appropriate that the emotional effects of assessment may be reduced to some degree. Performance anxiety cannot be disregarded as being influential on the result of this study but, as the scope of this study does not include that measurement, that aspect has not been measured.

<sup>11</sup> Extraneous variables as a limitation to this study are discussed further in Chapter 5.

- Measures to facilitate validity and trustworthiness have been taken to make the study as feasible as possible, however the limitations will be considered when analyzing the data and recording the findings.

### **1.11 POSSIBLE CONTRIBUTIONS OF THE STUDY**

The nature of the design used in this study, at most, allows for the generation of research ideas; however the results have allowed feedback to be given to the school at which the program is currently implemented. The study aimed to contribute to the generation of suggestions with regard to further research, which are discussed in Chapter 5.

### **1.12 SUMMARY**

The research processes to be undertaken during this study was outlined in Chapter 1. In the interest of sound research, the remaining chapters will be devoted to an extrapolation of those processes. Firstly, to dissect the current research pertaining to early reading development and, more specifically, on how teaching phonics could facilitate word level accuracy skills. Secondly, to illustrate the method and methodology used in this study were applied with the aim of achieving the utmost validity and reliability of both the method and the findings. Thirdly, the conclusions and findings are reported and possibilities for further research explored.

---oOo---



## 2.1 CHAPTER OVERVIEW

Chapter 2 presents a review of the THRASS Program and the elements of early reading development that the program was designed to address, namely, the development of skills necessary for competence in word level reading. Phoneme awareness, grapheme awareness and various instructional methods pertaining to early reading are explored, which are reviewed by means of the current theory relating to word recognition models. Various methods for reading instruction are reviewed in relation to how the aforementioned relate to the THRASS Program. Lastly, the literature relevant to the assessment of reading is considered, specifically the assessment of word level accuracy and how the NARA was used to gather data for this study.

## 2.2 INTRODUCTION

Bouwer (2000) describes reading as a communicative act that includes the interaction of mutually informative processes. The interaction and combination of these processes add constantly to the reader's understanding and thinking about the communication represented by text. Burns, Roe and Ross (1992) refer to reading as a life skill that combines nine aspects of the reading process to facilitate reading. Those aspects are sensory, perceptual, sequential, experiential, thinking, learning, associational, affective and constructive. Decoding and comprehension represent the two broad divisions of reading, however all nine aspects can be acknowledged as influential in the reading process.

It is argued that sensory, perceptual and sequential aspects feature most prominently in the decoding process of reading and the remaining six in the comprehension process of reading (Bouwer, 2000). Reading can thus be referred to as both an interactive and intra-active mode of communication. Intra-active as the reader is constantly making use of interactions between his or her own abilities, experiences, knowledge and skills to facilitate internal communication, understanding (semantics) and personal meaning of what is being read. And interactive because the

*... reader is conducting dialogue with the text in an endeavour to process the sense intended by the author and uses the information of the printed text to construct meaning (Bouwer & Jordaan, 2002:134).*

Bouwer (2000), Pike, Compair and Mumper (1997) purport that in reading running text readers make use of three sources of information, namely, semantic, syntactic and graphophonic information. Semantic information involves the reader's knowledge of meanings in the language of the text and associations with relevant experiences and existing content knowledge, which combine to form the frame of reference to construct the personal meaning of the text. Syntactic knowledge relates to the processing of the statements, as well as enabling the reader to constantly monitor their own reading and correct themselves if errors are made or words are misread. Graphophonic information underpins the decoding process, affording the reader access to each word in the text (Bouwer, 2000).

Prior to formal or structured reading instruction taking place, Gillon (2004) and Barone and Marrow (2003) state that the sound structure of words has to be understood and that this lays the foundation for successful reading instruction (Phillips, Clancy-Menchetti, & Lonigan, 2008). The sound structure of words is referred to as phoneme awareness.

It is the aim with the THRASS Program to teach this sound structure by teaching the 44 phonemes of the English language and aims to link knowledge of sound structures to symbol structures, i.e. the 120 key graphemes thus reinforcing the alphabetic principle and integrating the sound and symbol relationships necessary for word recognition and word level reading.

It is acknowledged that reading involves complex and interactive processes and that those processes have interrelated functions. It is, however, for the purposes of this study important to focus primarily on one aspect of the reading process, which will be the decoding process of reading. It is with this focus in mind that the literature regarding the decoding process of reading, and how this relates to the THRASS Program and early reading development, will be discussed.

### **2.3 THRASS IN THE SOUTH AFRICAN CONTEXT**

The THRASS Program, which was pioneered by British educational psychologist Alan Davies, has been welcomed as heralding the start of a new era in the teaching of phonics in South Africa (Mabotja, S., 2008, August. *THRASS ABSA Talk together project*. 11 August 2008, <http://www.thrass.co.uk/absa.htm>).

The developers of the THRASS Program, together with the South African promoters of that program, together with corporate sponsorship are pioneering projects that have the primary aim of increasing literacy in South Africa (Mabotja, S., 2008, August. *THRASS ABSA Talk*

together project. 11 August 2008, <http://www.thrass.co.uk/absa.htm>). That is to be achieved through the improved teaching and learning of English that will be achieved by partnering universities, primary schools and other organisations. The secondary aim is to provide a platform to celebrate the eleven national languages of South Africa, through free interactive software and printable calendar charts<sup>12</sup> (THRASS. 2008, August. *THRASS Keyfacts*. Accessed 11 August 2008 on <http://www.thrass.co.uk/absa.htm>).

To date over 5,000 teachers and student teachers in South Africa have earned the THRASS Accredited Certificate (THRASS. 2008, August. *THRASS Keyfacts*. Accessed 11 August 2008 on <http://www.thrass.co.uk/absa.htm>). And a number of both public and private schools have begun to implement the THRASS Program across their respective literacy curricula (THRASS. 2008, August. *THRASS Keyfacts*. Accessed 11 August 2008 on <http://www.thrass.co.uk/absa.htm>).

Education role players in universities, such as Dr. Jean Place, Head of Foundation Phase at Wits School of Education, say:

*The THRASS programme, enhanced with the new and innovative SING-A-LONG resources, will enable teachers to integrate their approach with all aspects of the new National Literacy Strategy. These resources and the THRASS approach to early reading will, I believe, impact significantly on literacy levels within South Africa* (THRASS. 2008, August. *THRASS Keyfacts*. Accessed 11 August 2008 on <http://www.thrass.co.uk/absa.htm>).

In South Africa, the following has been achieved by THRASS: Since the first course in 2005, the THRASS Program has been introduced in over 120 public and private schools. More than 5000 teachers and student teachers have completed the THRASS Accredited Certificate.

The THRASS Program is supported by a growing number of universities with foundation phase education faculties, who have included THRASS as a mandatory component of their curriculum. These include the University of Witwatersrand, University of Pretoria and the Cape Peninsula University of Technology, Fort Hare University and the University of KwaZulu-Natal (Mabotja, S, 2008, August. *THRASS ABSA Talk together project*. Accessed 11 August 2008 on <http://www.thrass.co.uk/absa.htm>).

---

<sup>12</sup> The interactive software and calendar charts were discussed in Chapter 5.

## 2.4 THE THRASS PROGRAM

### 2.4.1 INTRODUCTION

It is reported that *THRASS* is widely accepted as providing a revolutionary approach to teaching English and is used by teachers, parents, educational psychologists, speech and language therapists in thousands of schools in Africa, the UK, Europe and the Caribbean (Mabotja, S, 2008, August. *THRASS ABSA Talk together project*. Accessed 11 August 2008 on <http://www.thrass.co.uk/absa.htm>).

*THRASS* is a whole-school phonics programme for teaching learners of all ages and abilities using pictures and keywords. The programme has four phases and Ten Stages. *THRASS* teaches learners that when spelling we change phonemes to graphemes and when reading we change graphemes to phonemes. *THRASS* uses a multi-sensory approach (visual, auditory and kinaesthetic) to teach meta-cognitive strategies that empower learners with lifelong word-solving skills (THRASS. 2008, August. *THRASS Keyfacts*. Accessed 11 August 2008 on <http://www.thrass.co.uk/absa.htm>).

Davies (2006) states that beginning reading instruction has three levels: Word level teaching (often referred to as phonics), sentence level teaching and text Level teaching. The building blocks of word level teaching in the *THRASS* Program are phonemes (speech sounds) and graphemes (spelling choices); the building blocks for sentence level teaching are spoken and written words; and the building blocks for text level teaching are spoken and written sentences. All three levels require teachers to teach speaking, listening, reading and writing skills.

The *THRASS* Program is primarily designed for word level teaching, that is, the phonics aspect of teaching beginning reading skills (Davies, 2008). *THRASS* breaks away from the old phonics methodology of letter sounds, where children are taught that 'one letter makes one sound method' to a new phonics approach called the phoneme-grapheme (phonographic) method, which is phonics teaching that exposes children from the very beginning of the reading process at school to the 44 phonemes of spoken English and the graphs (one letter graphemes), digraphs (two letter graphemes), trigraphs (three-letter graphemes) and quadgraphs (four letter graphemes) of written English. According to Davies (2006), old phonics teaching of one letter makes one sound teaching, *limits learners, in that it is a system that allows only a small percentage of the 500 base words of English to be decoded - resulting in the learner being phoneme deprived*.

The THRASS Program advocates the use of both synthetic phonics and analytic phonics<sup>13</sup> and that synthetic and analytic phonics learning should take place through phoneme-grapheme teaching. That begins by exposing THRASS learners to 120 keywords (wholes) that each contains a bold key grapheme (part), representing a specific consonant phoneme (part) or vowel phoneme (part) that are explained, discussed, understood and assessed at the appropriate sub-stage or stage (Davies, 2006).

Children therefore do not only identify letters or decode words by the letter sound; they use their understanding of graphs, digraphs, trigraphs and quadgraphs to decode many English words forming the core of new phonics teaching (Davies, 2006). The THRASS Program is a four year program and is outlined as follows:

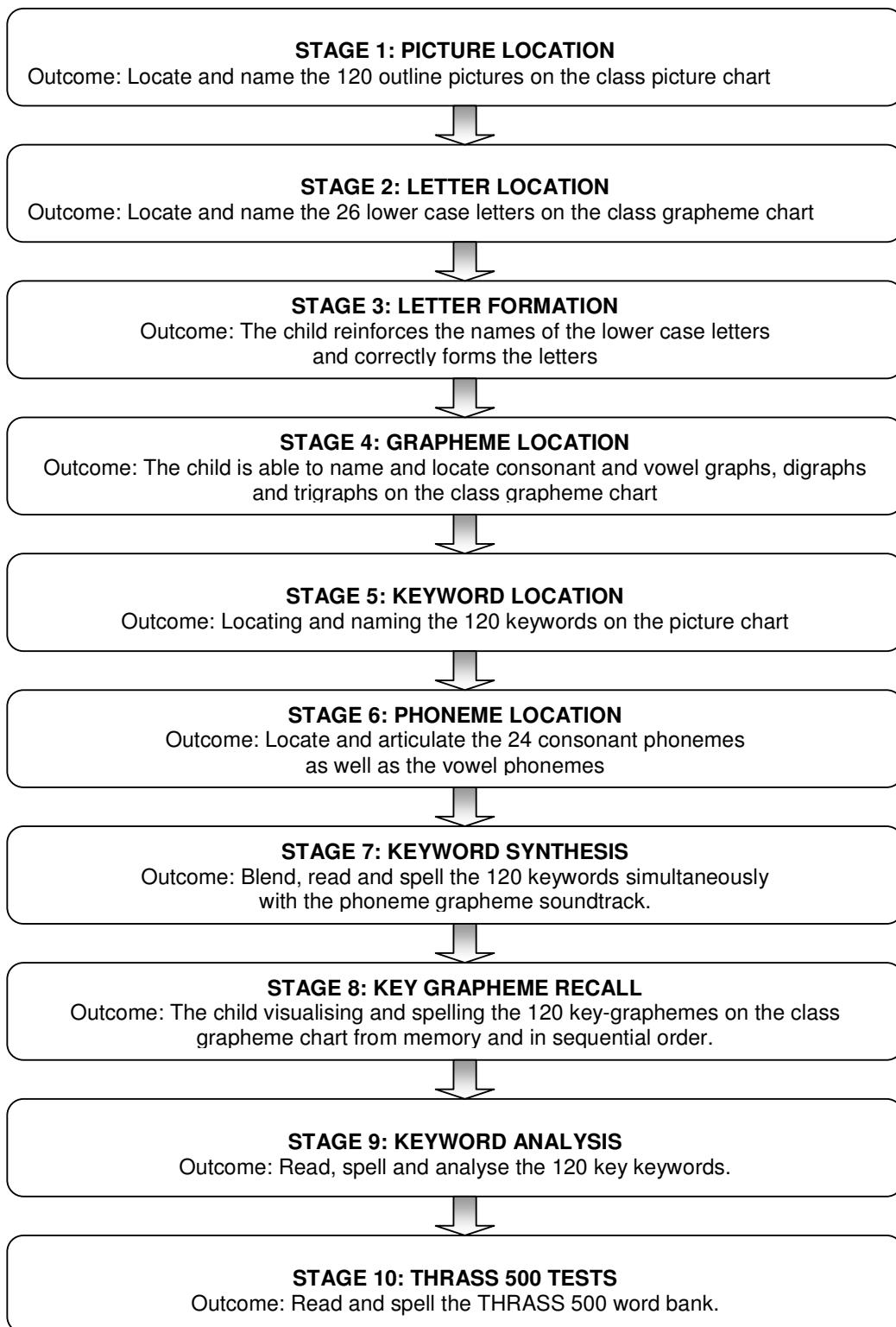
PHASES AND STAGES	
(The stages will be outlined in the following section):	
<b>PHASE P1</b>  Year 1 (Ideally 4 year olds should have acquired these skills*)	T1: Picture location
	T2: Letter Location
	T3: Letter Formation
	T4: Grapheme Location
	T5: Keyword Location
<b>Phase P2</b>  Year 2 (Ideally 5 year olds should have acquired these skills*)	T6: Phoneme Location
	T7: Keyword Synthesis
<b>Phase P3</b>  Year 3 (Ideally 6 year olds should have acquired these skills*)	T8: Keygrapheme recall
	T9: Keyword Analysis
<b>Phase 4</b>  Year 4 (Ideally 7 year olds should have acquired these skills*)	T10: THRASS 500 Tests
* Or begin one or more years later. For much older learners change the stage order, and the number of minutes for teaching, each day to suit the needs and abilities of the learners. For example try Stage order 2-10: T2, T3, T8 (which includes T1 and T5), T4, T6, T7, T9, T10 (Davies; 2006)	

The participants for this study were exposed to the first two years of the THRASS Program, namely Phase 1 and Phase 2.

#### 2.4.2 THE THRASS STAGES

<sup>13</sup> Synthetic and analytic phonics is discussed in section 2.7.

Children are exposed to the THRASS Program in a systematic way and through a series of progressive stages. Those stages are indicated in the figure below.



**Figure 2.1: The THRASS Stages**

As stated earlier in this chapter, the THRASS Program is a whole-school phonics program that teaches phoneme and grapheme awareness. And therefore phoneme awareness will be explored in the following section. Thereafter, the relationship between phoneme and grapheme awareness will be explored with relation to current theory pertaining to word recognition strategies.

## 2.5 PHONEME AWARENESS

Phoneme awareness is defined as the smallest unit of sound that influences the meaning of a word. For example the word *tree* has three phonemes /t/ /r/ /i/. If the first phoneme in tree /t/ was changed to /f/ the new word *free* would be heard. The separated phonemes are not heard during speaking, they are blended into syllables and individuals must be taught to perceive and understand that words comprise these individual sounds (Gillon, 2004; Truxler & O'Keefe, 2007).

According to Gillon (2004), the instruction of this phoneme awareness takes place through various methods, and some research speaks of a progression from syllable awareness, to onset-rime awareness to phoneme awareness, however that progression is not confirmed and requires further research. It would appear that research has established that phoneme awareness is critical for word recognition and written tasks to facilitate expression and meaning. Therefore, instruction in phoneme awareness must take place (Gillon, 2004; Truxler & O'Keefe, 2007).

Children develop phoneme awareness early in childhood by interacting in language rich environments. Nursery rhymes, children's songs and conversations automatically teach children about the sounds of language and how sounds are manipulated to form different words (Schumm, 2006). Some children do not just automatically grasp the understanding and manipulation of sounds in language. That may be due to lack of exposure to language rich environments or other physiological barriers such as sensory and perceptual difficulties.

Yopp (1988) suggests that phonemic awareness should be taught in a flexible manner and not as a progressive, step-by-step process. Rhyming, syllabification and on-set rime manipulations can be mixed and offered to children during sing-song time, play time, movement time and so forth and offered in an informal manner for the best results (Schumm, 2006).

The instruction of syllable and onset-rime awareness is beyond the scope of this study. However, phoneme instruction is relevant to this study as it is an aim of the THRASS

Program to teach phoneme awareness and phonics. Phoneme awareness has, at times, been confused with the term phonics. Gillon (2004) describes phonics as a teaching method of sound-letter correspondence and differentiates it from phoneme instruction. He states that traditional phonics teaching pays little attention to increasing an individual's awareness of the sound structure of words, namely, that words consist of phonemes and that some phonemes are represented by a combination of letters and not just that one letter makes one sound.

Gillon (2004) suggests that integrating phonics instruction with phoneme awareness demonstrates the best outcomes for reading. Davies (2006), on THRASS, agrees with Gillon; as right from the onset of the THRASS Program the 44 phonemes of the English language are taught (Davies, 2006).

As this study is investigating word level reading, namely, accuracy or word recognition, it is relevant to explore the way in which children recognise words. I have alluded that children use the phonemic structure of words as one strategy for word recognition. However, it is acknowledged that children recognise some words by sight. The combination of phoneme decoding, visual perception (grapheme awareness) and memory represent steps that are necessary in the process of learning to read (Share & Stanovich, 1995; Gillon, 2004). Word recognition models will, thus, be explored further in the following section.

## 2.6 WORD RECOGNITION MODELS

### 2.6.1 INTRODUCTION

*How does the awareness that a spoken word is comprised of smaller units (phonemes and graphemes) contribute to an individual's ability to accurately read and comprehend a connected text or to spell words correctly (Gillon, 2004:13)? That is a question posed by the author and is of relevance to this study.*

Word recognition is fundamental to the reading process and has a relationship with a child's phoneme awareness, grapheme awareness as well as an influence on the eventual meaning that is extracted from a text (Gillon, 2004; Gough & Tunmer, 1986; Perfetti, 1985; Stanovich, Nathan & Zolman, 1988).

Word recognition is defined as the association of a printed word with its spoken equivalent (Share & Stanovich, 1995). Word recognition theory is dominated by two main models of word recognition, namely, the dual-route model and the connectionist model. However, there

are related models such as the moairied dual-route model as well as the analogy model (Gillon, 2004).

The aforementioned models are briefly discussed within the context of their influence of the recognising and processing of words at the word level and the conceptualisation of the process of assisting a reader to translate the spoken word into its printed form. Those models will then be related to the THRASS Program and how the program incorporates some of those models to move towards constructive word level instruction.

### **2.6.2 THE DUAL-ROUTE MODEL**

The dual-route model suggests that there are two routes to recognising a word and processing its meaning; one is a phoneme route and the other a visual route.

According to Gillon (2004) a series of sub skills are required for the phoneme processing step to occur. The first is a skill known as graphemic parsing, which involves the analysing of the string of letters in their printed form that correspond to a single phoneme for example /sh/ two letters are parsed into a single phoneme. Following this, letter/sound translation rules are applied (grapheme-phoneme conversion). This allows for the phonology of the word to be accessed.

For word recognition to take place various cognitive sub-processes need to occur concurrently, for example, maintaining phonemic codes in working memory, assembling the phonemes into a complete phonemic representation and, finally, attaching meaning to the word (Gillon, 2004). For words in the English language that do not conform to regular grapheme parsing or grapheme-phoneme conversion rules, namely, irregular spelled words for example *sword*, the dual-route method purports that the word cannot be decoded using phonemic cues and phonics but must be visually encoded to recognise the word and, in turn, access its meaning (Gillon, 2004). According to Gillon (2004), the orthographic shape of the word, letter cues and legality of letter patterns may be used to access the orthographic representation of the word in memory, for example, recognising what the word looks like based on past experience and memory.

The THRASS Program teaches graphemic parsing and allows for grapheme-phoneme conversion by teaching children the 44 phonemes and 120 graphemes of the English language, one of which is the /sh/ phoneme. The THRASS Program does not term this graphemic parsing or grapheme-phoneme conversion but refers to it as phoneme-grapheme or phonographic instruction (Davies, 2006).

### 2.6.3 THE ANALOGY MODEL

The analogy model suggests that children learning to read access stored pronunciations of words and/or parts of words and use that knowledge when recognising similar words. The model proposes that children don't necessarily map each individual phoneme and its corresponding letter/s (graphemes) but make use of similar onset-rime patterns, for example, in recognising the word *fat* because of its spelling (grapheme usage) and phonemic similarities to known words such as *bat*, *cat* and *mat* (Fox,2000).

The THRASS Program makes use of spelling boxes which comprise the various graphemes used to spell the various phonemes. For example the /b/ phoneme can be spelt with a 'b' as in the word 'bird' or 'bat' and it can be spelt with a 'bb' as in the word 'rabbit'. The children are thus taught to store those two options for the /b/ phoneme in their memory, to assist them when recognising various words that contain the /b/ phoneme.

### 2.6.4 THE CONNECTIONIST MODEL

The connectionist model stresses the importance of phonemic knowledge in word recognition and proposes that both regularly and irregularly spelt words are processed through an interconnected system of orthographic, phonemic and semantic knowledge. *The connectionist model proposes that the relationships between spoken and written words are learnt through distributed patterns of activity represented by orthographic, phonemic and semantic processors* (Gillon, 2004:21). The understanding of the relationship between the spoken and written word takes place through *excitatory* and *inhibitory*<sup>14</sup> interactions among orthographic, phonemic, and semantic units, namely, the understanding of the connections between letters in words, the speech sounds those letters represent and the individual's vocabulary knowledge (Gillon, 2004).

To illustrate those interactions it is noted that at the early stages of learning to read, the /sh/ sound in the word *shop* may not be recognised and therefore not pronounced accurately. That may be due to limited phonemic knowledge and the many phonemic representations of the /sh/ sound. However what may take place is the *excitement* of just the /s/ sound, which results in inaccurate pronunciation. As the individuals phonemic knowledge improves and connections between grapheme and phonemes are strengthened due to the learning

<sup>14</sup> 'Excitatory' and 'inhibitory' are the actual words used by the author referenced. 'Excitatory' sounds/phonemes are the sounds that are provoked from memory and with experience and time the more appropriate sound is provoked from memory more often than other related sounds. The sounds that are inhibited are those sounds that initially would be provoked and then become those that are not related to the appropriate sound, through the process of learning and experience and are then inhibited from emerging from memory.

process (e.g. phonics teaching, syllable, onset-rime teaching and grapheme teaching), only connections from orthographic patterns close to the phonemic representations in words like *ship*, *shoe* and *show* will be activated or *excited*. With continued learning and experience other connections that are not related to the /sh/ sound will be *inhibited* (Gillon, 2004).

The connectionist model suggests that over time, with increased instruction and with increased practise, making connections between a particular letter or letter string and its phoneme and graphic form will result in stronger connections of sound/symbol relationships. That leads to the understanding that one phoneme can be represented by more than one grapheme e.g. the /f/ sound can be spelt with an *f* or a *ph*, depending on the word. As has been previously discussed, the THRASS program teaches connections between letters, letter strings and their corresponding phonemic form and, therefore, is supported by the connectionist model of word recognition.

#### **2.6.5 WORD RECOGNITION MODELS – CONCLUSION**

The models that have been discussed thus far suggest theories as to how individuals recognise words in isolation. Seldom are words found in isolation; they are more frequently joined with other words in the form of sentences, paragraphs, pages and books and are, therefore, placed within a context. That context allows the individual to use cues that surround the words to recognise unfamiliar words, i.e. other words or illustrations to assist in word recognition and meaning. That model of word recognition is referred to as the connected text model (Gillon, 2004). Those methods fall outside the scope of this study, as the THRASS Program is a word level program which does not make use of connected text.

This section has briefly discussed the relevant word recognition models that pertain to the word level instruction that the THRASS Program aims to address. It has become apparent that developing both auditory and visual skills and integrating the two are important for word recognition and word level reading accuracy. That process has been made evident by exploring the theory around the dual route model of word recognition. It has also been noted that children recognise similar sounds and symbols and use the similarities as analogies for recognising similar sounds and letters and letter combinations in words. That is evidenced by the analogy model of word recognition. Finally, through continued and repeated instruction and practise children begin to understand the connections between speech sounds and the letters that represent them.

The THRASS Program reflects the core elements of those models and it has been discussed that the THRASS Program incorporates the core elements of some of the word recognition

models and, therefore, makes theoretical sense. The way the THRASS Program attempts to facilitate that understanding of letter sound correspondence has been briefly discussed under Section 2.4 of this chapter, however it is still necessary to relate the THRASS method to instruction to the theory that pertains to reading instruction. That relationship will now be discussed in the following section.

## 2.7 FORMAL READING INSTRUCTION THEORY

### 2.7.1 INTRODUCTION

There are many perspectives on the teaching of reading, some of which have sparked major debate among researchers and theorists of reading instruction across the world (Chall, 1996). That debate rages on and concerns two opposing perspectives in reading instruction, namely, the bottom-up approach and the top-down approach.

### 2.7.2 THE BOTTOM-UP APPROACH TO READING INSTRUCTION

The bottom-up approach is based on the idea that learning to read is a process of being taught to break the code of reading in a systematic way, namely, the idea that good readers are good decoders (Schumm, 2006). A significant approach for the bottom-up perspective is referred to as synthetic phonics.

*Synthetic phonics emphasises part-to-whole instruction, because children are taught sounds in isolation and are then asked to blend the sounds together to form words, for example if children have been taught the sounds /b/, /a/ and /t/ they may be asked to blend them together to form the word /bat/ (Schumm, 2006:120; Wyse & Styles, 2007).*

Most reading practise and instruction initially takes place out of context, however later on books and passages are used to provide exposure to particular sounds and patterns being taught. Naturally occurring text is not typically emphasised during early bottom-up instruction (Schumm, 2006). However, that approach is regarded as the prime approach to establishing word recognition (Wyse & Styles, 2007).

The THRASS Program makes use of synthetic phonics in its approach to phonics teaching, however it expands upon that idea and teaches that one letter makes more than one sound, for example 'c' makes more than one sound: /c/ in the word *cat* and /s/ in the word *city*. The THRASS Program, therefore, teaches sounds in isolation and the blending of sounds but, unlike synthetic phonics, it does attempt to teach all the sounds a letter can make.

### 2.7.3 THE TOP-DOWN APPROACH TO READING INSTRUCTION

The top-down approach, also known as the whole language approach, is a philosophy which holds that learning to read is similar to the acquisition of natural language. Children learn to read by exposure, reading along with the teacher, and by guessing words using the context, pictures, and other cues (Gillon, 2004; McGuinness, 2004) and, therefore, the top-down approach views reading as primarily a meaning making activity influenced mainly by the reader's evolving knowledge of the spoken word. The supposition that good readers are good meaning makers informs this perspective.

The top-down approach utilizes analytic phonics, namely, whole-to-part-to-whole instruction (Schumm, 2006). That approach involves giving children whole words with similar patterns and then teaching the children to make generalisations about the parts of words that are similar. For example, when teaching the children the word *bat* they are reminded of the beginning of the word *boy* and the ending of the word *cat*. They are then encouraged to recognise the similarities and asked to join the /b/ sound of the word *boy* and the /a/t/ sounds of the word *cat* and make the new word *bat*. With the analytic approach the words are introduced in context and children make use of contextual cues as well as naturally occurring text. The idea is that as children become familiar with word parts they are able to generalise their usage and form other words with similar patterns (Fox, 2000; Schumm, 2006). THRASS makes use of analytic phonics in that it presents the child with the whole word first and then proceeds to break up the word into its relevant parts (phonemes) in the form of their graphical representations, namely, graphs, digraphs, trigraphs and quadgraphs.

### 2.7.4 FORMAL READING INSTRUCTION – CONCLUSION

Whilst there is still argument pertaining to which method is the best, it is clear that both the sound aspect and the visual aspect of learning to read, namely, the sound symbols relationship, must be addressed. THRASS is a program that appears to integrate aspects of the two approaches illuminated upon above and integrates the usage of both synthetic and analytic phonics, which appears to be in agreement with Schumm (2006) when he states that phoneme awareness activities have a great effect on success in reading if they are combined with letter/sound correspondence, namely, phonics instruction.

Whilst the teaching method is important in the facilitation of reading skills, continual assessment of progress and evaluation of the teaching method plays a crucial role in establishing those skills. The assessment of reading thus forms the subject of the following section.

## 2.8 ASSESSMENT OF READING

### 2.8.1 INTRODUCTION

There is no doubt that assessment is necessary for evaluating and monitoring an individual in terms of acquired skills and planning of individual needs, as well as for the evaluating of instructional programs. Schumm (2006) considers three broad questions for the assessment of reading, which will form the basis of this discussion:

1. What are the purposes of assessment and reading?
2. What expertise is necessary to conduct reading assessment and to analyse assessment results?
3. What instruments can be used for assessment?

### 2.8.2 THE PURPOSE OF ASSESSMENT

The initial purpose of reading assessment starts with a screening that helps to establish initial grouping of the student. Secondly, assessment can be used to identify the individual's strengths and areas in need of improvement in reading and writing. Assessments to address this need are referred to as diagnostic assessments (Hudson *et al.*, 2005; Schumm, 2006). Diagnostic assessment may address specific academic areas as well as other factors, like cognitive, neurological, physical, educational, language, social/emotional and family factors. Thirdly, ongoing monitoring of the individual's progress is facilitated by assessment (Foxcroft & Roodt, 2001). The main purpose for this is to inform instructional method, both in terms of effectiveness of method and knowing whether changes need to be made to the process. The fourth purpose of assessment is for determining the individual's grades or marks, promotion or placement in a special needs environment, namely, remedial classes or a specialised education setting. Standardised tests are often used for that type of assessment to determine grade-level promotion (Schumm, 2006). The fifth and sixth purposes for assessment are based on evaluation and accountability issues. Student outcomes can be used to evaluate the effectiveness of instructional programs and long term trends in reading performance (Hudson *et al.*, 2005; Schumm, 2006).

For this study the purpose of assessment is in line with the fifth and sixth purposes listed above. This study has made use of assessment results of a group of children, to evaluate the effectiveness of the THRASS Program in developing word level accuracy skills and improved word recognition skills.

## 2.8.3 EXPERTISE NECESSARY TO CONDUCT READING ASSESSMENTS AND ANALYSE RESULTS

It is important for a teacher, counsellor, therapist or test administrator to have the competencies necessary when assessing children. According to the international reading association, adhering to a code of ethics is encouraged for ethical practise to occur. The aforementioned association's code of ethics states:

*Professionals in reading must possess suitable qualifications for engaging in consulting, diagnostic, or remedial work. Unqualified persons should not engage in such activities except under the direct supervision of one who is properly qualified. Professional intent and the welfare of the person seeking the services should govern all consulting or clinical activities such as counselling, administering diagnostic tests, or providing remediation (International Reading Association (1996) Code of Ethics, 2008. [www.reading.org/association/about/code.html](http://www.reading.org/association/about/code.html)).*

A qualified teacher is able to administer the NARA, providing the teacher is familiar with the administration procedures which are set out in the manual for the NARA (Neale, 1997:17). The following considerations have been given for the test user in interpreting the data received from using the NARA (Neale, 1997):

- Changes in child growth and development require time and must be taken into account in any interpretations of development.
- Age data represent the general or average trend for a particular aspect of behaviour observed from large numbers of children and considerable variations occur among individuals in attaining different stages of development.
- Information processing is dependent on the integrity of the individual's central nervous system, namely, senses, motor co-ordination, control and speed of execution.
- Language and communication skills meet personal and social needs and are subject to elaboration through use, instruction, and social intercourse.
- Child rearing practices and socialisation influence the pattern of social and emotional growth but, increasingly, the interests, pursuits, and information processing styles of the individual play a significant role.
- Adaptation and learning constitute a fourth system in child development, interweaving with physical development, linguistic development, and social and emotional maturation.
- The foregoing processes are interwoven according to the individual's maturational tempo. Auditory and visual signs are synthesized in language. In reading they must

be synthesized speedily according to arbitrary, given, conventional sequence, and must take into account varying styles of print.

- The motor functions of the central nervous system for processing patterns of light and sound, and for selecting and discriminating between similar patterns of stimuli are individual and individual differences are inevitable.

Many variables are at play when considering the development of reading ability in children. Each child has his or her own set of variables – genetic abilities as well as those that have been developed through social interaction. Those have been considered in the development of the NARA and are reflected in the norms of that measure.

#### **2.8.4 INSTRUMENTS FOR ASSESSMENT**

Assessment instruments can be categorised into three main groups: norm-referenced tests, criterion-referenced tests and informal tests.

- **Norm-referenced tests**

Norm-referenced tests compare the individual student's performance with that of his peers of the same age and/or grade level. When those measures are developed the test is given to a sample or norming population of students. Norm-referenced tests can normally be administered in a group or individual basis and are typically used for instructional programme evaluation as well as screening tools (Schumm, 2006).

- **Criterion-referenced tests**

Criterion-referenced tests assess the student's mastery of specific skills or goals. The results tend not be applicable across populations and more closely linked with curriculum outcomes or specific individual skills or competencies. The individual's results are thus not compared to a norm group but to the criteria being measured (Foxcroft & Roodt, 2001). Criterion-referenced tests are mostly used to inform individualised instruction, grouping for instruction and individual performance.

- **Informal tests**

Informal assessment tools are critical for an all-inclusive assessment program. The administration and results tend to be flexible for informal measures and usually include both teacher and individual evaluation of reading. Of importance is that the measures chosen, adapted and adopted generate the kind of information that is relevant for an organisation's vision for reading and writing (Schumm, 2006).

For the purposes of this study, a norm-referenced test was used, namely, the NARA. The selection of a norm-referenced test is due to the aim of this study being to evaluate the THRASS instructional program and its influence on the word level accuracy skills of a group of Grade 2 learners. Word recognition, as has been discussed, includes visual skills, contextual skills as well as decoding skills. THRASS is referred to as a word-level phonics program and thus I have used the accuracy sub-scale on the NARA to ascertain the effect the exposure to the THRASS Program has had on the word level reading of that group of Grade 2 learners.

#### 2.8.5 THE NEALE ANALYSIS OF READING ABILITY-REVISED (NARA)<sup>15</sup>

The NARA was developed by Marie D. Neale in the 1950's and has, since that date, undergone several revisions, the latest being in 1997. Overall 2000 children have been individually tested to obtain the normative data. The selection of passages for the assessment measure were carefully written and graded, by controlling the vocabulary, the syntactic complexity, and the length of the narrative. The selection of words in the original passages was based on word lists and word frequency counts (Neale, 1997). In the development of the revised Australian edition, books containing high frequency words were consulted to ensure that the words chosen were in current use by children. They were also checked against contemporary basic instructional schemes of reading (Neale, 1997). Several procedures were adopted to ensure that each passage is of appropriate difficulty, and that they discriminate satisfactorily between children of differing abilities. The analysis of the Australian standardisation sample on the revised forms attempted to reflect the variations of difficulty associated with each passage (Neale, 1997).

For this study the accuracy scale of the NARA has been used. The assessment measure makes use of miscue analysis and assesses accuracy by recording the child's errors. The term *error* is reserved for inaccuracies in reading which some current researchers refer to as *miscues* (Neale, 1997). In the NARA those errors are viewed as the frequency count of difficulties that are exhibited in a child's oral reading. *They are used for normative purposes to obtain an objective measure of the accuracy with which a child recognizes words* (Neale 1997:8). The categories of errors are defined for qualitative use by test administrators for graphophonic, semantic and syntactic error analysis. The additional categories of errors are defined as: Mispronunciations, Substitutions, Refusals, Additions, Omissions and Reversals.

---

<sup>15</sup> The validity and reliability of the NARA was discussed in Chapter 3, Section 3.5.

- **Mispronunciations**, or decoding errors, are words that are wrongly pronounced or distorted and only partially decoded. They afford information on the way the child attempts to decode the main features of the word.
- **Substitutions** are real words that are used instead of the word in the narrative e.g. ‘they’ for ‘then’, ‘her’ for ‘here’, ‘realised’ for ‘released’ etc.
- **Refusals** are words which the child fails to make an attempt at reading or does not recognise and are unable to attempt.
- **Additions** are where words or parts of words are inserted into the text.
- **Omissions** are errors where words or parts of words are omitted from the text.
- **Reversals** are technically substitutions but are recorded separately because of their diagnostic value, e.g. ‘on’ for ‘no’ etc.

The analysis of those errors is asserted as a primary way for acquiring insight into the reading strategies children use. The essential aim for the analysis is to gain some understanding of any patterns that may be occurring at the graphophonic, syntactic or semantic level. The view of the NARA suggests that the more proficient the reader the fewer mistakes will be made (Neale, 1997). The accuracy sub-scale of the NARA is of relevance for this study. Accuracy of reading will be discussed in the following section<sup>16</sup>.

#### 2.8.6 ASSESSMENT OF READING ACCURACY

Word-level reading accuracy refers to the ability to recognize or decode words correctly. Thorough understanding of the alphabetic principle, the ability to blend sounds together (Ehri & McCormick, 1998; Hudson *et al.*, 2005), and knowledge of a large bank of high-frequency words is required for word-reading accuracy (Byrne & Fielding-Barnsley, 1989).

When words cannot be read accurately from memory as sight words, they must be analyzed. In order to accurately decode words,

*... readers need to be able to accurately (a) identify the sounds represented by the letters or letter combinations, (b) blend phonemes, (c) read phonograms (common patterns across words), and (d) use both letter–sound and meaning cues to determine exactly the pronunciation and meaning of the word that is in the text (e.g., knowing how to correctly pronounce bow in two different sentences: The dog had a bow tied around her neck. The bow of the ship was tall) (Hudson *et al.*, 2005:703).*

---

<sup>16</sup> Reliability and validity of the NARA was discussed in chapter 3, Section 3.5.

Assessment of children's word-reading accuracy can take various forms. Listening to oral reading and counting the number of errors per 100 words can provide invaluable information for the selection of appropriate text for various instructional purposes for an individual or group of students. A running record and miscue analysis provides more detailed information about the child's accuracy (Clay, 1984 & 1993; Hudson *et al.*, 2005). Through careful examination of error patterns, a teacher can determine which strategies the child is utilising and which strategies the child is failing to use. Observing a student's attempts to figure out an unknown word may yield evidence of phonemic blending, guessing based on context, or a combination of decoding and contextual analysis. Those observations may provide information about areas in need of further instruction to improve word-reading accuracy.

As established previously, the purpose of this study is to investigate both the number as well as the type of errors made on the pretest and posttests using the NARA, which were administered one year apart, to investigate the possible effect the THRASS Program may have had on the word level accuracy skills of the research participants.

## 2.9 CONCLUSION

The current literature relating to aspects of reading development, reading instruction and reading assessment relevant to this study was reviewed in Chapter 2. That it is beneficial for children to establish an awareness of the phoneme structures of words before formal reading instruction is to take place was discussed, as was the need for the sound/symbol relationship to be taught. It was established that integrated approaches combining synthetic and analytic phonic approaches that involve the orthographic processor (letter forms and sequences), the phoneme processor (individual phonemes), meaning processor (prior knowledge and vocabulary), and the context processor (ongoing construction of the meaning of text) are most beneficial for children learning to read. It would appear the THRASS Program does integrate those components.

The investigation of the possible effect that the THRASS Program has on word level reading that is the primary concern in this study. Through the assessment of the reading accuracy of children, it is the aim in this study to investigate the influence of the THRASS Program on the word level reading of a certain group of Grade 2 children. In Chapter 3 the method and methodological approach used to test that effectiveness will be considered.

---oOo---



### 3.1 CHAPTER OVERVIEW

The methodological approach undertaken for this study is outlined in Chapter 3. Due to the nature of the quantitative methodological design the ontological and epistemological paradigm was positivist in nature. This research study utilised performance-based assessment to track the progress of accuracy in reading of Grade 2 learners following their exposure to the THRASS Program. The collection of data, reduction of that data, data display, and data analysis are described, to form a concise understanding of the presentation of the research results. Descriptive and inferential statistical analysis of results – which includes the employment of parametric statistical analysis, such as variability, differences and comparative deviations from the average – were obtained and evaluated. The validity of the findings, measurement validity and validity of results of the study are then explored; followed by a discussion on the reliability of the study. To conclude, the threats to internal and external validity of the study are also examined.

### 3.2 RESEARCH PARADIGM

Paradigms are all encompassing systems of inter-related practice and thinking that define for researchers the nature of their enquiry along three dimensions: ontology, epistemology and methodology (Terre Blanche & Durrheim, 1999). Ontology specifies the view taken of reality that is to be studied and what can be known about it. Epistemology specifies the nature of the relationship between the researcher and what can be known, and the methodology refers to how the researcher will practically go about studying whatever he believes can be known.

As was mentioned in chapter one it is acknowledged by the researcher that the research perspective or paradigm within which research is conducted usually informs the methodology of the research. However in this specific study the research was conducted *ex post facto*, after the assessments results were produced. The research paradigm for this study was thus informed by the research methodology, which takes the form of statistical evaluation of the results obtained from the pretest and posttests administered. The data collection method is positivist in nature using quantitative data.

The positivist view is described by referring to the goal of knowledge; that is, to describe the phenomena that we experience. And the purpose of science is to describe the phenomena in

terms of what we can observe and measure and that the truth or reality is governed by and limited to the observation and measurement methods. The aim of the positivist researcher is to see science as a way to get to the truth and to understand it well enough to be able to predict and control it. The view is often considered deterministic and governed by the law of cause and effect. Postulated theories are based on deductive reasoning derived from controlled experiments and tests (Trochim, 2006).

Through positivist foundations, in order to assess the effectiveness of learning,

*external actions - what students could actually show they could do - became the focus of assessment. From observation of these external actions, deductions have been made about the effectiveness of internal actions that underlie them (Donald et al, 2002:120).*

For the purpose of this study a pre-experimental design, namely, a one-group pretest-posttest design, was applied, which consisted of quantitative data collection/analysis methodology. The aim was to test the hypothesis to identify and explain possible influences the exposure to the THRASS Program had on progress in word level reading accuracy (Terre Blanche & Durrheim, 1999). Those elements of the research process are discussed as part of the section to follow.

### **3.3 RESEARCH DESIGN**

Research design is the general plan for setting up and testing a specific hypothesis or research question (Thompson & Panacek, 2006). There is often no single best design to answer a research question. Rather, there are usually multiple research designs that can be used to approach a given research question, each presenting its own unique advantages and disadvantages.

It is important to clarify the aim of one's research, as that often informs the most appropriate design for a research study. The aim of this research was to start a process of research i.e. to start generating ideas on how the THRASS Program can be researched in the South African context and, more specifically, arrive at some conclusions on the possible efficacy of the THRASS Program in this specific private school (McMillan & Schumacher, 2001).

With this aim in mind it was decided to use a pre-experimental design in the form of a pretest-posttest design. Pre-experimental designs are termed as such as they are without two or more characteristics of those necessary for experimental research, namely, manipulation, control, randomization, independent variable, dependent variable and

extraneous variables. As a consequence of this few threats to internal validity are controlled. However this does not mean *the designs are uninterpretable and shouldn't be used as there are certain cases in which threats can be ruled out on the basis of accepted theory, common sense, or other data* (McMillan & 2001:304).

McMillan and Schumacher (2001) state that pre-experimental designs are such that they fail to rule out rival hypotheses and, as such, that increases the difficulty of making reasonable causal inferences. It should be noted that it was never an intention of this study to make any such inferences. The aim of this study, as mentioned previously, was to generate ideas for further research, which is in agreement with McMillan and Schumacher's statement that *pre-experimental designs are best used, perhaps, as a way of generating ideas that can be tested more systematically* (McMillan & Schumacher, 2001:304).

In the presentation of this design the following notational system will be used to provide information for understanding of the design:

R	Random assignment
O	Observation, a measure that records observations of a pretest or posttest
X	Treatment conditions (subscripts 1 through n indicate different treatments)
A, B, C, D, E, F	Groups of subjects, or, for single subject designs, base-line, or treatment conditions (McMillan & Schumacher, 2001:304).

In the one-group pretest-posttest design a single group of subjects is given a pretest (O), then the treatment (X), and then the posttest (O). The result that is examined is the changes recorded from pretest to posttest. While the researcher can at least observe a measure of change with this design, it must be acknowledged that there are still various rival hypotheses that may be applicable (McMillan & Schumacher, 2001:304).

The one-group pretest-posttest design for this study was thus:

Group	Pretest	Treatment	Posttest
Grade 2 learners at a private school	NARA Accuracy score January 2007	THRASS Program for 12 months	NARA accuracy score January 2008.
A	O	X	O
→			
Time: January 2007 to January 2008			

The table was adjusted according to McMillan and Schumacher (2001:305)

Donald *et al.* (2002), and Foxcroft and Roos (2001) suggest that in order to track functioning and progress effectively it is important to conduct performance-based assessment in conjunction with the recording of systematic observation, which process enables a process of accurate review. It is due to that assertion this study includes quantitative data collection methods and subsequent systematic statistical evaluation of the results. The performance based assessment was conducted with the norm-referenced testing measure, namely, the NARA. A pretest took place before the THRASS Program was implemented, with the results being statistically compared with the posttest results obtained one year later, applying the same norm referenced test. I administered both tests personally. The results were analysed and averages established.

### **3.4 METHODOLOGICAL PARADIGM**

The research methodology is quantitative in nature. The advantage of a quantitative data collection method would be the *demonstration that accurate and observable progress could be statistically demonstrated by means of comparison* (Neuman, 1997:56). Quantitative data has been collected by means of a norm referenced test. Terre Blanche and Durrheim (1999) specify that quantification, conceptualization and operationalisation are critical for valid and reliable quantitative measurement to take place.

Quantification is referred to as the process of assigning numerical value to an individual's experience or variable. In this study the variable of word level accuracy was quantified. Reading accuracy has been conceptualised in the literature as *the ability to recognize or decode words correctly* (Hudson *et al.*, 2005:703) and operationalised through the use of a measurement instrument, deriving accuracy scores through the NARA. The accuracy scores have been utilised for this study as they pertain to the research question of the influence of the THRASS Program on the word level accuracy skills of the group of Grade 2 children. The accuracy scores were quantified by recording the number of errors an individual makes on each of the reading passages given as part of the assessment measure. This score is then tallied at the end of the administration of the measure and an accuracy raw score is produced. The pretest and the posttest results of the accuracy scores, as well as the pretest and posttest error category scores have been statistically analysed, to indicate and validate an increase or decrease in the specific construct of reading accuracy as measured by the NARA (Neale, 1997).

#### **3.4.1 QUANTITATIVE MEASUREMENT VALIDITY**

Quantitative measurement validity is the degree to which the measure does what it is intended to do. In other words, there must be a 'good fit' between the conceptual and operational definitions of the construct, and the instrument should be usable for the particular purpose for which it was intended (Balnaves, 2001; Tashakkori, 2008). In this study it is intended to measure reading accuracy (conceptual definition) and the possibility of a change in reading accuracy from the pretest and posttest analysis (operational definition). It is acknowledged that the purpose of this measurement is to observe whether there was a significant increase in the reading accuracy of the target population within a period of one year and not to generalise the results to populations in other contexts. The purpose of this measurement was to be descriptive of the Grade 2 population at a specific school and not descriptive of other school populations. Terre Blanche and Durrheim (1999) indicate three kinds of measurement validity: criterion-related validity, content validity and construct validity.

Criterion-related validity is the degree to which a measure is related to some other standard or criterion that is known to indicate the construct accurately. This form of validity is established by comparing the measure with another measure of the same construct (Balnaves, 2001). Criterion-related validity has been established for the NARA and considers comparing measurement instruments and the measured variable, which is discussed in the section that follows.

Content validity is determined by establishing the extent to which a measure reflects a specific domain of content (Neuman, 1997). Content validity is specifically important for tests of knowledge. It has been assumed that there is a close relationship between the concept of reading accuracy and the operational measurement of accuracy on the NARA and, thus, it is felt that content validity has been established with the analysis of the results of the data.

Construct validity of a measure involves both the theoretical and empirical task of determining the extent to which a measure of a construct is empirically related to other measures with which it is theoretically associated (Terre Blanche & Durrheim, 1999). In establishing construct validity, a relationship between theoretically associated constructs is sought and

*... involves three steps (a) specify a set of theoretical relationships between constructs; (b) test those hypotheses empirically; (c) interpret the pattern of the relationships in terms of how they clarify the construct validity of the measure* (Terre Blanche & Durrheim, 1999:87).

From the literature it is apparent that the THRASS Program develops both phonemic awareness as well as sight word vocabulary. Accuracy of reading is related to phonemic

awareness and sight word vocabulary. Therefore the constructs addressed by the THRASS Program and indicated by reading accuracy appear to be related. The pretest and posttests tested the hypothesis: by exposing a group of Grade 2 children to the THRASS Program there will be a significant difference observed between the average accuracy age observed on the pretest and the average accuracy age observed on the of posttest; which was indicated in the findings of this study. The results were then analysed to establish the extent of the relationship between exposure to the THRASS Program and reading accuracy, thus indicating construct validity (Terre Blanche & Durrheim, 1999).

Reliability refers to the dependability of the measurement instrument, which is the extent to which the instrument yields the same results on repeated trials (Tashakkori, 2008). The reliability of NARA has been established and will be discussed further in the following section. The validity and reliability of the findings observed in this study have been described in the section that follows.

### **3.5 THE VALIDITY AND RELIABILITY OF THE NARA**

In the evaluation of a test, it is important to know how useful and meaningful the scores, which the test generates, are; that is referred to as the concept of validity. The various means for demonstrating validity have been divided into three categories – content related, criterion related, and construct related evidence (Neale, 1997).

Content validity relates to the extent to which the task content of a test can be a representative sample of the skill or ability the test is designed to measure and how closely the behaviour demanded by the test resembles real life activity in that specific skill (Neuman, 1997). The NARA requires the child to read out loud and answer comprehension questions; those are considered important aspects of everyday reading and, thus, the NARA can be considered valid. To further ensure validity, steps were taken at the developmental stage of the test to select age appropriate reading passages (Neale, 1997).

Criterion related validity of a test is attributed to the relationship a test has to other assessment measures that claim to measure the same skill (Balnaves, 2001; Neuman, 1997). That type of validity is often separated into concurrent and predictive validity. That indicates how well the scores on a test relate to performance on similar tests at the same time (concurrent) or at a later date (predictive). Evidence for concurrent validity of the NARA have been gathered by correlating its subtest scores with scores from other tests which are accepted as equivalent measures of reading and verbal ability (Neale, 1997).

The predictive validity of the NARA has been well documented. With respect to the Australian revised edition it was found that the results given at the end of year one correlated significantly on all three scores with reading achievement by the same sample of children tested at the end of year two (Neale, 1997).

Construct-related *validity of a measure involves both a theoretical and empirical task of determining the extent to which a measure of a construct is empirically related to other measures with which it is theoretically associated* (Terre Blanche & Durrheim, 1999). The construct that is measured by the NARA is reading ability (Neale, 1997). For developmental tests like the NARA, increases in scores with age can be taken as evidence of construct validity. The consistent rise in average scores over the age groups shows that age differentiation of reading ability has been achieved (Neale, 1997).

The reliability of the NARA has been reported on three main types of reliability: stability (parallel forms) reliability, internal consistency, and standard error of measurement of the test.

Stability reliability is considered to be the most important kind of reliability, which is usually determined by the test-retest method. This was made possible by the availability of parallel forms in the NARA and made it possible to report coefficients of stability that are not contaminated by practise on the same test (Neale, 1997). Internal consistency of a test estimates the extent to which its items appear to be measuring the same skill (Neuman, 1997). A generalised alpha coefficient was used for continuous scores on the accuracy scale. Tables reflecting the internal consistency coefficients are reflected in the manual of the NARA.

All educational measurement has some degree of error (Neale, 1997). Estimates of errors made can be used to describe the range of scores the individual would obtain if given the test again. The standard error of measurement is a reflection of the consistency of performance in that it shows the extent of variation due to test error for each score achieved (Neale, 1988). For the NARA the standard error of measurement has been calculated using the parallel form reliability coefficients and the standard deviations of the raw scores for the various components of the test. The standard errors of measurement in raw score units associated with the three aspects of reading for each form across the entire range are reported in the NARA manual (Neale, 1997).

In South Africa the NARA is used by numerous institutions as a means to assess reading. It has, furthermore, been used as part of numerous studies in South Africa. One such study

was to consider the relationship between the Senior South African Individual Scales – Revised (SSAIS-R) subtests and reading ability. That study was published in 2008 and it investigated the relationships between different cognitive functions as measured by the selected sub-tests of the SSAIS-R, namely, vocabulary, similarities, story memory and coding. Moderate, significant correlations were found between the Vocabulary and similarities sub-tests and components of the NARA (Cockcroft & Balckburn, 2008)<sup>17</sup>.

The unique attributes of this study include combining the results of a group of pupils who have all been exposed to a new independent variable i.e. the THRASS program. Observations taken from the data produced on the accuracy scale of the NARA have been statistically analysed according to the norms on the NARA for the children within the sample.

### 3.6 SAMPLING

The participants were selected from a school based on the East Rand. It included the entire population of Grade 2 learners. Those learners are demographically of mixed race and gender. Their ages ranged from 7 years 3 months to 8 years 7 months on the pretest, and 8 years 3 months and 9 years 7 months on the posttest. Convenience as well as purposive sampling was applied.

Convenience sampling was used to find participants in this study, as they were found on the premises where I am employed, which made them easily accessible. The reading tests that were carried out are carried out annually and were, therefore, available as the source of data. Permission was obtained from the school authorities to use the data.

Purposive sampling was also used. De Vos (2000) described that form of sampling as a *purposeful, systematic method by which controlled lists of specified populations within geographic districts are developed and detailed plans are designed to recruit adequate numbers of cases within each of the targets* (De Vos, 2000:383). I identified the two Grade 2 classes at the aforementioned school as possible participants. The reason for this was that Grade 2 children can read to a standard that allows assessment with the NARA. They were, furthermore, already exposed to the THRASS Program in 2007 and, thus, it was convenient to use them. In contrast, the Grade 1 pupils were not able to read as yet and the Grade 3 pupils had already been exposed to two years of other forms of reading instruction. Therefore, it seemed more appropriate to use the Grade 2 group for this study.

### 3.7 PLANNING FOR DATA ANALYSIS

---

<sup>17</sup> The article mentioned can be accessed on [http://www.sabinet.co.za/abstracts/sapsyc/sapsyc\\_v38\\_n2\\_a9.xml](http://www.sabinet.co.za/abstracts/sapsyc/sapsyc_v38_n2_a9.xml).

De Vos (2000:48) describes the data analysis stage as the '*ponder-and-check stage*', as efforts proposed by Huberman and Miles (1997) indicates three associated sub-processes that may occur before, during and after data collection. The main steps that can be followed during analysis and interpreting qualitative data are outlined below.

### **3.7.1 DATA REDUCTION**

I have reduced the data from various data collection methods: data from initial pretest results from standardized media versus the data from standardized media received in the posttest. An error analysis has been undertaken based on the guidelines of the NARA. The data has been organised and listed according to individual scores for reading accuracy and an average reading age for the group has been established. The individual error categories have been tallied and have been compared in the data analysis section.

### **3.7.2 DATA DISPLAY**

The data display step involved thorough organisation of concise data. That included finding norm scores, tabulating averages and a statistical analysis of the standardised results. The display process adds to the creation of a foundation that stimulates further examination and contemplation of the data that can assist in simplifying the final procedure of drawing conclusions from the data.

### **3.7.3 DATA ANALYSIS**

Descriptive and Inferential analysis of the data were used. Descriptive analysis aims to describe the data by investigating the distribution of the scores on each variable and by determining whether the scores on the different variables are related to each other. In this study, descriptive analysis was done to gain an initial impression of the data.

The first type of descriptive statistics applied is called a frequency distribution, the aims of which was to summarise the scores or observations obtained on a single variable and represent them in the form of a graph and frequency table. It is a graphical representation of the number (frequency) of subjects which obtained a particular score on a variable. Graphic representations of the recorded reading ages for the sample scored in the pretest and scores from the posttest have been made in this study, to observe any pattern or difference. The purpose of using the frequency distribution was to gain information regarding the distribution of scores. Adding to the graphical representation of the data, the frequency distribution

allowed the researcher to describe the shape, variability and the central tendency of the distribution (Terre Blanche & Durrheim, 1999).

Measures of central tendency are estimates of the centremost score in a distribution and have been included as part of the descriptive data analysis. There are three different measurements of central tendency – the mode, median and average – and they are each represented by a single number to best represent the data collected for a variable (Gorard, 2003). This study has compared two averages; the first taken from the pretest and the second from the posttest. The differences observed from this comparison were further analysed statistically.

Calculations of the variance and standard deviations from the average form part of Chapter 4, in which the results of the measurements are considered.

*Measures of variability estimate the degree to which the observations for a variable are dissimilar to each other. They estimate the degree to which the scores spread out or vary from the measure of central tendency* (Terre Blanche & Durrheim, 1999:106).

Inferential statistical analysis of data allows the researcher the opportunity to discover possible relationships between variables and the real world, for the population and not only the sample (Terre Blanche & Durrheim, 1999). In an attempt to generate ideas for further research I have done some inferential statistical analysis. Gorard (2003) stipulates certain criteria for inferential statistical analysis to be made possible. Firstly the sample must be representative of the population; secondly certain assumptions about the sampling distribution must be satisfied in order to draw inferences about populations (Gorard, 2003). According to Terre Blanche and Durrheim (1999) inferential statistics have two purposes, (a) to estimate population parameters and (b) to test hypothesis.

The hypothesis for this study was: The THRASS Program significantly improves the word level accuracy skills of a group of Grade 2 learners over the period of one year; as opposed to: The THRASS Program does not significantly improve the word level accuracy of a group of Grade 2 learners over a period of one year. This was measured by comparing the average scores on the pretest with the average scores on the posttest and the difference analysed. Should a significant difference be found, it will be assumed that was due to the THRASS Program. However that difference could have been by chance or could have been attributed to other extraneous variables.

Inferential statistical analysis provides an estimate of the probability that there is no difference between the sample averages. If the probability is very small (usually less than 0.05 or 5%) then a conclusion can be made that there is a significant difference between the averages of the sample. The term significance means that the difference between our samples is of such a magnitude that it is unlikely to have occurred by chance (Gorard, 2003; Terre Blanche & Durrheim, 1999:118).

As I studied the averages revealed in this study, I realised the data is of a ratio nature, mathematical operations are possible and, thus, parametric statistics are appropriate for this data. As the population/sample size for this study is above 30 participants it can be assumed that the population distribution is normal in nature (Steyn *et al.*, 1994). However, as the population sample is still regarded as small, the decision was made to utilise distribution free tests - these do not depend on strict assumptions regarding the distributions and are namely the Friedman Test and the Kruskal-Wallis test (Steyn, *et al.*, 1994).

The aim of the Kruskal-Wallis Test is to accommodate the statistical analysis of two independent samples, namely, the two Grade 2 classes, allowing the comparison between the two samples. Once the results of the Kruskal-Wallis test revealed that the two samples are comparable and able to be used as a unit of data without any threat of teacher effect, the Friedman Test was employed. The Friedman Test accommodates the use of two sample groups associated to each other, namely, the pretest and posttest groups. M. Friedman developed that test to enable the statistical comparison of averages of both pretest and posttest groups and test the hypothesis related to the average (Steyn, *et al.*, 1994).

Following the employment of those statistical measures, the T-test was conducted to establish the degree of difference in the averages related to the pretest and posttest, and the probability of the degree of difference being statistically significant. The three statistical measures employed enable the testing of hypotheses and the determination of the statistical significance of the results obtained by the study. The results of these statistical tests are discussed in the following section. Once the data has been analysed it needs to be verified and certain conclusions made. There is a process which should be followed to ensure that what has been revealed by the data analysis is sound and truthful. That process will now be discussed.

### **3.7.4 CONCLUSION AND VERIFICATION OF THE DATA**

When planning or designing a research study the adequacy of the research must also be considered, i.e. the validity and reliability of the research must be maximised. The validity of the data will now be discussed (Terre Blanche & Durrheim, 1999).

### 3.7.4.1 Validity

Complete internal and external validity are exceptionally difficult to achieve and is often dependent on the type of research design. Internal validity is often achieved in studies where great control is exercised over the study conditions e.g. in a laboratory, however, that setting may well threaten the external validity of the representation of the findings outside that artificial environment. Yet, the converse is also often true, i.e. that if the external validity is achieved this often threatens the internal validity of the findings of the research. It is, thus, often the case that the researcher will be content with satisfactory internal validity and some degree of external validity (Terre Blanche & Durrheim, 1999).

When attempting to judge the internal and external validity of a study certain aspects can be taken into account. When judging internal validity, consideration of as many alternative explanations for the findings should be undertaken. The researcher should question whether the conclusions drawn could be explained by alternative explanations or rival hypothesis. External validity is somewhat easier to evaluate and usually three aspects of the design are scrutinised:

- (a) The research participants and their relationship to the population. The uniqueness of the population should be considered to ascertain if the findings made in respect of them are generalisable to other populations and if appropriate sampling techniques were used to ensure representativeness.
- (b) The Independent variables or manipulations should be examined to determine if they are highly specific or can be generalised to other operationalisations.
- (c) The findings on the dependent variable must be assessed to establish the ability to generalise to other similar variables. In the research design of this study attempts have been made to reduce threats to internal and external validity.

According to Tredoux (Terre Blanche & Durrheim, 1999) threats to internal validity include, firstly, the threat of co-varying events (history), i.e. that something else in the environment could have caused or been influential in the change that has been observed or described in the independent variable – in this case the word level accuracy skill of certain Grade 2 learners. Even though two classes and two teachers were used as the sample, teacher effect was statistically dismissed, which would then substantiate validity. Using the normative scores of the NARA for both the pretest and posttest allows for an accurate comparison, as

the internal and external validity of that measurement instrument has been established. For example, there is an expected change in reading accuracy to occur in one year, i.e. from the time the pretest was administered to the time the posttest was administered. Many variables have been considered in this expected change and that score was established on the normative scores of the NARA. Therefore, if there is a change over and above the norm scores of the NARA, there is a chance that the change may have been caused by the exposure to the THRASS Program. That inferential probability has been analysed and will be considered as part of the following section when the results will be discussed.

Secondly, the threat due to natural change or maturation must be considered. Children have a natural trajectory of cognitive development (Mertens, 2004; Robeck & Wallace, 1990). Any change or improvement in the reading accuracy scores for this study may have been due to maturation of the group as a whole. Once again that threat is controlled by the use of the NARA which is a standardised test.

The third threat is referred to as the reactive effects to participating in a study (testing). Testing and exposing children to the conditions of a testing situation can have an effect on the results of the study. The participants for this study were not singled out for any specific reason. They were not chosen on the basis of their reading ability. At the school the NARA is used annually and routinely to assess the children's reading skill. So the children had been exposed to the test at previous developmental levels, as well as the tester on numerous occasions and are thus familiar with the process. The test consists of two separate forms which include different texts, which are considered to be at the same developmental level, i.e. the norms for both forms have been established with no significant difference being detected. Form 1 was used for the pretest and Form 2 on the posttest. I, furthermore, ensured that ethically correct assessment practises were followed, to minimise the effect of the test situation on the participants (Foxcroft & Roodt, 2005). Therefore it is felt that the reactive effects of this study were minimal.

The reliability of the measurement instrument otherwise referred to as instrument decay is the fourth threat to internal validity. The validity and reliability of the measurement instrument is critical to the internal validity of the study – the NARA has been discussed earlier in this study and its validity and reliability has been established, therefore instrument decay as a threat to internal validity does not apply to the NARA.

The fifth threat identified is the statistical regression of extreme subjects (regression to the average). That can occur if the participants chosen for the study are at the extreme of the criterion measure e.g. if one chooses a sample of fifteen readers who performed the worst on

a particular measure and used them for a study an error would be made, as there is a likelihood that there will be a positive change in the absence of any intervention. The reason provided for that is that all scores exhibit a random variability and those scores at the bottom of such a set are there, in part, because that variability has not favoured them on this occasion (Terre Blanche & Durrheim, 1999). That threat has been countered in this study, as the sampling is not on the extreme and random variability will be applicable. Participant drop out can affect the internal validity of a study. However, for this study the sample began at 36 children and ended with the same number, with the same children being tested (Terre Blanche & Durrheim, 1999).

Threats to external validity are identified and are discussed as follows. Firstly the question that needs to be answered is whether there are any aspects of the way in which the participants or subjects were selected that could threaten the generalisability of the observed results. For this study there may indeed be a threat to external validity, as the subjects, although all from one grade, they were, however, taught by two different teachers. If teacher effect is dismissed, the validity of the findings will be increased. One way to improve the external validity is to replicate the study with a different sample. That will be discussed later as part of the recommendations for further research.

Generalisability to other operationalisations of the intervention or similar forms of the intervention is to be considered. Should generalisability not be established that may present as a threat to the external validity of this study, which elicits opportunities for further research. In this study the effect of the THRASS Program (independent variable) on the word level accuracy (dependent variable) of a group of Grade 2 learners, using the NARA as a standardised test, was measured. The use of another measurement instrument could elicit further enquiry and the results on that measure could be compared with the results from this study and observations noted. That threat has been countered, to some extent, by using a standardised test with strong construct validity; however, that threat has produced another possibility for further research (Terre Blanche & Durrheim, 1999).

### **3.7.4.2 Reliability**

Reliability is the degree to which the results are repeatable and dependable (Creswell, 2003). The reliability of the data collection method, i.e. the NARA, has been established. Therefore, it is said that the results are dependable. However, regarding the repeatability of the test with the same sample that will not be possible as it is not possible to unexpose the sample to the THRASS Program. However, a similar sample could be used. That is discussed under recommendations for further research in Chapter 5.

### 3.8 CONCLUSION

Chapter 3 placed the research for this study within the positivist theoretical framework and explored quantitative methods used to obtain and analyse the data received from the performance based assessment administered. The data has been described by making use of descriptive statistical methods and the possible inferring of the results has been explored using appropriate parametric tests. The various related variables to the reading process have been considered with reference to their possible effect on the findings of this study, which is further explored in Chapter 4.

While the validity and reliability of the measurement instrument as well as possible validity and reliability issues pertaining to the results of data analysis were dealt with in Chapter 3, those issues are discussed further in Chapter 5 in which the findings of this research are discussed.

---oOo---



## **4.1 INTRODUCTION**

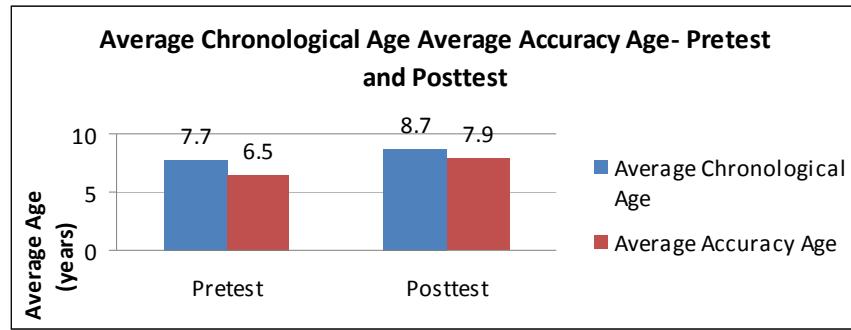
The focus of Chapter 4 is on the presentation and interpretation of the results of the quantitative analysis of reading accuracy scores taken from the NARA. Those results have been scrutinized and statistically analysed to identify patterns and the probability of the influence of the THRASS Program on the word level reading accuracy of the participating group of Grade 2 learners. Firstly, a descriptive analysis of the data will be presented, followed by an exploration of the various statistical methods used to compare the pretest data with the posttest data on a scientifically sound basis.

## **4.2 DESCRIPTIVE ANALYSIS**

The aim of descriptive analyses is, among others, to describe data by investigating the distribution of scores on each variable and determining whether the scores on the different variables are related to each other (Terre Blanche & Durrheim, 1999). The first of the descriptive methods used in the exposition of this study are the various frequency distributions that have been obtained. The frequency distributions are presented in the form of bar charts, with the aim of graphically representing the data and summarise the scores or observations obtained on a single variable. For each of those frequency distributions, measures of central tendency and spread are calculated and presented. The measures of location, which represents the central tendency that were reported, include, the mode, which is the most frequently occurring score in the distribution; the median, which is the middlemost score in the data set which has been ranked from lowest to highest; and the mean, which is the average of all the values in the data set. The measures of spread which were reported are the range which represents the difference between the value of the largest and smallest observation on a variable (Neuman, 1997) and the standard deviation which represents the average distance of the observation's spread around the mean.

### **4.2.1 BAR CHARTS**

Figure 4.1 graphically represents the average chronological age and accuracy age for all participants on the pretest as well as the posttest.



**Figure 4.1: Average chronological age versus average accuracy age-pretest and posttest**

It is observed on the pretest that the average chronological age for this group of learners is 7.7 years and the average accuracy age as measured by the NARA is 6.5 years. The modal chronological age is 7.3 years and the modal accuracy age is 6 years. The median chronological age is 7.7 years and a median accuracy age is 6.4 years. The chronological age range on the pretest is 1.5 years and the accuracy age range is 3.2 years. On the posttest the average chronological age of the participants is 8.7 years and the average accuracy age is 7.9 years. The modal chronological age is 9.0 years and the modal accuracy age is 8.4 years. On the posttest the median chronological age is 8.8 years the median accuracy age is 7.7 years. The range of the chronological age is 1.6 years and the range of the accuracy age is 5.3 years. The central tendencies and measures of spread for the selected data are presented in Table 4.1.

**Table 4.1: Central tendencies and measures of spread for the pretest and posttest ages**

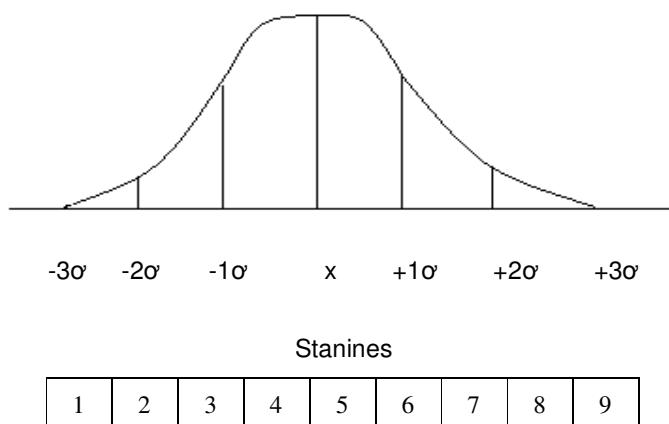
	Pretest		Posttest	
	Chronological age (yrs)	Accuracy age (yrs)	Chronological age (yrs)	Accuracy age (yrs)
<b>Average</b>	7.7	6.5	8.7	7.9
<b>Median</b>	7.7	6.4	8.8	7.7
<b>Mode</b>	7.3	6.0	9.0	8.4
<b>Range</b>	1.5	3.2	1.5	5.3
<b>Standard Deviation</b>	0.39	0.86	0.39	1.35

According to the NARA, the closer a child's result on the test is to its chronological age the higher the likelihood that the child is functioning at its level (Neale, 1997). The observed differences of those results will be subjected to statistical analysis, to ascertain whether there

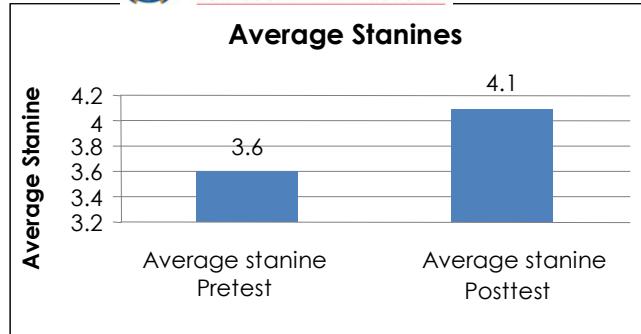
are any statistical significant differences. It is evident from the graphical representation that there was an improvement in the average accuracy age of this group of learners from 6.5 years to 7.9 years. That improvement will be subjected to further statistical analysis to establish significance and interpreted in the following section.

#### 4.2.1.1 Stanines

Stanines have been included as part of the research associated with the NARA. Stanines are bands of values or broad units derived from the conversion of individual raw scores on an assessment to comparative scores obtained from a normative sample's performance (Neale, 1997). Stanines involve the representation of those converted scores to a frequency distribution divided into nine categories or deviations from the average. Each stanine is equal to approximately one-half a standard deviation in width and the average being the midpoint of the fifth or middle stanine (Foxcroft & Roodt, 2001). That enables an illustration of deviation from the average in quantifiable, concrete units. Normally distributed scores are divided into nine units of varying size. The nine units of this scale (see Figure 4.2a) have essentially equal distances along the base line of the curve of the normal distribution. Stanines are useful for reporting differences in broad general terms. The rationale for stanine usage in this study lies in the attempt to gauge the progress of the group or population under study. The stanines are illustrated on a normal distribution in Figure 4.2a. The stanine table attached to Figure 4.2a represents the average accuracy stanines for the population of Grade 2 learners on the pretest and posttest. The results of the research indicate that the average accuracy stanine has altered from 3.5 to 4.1 in the period of a year, as represented by Figure 4.2b. That equates to almost half a standard deviation. Once again, that is not necessarily statistically significant; however it is indicative of change.



**Figure 4.2a: Normal distribution curve and stanine illustration**



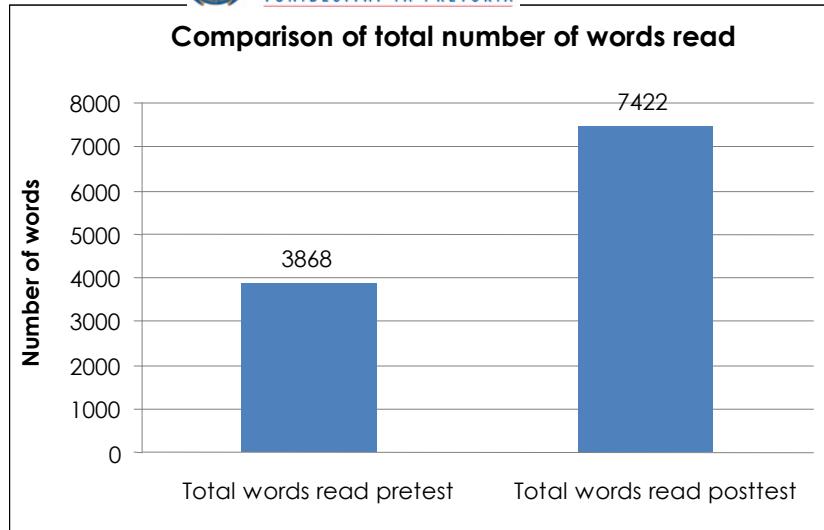
**Figure 4.2b: Bar chart of average stanine for pretest versus posttest**

It is observed in Figure 4.2b that there has been an increase in the average accuracy stanine from 3.6 to 4.1 in a period of one year. Despite the increase in the chronological age of the participants, it is not necessarily the case that each participant will move to the next accuracy stanine, as is indicated by this study. However, as is observed there has been an increase in the average accuracy stanine for the group. That result will be subjected to further statistical analysis to ascertain the significance of this change.

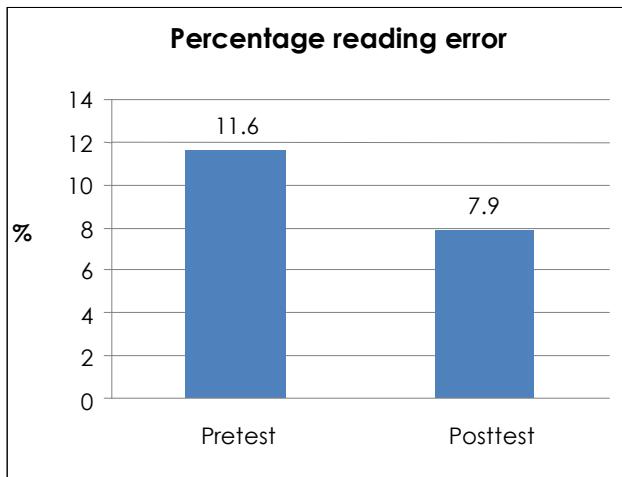
#### **4.2.2 FREQUENCY DISTRIBUTIONS OF ANALYSIS OF READING ERRORS**

The intention of the error analysis conducted within this study was to compare the total number of words read with the total number of errors made on both the pretest and posttest. Thereafter, the percentage errors on the pretest will be compared to the percentage errors made on the posttest and, ultimately, the individual error categories will then be compared. The error categories will be compared collectively and statistically evaluated. Subsequently, each individual error category will be presented.

The total number of words read was compared to the total errors made on the pretest and the posttest:

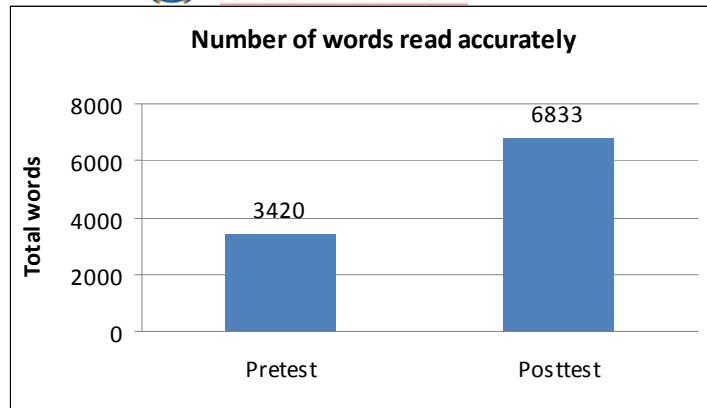


**Figure 4.3: Comparison of total number of words read: pretest versus posttest**

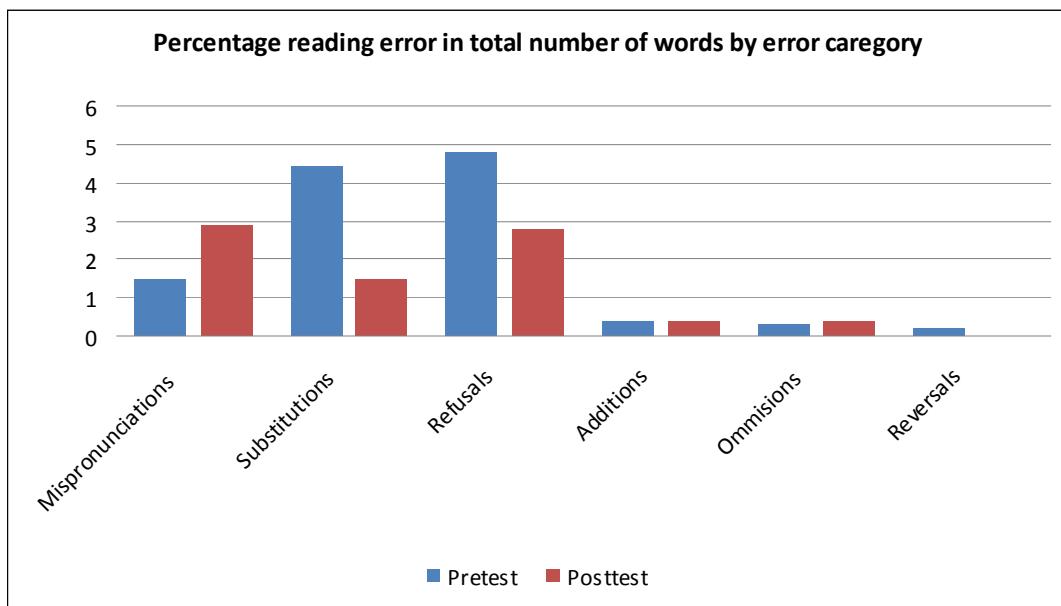


**Figure 4.4: Percentage reading errors for the pretest versus posttest**

It is noted that the distributions indicate an increase in the total number of words read from pretest to posttest. It is also evident that a decrease of 3.7% is detected in the percentage of errors recorded from the pretest to the posttest. Figure 4.5 represents the number of words read accurately, i.e. the total number of words minus the total number of errors.



**Figure 4.5:** Number of words read accurately on pretest and posttest



**Figure 4.6:** Percentage reading error in total number of words read for pretest and posttest by reading error category

For this group of learners it is evident that the number of words read accurately increased by a total of 3413 words from the time of pretest to the recording of posttest results. Most of the errors made on the posttest were mispronunciations, followed by refusals and then by substitutions. For closer investigation, the results from the pretest are presented together with the results from the posttest for purposes of a clear comparison.

As can be observed, the percentage of mispronunciations increased on the posttest and the percentage of substitutions as well as refusals both decreased. It is beyond the scope of this

study to statistically analyse the reading errors, however the differences are noted and are recorded as areas requiring further research.

### 4.3 INFERENTIAL STATISTICAL ANALYSIS

Inferential statistics are used for two main purposes: to test hypotheses and to estimate population parameters (values that describe the properties of the population). Testing of hypotheses involves the validation of the assumed (unknown) values of the parameters which describe the population by making use of the observed data gathered from the population via a sample. Inferential data analysis allows the researcher the opportunity to discover possible relationships between variables and the real world, for the population and not only the sample (Terre Blanche & Durrheim, 1999). Inferential tests are broken up into two broad categories: parametric tests and non-parametric tests. Parametric statistical techniques, however, make use of the assumption that the data follow a certain distribution (mostly the normal distribution) and also allow mathematical operations as the data considered is at the interval or ratio scale. The average, for example, is a measure of central tendency, and is determined by adding scores and dividing by the number of observations ( $n$ ). Non-parametric statistical techniques focus on the order or ranking of data and do not take cognisance of the numerical properties of numbers at the interval and ratio scales; there is also no assumption of underlying distributions to which the data must adhere. Various parametric and non-parametric tests have been performed on the data within the framework of this study in order to attempt to establish associations between variables as well as to establish any significant differences between variables.

#### 4.3.1 TEACHER EFFECT

Since data from two different classes were used it becomes necessary to ascertain whether any significant differences exist between the data of the two classes, as such differences may affect the internal validity of the study. The pretest and posttest accuracy scores for both classes were, therefore, compared and analysed for any significant differences. For that purpose the Kruskal-Wallis test was employed because the number of observations in both classes was less than 30; the observations in the classes were independent and it could not be assumed that the observations were from two normally distributed populations. The pretest hypothesis is formulated as follows:

- Null Hypothesis:  $H_0$ : The average pretest accuracy score of class 1 = the average pretest accuracy score of class 2.
- Alternate Hypothesis:  $H_A$ : The average pretest accuracy score of class 1 is not equal to the average pretest accuracy score of class 2.

**Table 4.2: Kruskal-Wallis test to detect teacher effect in the pretest scores**

Class	Average (Months)	Standard deviation (Months)	Count (n)
Class 1	76.82	9.93	17
Class 2	77.89	11.03	19

The exceedance probability generated by the Kruskal-Wallis test was 0.96, which is greater than 0.05. Therefore the Null Hypothesis ( $H_0$ ) is not rejected and it is concluded that the two classes do not differ with regard to their average accuracy age on the pretest.

The posttest comparison is as follows:

- $H_0$ : The average posttest accuracy score of class 1 = the average posttest accuracy score of class 2.
- $H_A$ : The average posttest accuracy score of class 1 is not equal to the average posttest accuracy score of class 2.

**Table 4.3: Kruskal-Wallis test to detect teacher effect in the posttest scores**

Class	Average (Months)	Standard deviation (Months)	Count (n)
Class 1	94.41	16.36	17
Class 2	95.16	16.50	19

The exceedance probability generated by the Kruskal-Wallis test was 0.862 which is greater than 0.05. Therefore the Null Hypothesis ( $H_0$ ) is not rejected and it is concluded that the two classes do not differ with regard to their average accuracy age on the posttest.

The classes, therefore, do not differ significantly from each other at the beginning or end of the study; thus, there is no indication of a significant teacher effect and this permits the data of the two classes be grouped together.

#### **4.3.2 COMPARISON OF THE PRETEST AND POSTTEST REGARDING AVERAGE ACCURACY**

The average accuracy scores of the pretest and the posttest were compared using the Friedman Test. The Friedman test was employed for this purpose because the observations within the class (pretest and posttest) were dependant and it could not be assumed that the observations were from two normally distributed populations. It can be concluded that there

is a difference in the averages or the accuracy scores between the pretest and the posttest, however this difference is not statistically significant.

The Hypothesis is as follows:

- Ho: The average pretest accuracy score = the average posttest accuracy score.
- HA: The average pretest accuracy score is not equal to the average posttest accuracy score.

**Table 4.4: Friedman test on the accuracy between the pretest and posttest (n=36)**

Accuracy Score (Converted)	Average (Months)	Standard deviation (Months)
Pretest	77.39	10.39
Posttest	94.81	16.20

The Friedman test statistic is 0.94, with an exceedance probability value of 0.330 which is not significant at either the 10% or 5% level of significance.

The Null Hypothesis can therefore not be rejected and, although some difference between the pretest and the posttest is noted, the difference is not statistically significant.

#### **4.3.3 INVESTIGATING PROGRESS OF LEARNERS BETWEEN PRETEST AND POSTTEST RESULTS**

Contingency tables (cross tabulations) were set up for the learners in Class 1 (n=17) as well as those in Class 2 (n=19), to apply the Chi-Squared Procedure. The learners were categorized on the difference between their accuracy age scores and their chronological age for the pretest and the posttest. If the difference was less than zero (accuracy age less than chronological age) the difference category was recorded as -1 (negative difference), otherwise it was recorded as +1 (positive difference). The contingency table consists of the matching frequencies (cell counts) for the different categories in which learners can be classified. The Chi-Squared procedure is executed under the assumption that the rows and columns are independent of each other and it is used to detect significant shifts in frequencies between the different cells of the contingency table. The contingency tables are presented in Tables. 4.5 and 4.6.

**Table 4.5: Contingency table and frequencies of posttest and pretest categories for Class 1**

		Pretest Category		Total
		-1	1	
Posttest	-1	13	0	13
	1	2	2	4
Total		15	2	17

From the contingency table it is clear that the category status of 15 (13 +2) learners stayed the same, indicating no improvement or deterioration in reading ability. No learners changed from a +1 pretest category to a -1 posttest category, which would have signalled deterioration in reading ability. The status of two learners changed from a -1 pretest to a +1 posttest category, which signals an improvement in reading ability.

The Chi-Squared test statistic had a value of 7.37 and an exceedance probability of 0.007, while the exceedance probability of the Fisher's Exact Test was 0.044, which is less than 0.05 (5% level of significance), indicating that a significant shift occurred between the frequencies in the contingency table. It can thus be concluded that the improvement of the two learners represents a significant event.

**Table 4.6: Contingency table and frequencies of posttest and pretest categories for Class 2**

		Pretest Category		Total
		-1	1	
Posttest	-1	15	0	15
	1	1	3	4
Total		16	3	19

From the contingency table it is clear that the category status of 18 (15 +3) learners stayed the same, indicating no improvement or deterioration in reading ability. No learners changed from a +1 pretest category to a -1 posttest category, which would have signalled deterioration in reading ability. The status of one learner changed from a -1 pretest to a +1 posttest category, which signals an improvement in reading ability.

The Chi-Squared test statistic had a value of 13.36 and an exceedance probability of <0.001, while the exceedance probability of the Fisher's Exact Test was 0.004, which is less than

0.05 (5% level of significance), indicating that a significant shift occurred between the frequencies in the contingency table. It can thus be concluded that the improvement of the one learner represents a significant event.

It can thus be concluded that in both classes there were significant improvements in reading ability of learners between the pretest and posttest.

#### **4.3.4 COMPARISON OF THE AVERAGE STANINE SCORES ON THE PRETEST AND THE POSTTEST**

The stanines have been calculated for the group and are indicated by  $\pm$  half a standard deviation from the average of a normal distribution. The pretest accuracy average for the participants indicates a stanine average of 3.5 and on the posttest; a stanine of 4.1 was observed. There was, therefore, an improvement in reading ability of the group as a whole. The comparison of the stanine averages from the pretest and posttest, to determine whether the difference is significant at the 5% level, is set out below:

- $H_0$ : The average pretest stanine is equal to the average posttest stanine.
- $H_A$ : The average pretest stanine is not equal to the average posttest stanine.

**Table 4.7: Statistical analysis of pretest stanines versus posttest stanines (n=36)**

Stanine	Average	Standard deviation
Pretest	3.5	1.18
Posttest	4.1	1.59

The Friedman test statistic value is 2.12 with an exceedance probability of 0.146, which is greater than 0.05 and therefore the Null Hypothesis cannot be rejected; indicating that although a difference has been observed, the difference is not statistically significant.

#### **4.3.5 COMPARISON OF READING ERRORS ON THE PRETEST AND POSTTEST**

As error analysis does not form part of the main focus of this study, the results are included here as they elicit an aspect for further enquiry and will be included as part of the recommendations for further research. Various comparisons have been made between pretest and posttest values and observations are for descriptive purposes only, as they do not inform the results of this study as the change in age has not been taken into account. It is recognised that there are many factors that could have contributed to the recorded differences. Those factors do not form part of the scope of this study. It is, however, worthy

to note those differences. Comparisons have been made regarding the number of words read on the pretest and posttest as well as the types of errors made.

#### 4.3.5.1 Number of words read

An analysis was done on the number of words read by learners under the same conditions during both the pretest and posttest, to test for a significant difference. The null and alternative hypothesis was set up as follows:

- Ho: The number of words read on the pretest is equal the number of words read on the posttest.
- Ha: The number of words read on the pretest is not equal to the number of words read on the posttest.

**Table 4.8: Comparison between the number of words read during the pretest and the posttest (n=36)**

Number of words read	Average	Standard deviation
Pretest	107.4	55.51
Posttest	206.4	104.50

The Friedman test was applied, which yielded a test statistic value of 21.78 with an exceedance probability of <0.001, which is significant at the 5% level of significance. Therefore, the Null Hypothesis is rejected in favour of the Alternative Hypothesis. The average number of words read during the pretest is significantly different from the average number of words read during the posttest.

#### 4.3.5.2 Types of errors made on the pretest and the posttest

The error analysis was alluded to in the literature of this study, especially in that various reading errors are attributed to various aspects of the reading process. According to the NARA, errors are categorised according to grapho-phonetic errors, syntactical errors and semantic errors e.g. a mispronunciation is classed as a decoding error or grapho-phonetic error and an addition as a syntactical error. Error analysis is not the primary focus of this study; however the observation of the results inspires further inquiry and opportunities for further research. A test procedure making use of the differences (pretest % subtracted from the posttest %) over all the learners in each category of reading error can be employed to gain some insight in that matter. The T-test was applied to the differences (posttest % – pretest %) found on the pretest and posttest percentage of reading error, per category of reading error and the statistical significance was tested (it must be noted, however, that the

age difference and time frame of one year has not been taken into account). The results are as follows:

- Ho: Percentage error (category e.g. mispronunciations) on the pretest is equal to the percentage error (category e.g. mispronunciations) on the posttest.
- Ha: Percentage error (category e.g. mispronunciations) on the pretest is not equal to the percentage error (category e.g. mispronunciations) on the posttest.

**Table 4.9: Comparison between the different types of reading errors made during the pretest versus the posttest (n=36)**

Type of reading error	Average difference (%) (Posttest% – Pretest%)	Standard deviation of difference (%)	t-Value	Exceedance Probability (p-value)
Mispronunciations	1.69	2.09	4.86	0.001*
Substitutions	-3.64	5.27	-4.14	0.002*
Refusals	-1.985	4.85	-2.44	0.020*
Additions	0.18	0.91	1.19	0.243
Omissions	0.13	0.64	1.24	0.225
Reversals	-0.20	0.55	-2.15	0.039*

\* Significant differences on the 5% level of significance

The Null hypothesis is rejected in favour of the alternative hypothesis for the categories of mispronunciations, substitutions, refusals and reversals, because their p-values are less than 0.05. The average of the above mentioned categories differ at the 5% level of significance. The Null Hypothesis cannot be rejected for the additions and omissions error categories and, therefore, it may be concluded that those means do not differ significantly from each other.

#### 4.4 SUMMARY OF QUANTATIVE ANALYSIS

Descriptive analysis enabled the development of frequency distributions and central tendencies in order to obtain, present and evaluate comparative data with regard to the results available for this study. Inferential statistics were employed to reinforce the conclusion that a teacher effect between the two classes was not statistically significant, enabling the data from those classes to be combined in the evaluation of further results. From the analysis of this data it became evident that an improvement in the average accuracy age and stanine scores occurred between the pretest and posttest. Further statistical analysis clarified that the difference was not statistically significant. Two specific observations in Class 1 and one in Class 2 indicated significant improvements between the pretest and posttest ages

(accuracy – chronological). The statistical analysis yielded results which indicated a significant increase in the total number of words read from pretest to posttest.

Reading error analysis was included in the reporting of this study, for purposes of interest and in an effort to assist in generating further research ideas. Reading error analysis manifests a decrease in the percentage of reading errors recorded from the pretest to the posttest. The number of words read accurately increased from the pretest to the posttest results. As is evident from the posttest, the majority of the reading errors recorded were detected in the areas of mispronunciation, followed by refusals and then by substitutions. It is recorded that the percentage of reading errors recorded on the pretest were less than those on the posttest in the categories of substitutions, refusals and reversals, with more reading errors being recorded on the posttest in the categories of mispronunciations, additions and omissions. On further statistical analysis, differences on the categories of mispronunciations, substitutions, refusals and reversals indicated statistical significance.

#### **4.5 CHAPTER SUMMARY**

In Chapter 4 the results of the data received from the NARA for the population (two classes from a specific school) of Grade 2 learners were investigated. The data was presented in the form of bar graphs and descriptive statistics were calculated, to describe and investigate the distribution of the data, as well as investigate if and how the data are related to each other. Thereafter various inferential statistical tests were applied to that data to test various hypotheses and calculate the exceedance probability values (p-values), to determine significance. What emerged from that data analysis is that, due to time that lapsed between the two tests, a difference has been observed between the pretest and posttest scores over and above the expected difference of one year.

However, the overall difference yielded an exceedance probability (p-value) above the 5% level and, therefore; it is not regarded as a statistically significant change. It can thus not be concluded that the difference between pretest values and posttest values was only caused by the influence of the THRASS Program. Further discussion of those results and possible causal influences in score difference will be addressed, as will as opportunities for expanding the scope of this research in Chapter 5, which concludes this study.

---oOo---



#### 5.1 INTRODUCTION

Learning to read is a critical skill that plays a significant role in education. As was mentioned in the rationale of this study, the management of the school where I am currently employed has perceived a decline in the standard of reading among the children in the last few years. To address that situation, it was decided to implement the THRASS Program at the school. The purpose of this research was born from a need expressed by the management to determine if the implementation of the THRASS Program had any influence on the development of learner reading ability in that specific school.

The aim of the THRASS Program is to develop children's phonemic awareness and assist them in becoming phoneme rich, thereby improving their decoding and word recognition skills. I have made reference to my interest in this study of children's learning in general, which interest underpinned my decision to further my studies in educational psychology. While furthering my studies I developed a particular interest in children learning to read and have witnessed some of the difficulties some children experience with the acquisition of certain learning skills. The effectiveness of the THRASS Program is, therefore, of particular interest to me and has informed the rationale of this study.

This study has included the administration of a pretest and a posttest, utilising a norm-referenced test administered prior to the implementation of the THRASS Program and re-administered one year following its implementation. The norm-referenced test utilized was the NARA. The accuracy subtest of the measure was used. Reading accuracy comprises the ability to decode words, recognize words and articulate words with phonemic accuracy. Those are skills that the THRASS Program aims to develop. Thus reading accuracy was the focal point of this research.

The sample selected included children about to commence with Grade 2. The posttest was conducted at the time the sample commenced their Grade 3 schooling year. The sample/target population comprised all the Grade 2 learners at the school who were placed in two different classes and taught by two different teachers. The reading accuracy results from the pretest were then compared statistically with the results of the posttest to observe any significant difference. The outcomes of the statistical analysis will now be briefly discussed.

## 5.2 DISCUSSION OF RESEARCH RESULTS

From the results recorded in Chapter 4, it is accurate to say a change has been recorded in the word level accuracy norm score for the sample over the period in which the participants were exposed to the THRASS Program. As stated previously, the primary interest in this study is the average accuracy scores for the sample and for that reason the primary focus of this discussion is on the accuracy averages of the group taken from the pretest and posttest results. The averages (in years) for the pretest and posttest are 6.5 ( $SD = 0.86$ ) and 7.9 ( $SD = 1.35$ ), with an exceedance probability value of 0.330 and a difference of 4 months. As alluded to earlier, it is clear there has been a positive change in the averages of the norm scores produced on the posttest. The NARA is a norm based test and, therefore, it is expected that a proportional increase in the results from the pretest to the posttest would occur. However, it is evident that there has been a 4 month increase in the average despite the consideration of age-related norms. That increase, however, has been subjected to various statistical tests and was found to be not statistically significant. Various causal factors may be responsible for that increase and warrant further discussion. Some of those will be addressed within the limitations of the study and as part of the recommendations for further research. However, the introduction of the THRASS Program cannot be disregarded when considering the role played in differences recorded. As a whole the group has improved on their accuracy of reading, which is further indicated by the movement in the stanine from 3.6 to 4.1 – however, that result has proved not to be statistically significant.

From the contingency tables presented in section 4.3.3 for class 1 of the previous chapter the status of two learners changed from a -1 pretest to a +1 posttest category, which signals an improvement in reading ability. The Chi-Squared test statistic had a value of 7.37 and an exceedance probability of 0.007, while the exceedance probability of the Fisher's Exact Test was 0.044, which is less than 0.05 (5% level of significance), indicating that a significant shift occurred between the frequencies in the contingency table. It can thus be concluded that the improvement of the two learners represents a significant event.

It can be stated with confidence that a change in the average accuracy scores has been observed, however, it cannot be stated with confidence that the THRASS Program was solely influential in that change.

Although this study is not primarily concerned with error analysis, a change in error patterns has been observed which elicited some interest. An increase in the number of words read as well as the number of words read accurately has been observed. The NARA does not provide stanines for the change in the number of words read or the change in error category.

The results of those descriptions cannot be statistically analyzed, as the change in age, i.e. 12 months, has not been taken into account. However, if one considers the percentage of error, calculated by taking the total errors per category, dividing it by the total number of words read and presenting this answer as a percentage, some statistically significant results have been observed. A significant difference has been observed in mispronunciations (4.86), substitutions (-4.14), refusals (-2.44) and reversals (-2.15). The positive value indicates that there were more recordings of that particular error on the posttest and the negative value indicates that there were more recordings of that particular error on the pretest. Thus on the pretest the most errors that occurred were substitutions and on the posttest that shifted to mispronunciations. Both those types of errors are referred to as decoding errors, however on reflection it would seem to indicate that on the pretest more words have been guessed, whereas mispronunciations would imply that the words have been recognised and attempted. This is, however, speculative and calls for further research.

The change in error pattern is worthy of further inspection, particularly when considering that the THRASS Program is designed to improve phonetic awareness and word level reading. That shift may be a normal progression due to maturation, increased exposure to reading and other factors. Further research in this area is required for conclusions to be drawn.

### **5.3 LIMITATIONS OF THE STUDY**

Terre Blanche and Durrheim (1999) assert that we can only partially understand accounts of the world by referring to the facts and must look beyond the empirical evidence into the background knowledge that makes that evidence believable. It is; therefore, appropriate to consider the background knowledge and influences that may have had an impact on this study and the result thereof. Those influences are discussed below and are considered important as possible influences of the validity of the findings. It is also important to contextualise the understanding and interpretation of the findings.

#### **5.3.1 ABSENCE OF CONTROL GROUP**

Within this study no control group was utilized as a method of comparison, which may be perceived as a limitation to this study. The introduction of a control group may have expanded the research focus, as two different methods of teaching would need to be explored, which may have created a change in research focus. However, a control group may have enabled the research to identify the THRASS Program as the only causal influence in reading accuracy, should a change have been observed.

### **5.3.2 TIME ALLOWED FOR THRASS IMPLEMENTATION**

Another potential limitation of this study includes the degree of time allowed for the implementation of the THRASS Program prior to measurement of results. The THRASS Program is designed as a four-year curriculum and, therefore, in order to adequately evaluate the program effectively, it can be asserted that total exposure to the curriculum is necessary for holistic evaluation. It has, therefore, been suggested under the section on the opportunities for further research that this be considered. Longitudinal studies would provide the opportunity for that holistic evaluation, despite the limitations created by potential loss of subjects over time and increases in extraneous influential variables over time.

### **5.3.3 MATURATION**

It is important to take cognisance of the fact that development and maturation in children involves the integration of various skills and processes (Donald, *et al.*, 2002). Physical development is, therefore, likely to play a role in the ability of a subject to process information adequately. It has not been possible within the scope of this study to include a concise investigation of appropriate overall development of the subjects, as that lies beyond the scope of this study but must be considered as playing a possible role in general performance.

### **5.3.4 INDIVIDUAL DIFFERENCES**

It is acknowledged that despite the utilisation of age-related normative information on the NARA that is likely to represent the general development in large numbers of children. Due to individual development being unique and progression of skills occurring at individually determined rates, it is important to consider that even those age-related norms may not be representative of each student's degree of development at any given time. Age-related norms do, however, provide the best available measure of overall expected rates of progression and have, therefore, been utilised within this study.

### **5.3.5 EXPOSURE TO OTHER LEARNING OPPORTUNITIES**

It is possible that participants involved in this study may have been exposed to learning opportunities beyond the confines of the schooling environment, which are not possible for this research to assess. The degree to which parents are involved in the development of reading ability or the exposure to other reading programs and interventions, are beyond the scope of this study but could play a significant role in the development of reading skill. It is,

therefore, acknowledged that learning takes place in more than one context that could have influenced the development recorded within the study.

### **5.3.6 SUBJECT REACTION TOWARD ASSESSMENT**

It is considered that assessment, in any form, elicits potential emotional reactions in subjects. Due to the subjects of this study being familiar with me, it is considered that the emotional effects of the assessment may have been reduced to some degree. However, performance anxiety cannot be disregarded as influential in the results of this study but they have not been measured, as the scope of the study does not include that form of measurement.

### **5.3.7 NARA IN SOUTH AFRICAN CONTEXT**

The NARA is a British test that was standardized according to British norms. This may be a limitation for this research, as the test material and passages may contain aspects that are unfamiliar to children in the South African context. However, the participants for this study are English speaking and have English as their first language. It has been observed that the group of participants has scored significantly lower on the accuracy subscale of the NARA, when compared to their chronological age. It is not part of this study to investigate that phenomenon, however, I have observed through applying the NARA that the children at the school I am at appear to catch up with the norms and the gap between the chronological age is bridged. It is thus my opinion that the actual accuracy age score achieved is not as significant as the observed increase in the average accuracy age scores. That is, however, information gained from my personal experience and is acknowledged as being an unscientific opinion which could be transferred to recommendations for further research.

### **5.3.8 THRASS IN SOUTH AFRICAN CONTEXT**

The THRASS Program was developed in the United Kingdom, without research being done as to the applicability to the South African learner population. That may be a limitation to this study as well as to the learners, as there may be some unfamiliarity with some of the words and sounds which form part of that. However, the learners who participated in this study are all first language English speakers and their language of education is English. One of the reasons that this study was undertaken has been to attempt to begin the research on the applicability of the THRASS Program for South African learners and it is hoped that this study will inspire further research.

## **5.4 CONTRIBUTIONS OF THE STUDY**

Enlisted below are possible contributions of this study:

- This study has contributed to a process of research on the THRASS Program in South Africa. It is hoped that this process will be continued.
- The various questions that have emerged from this study can be investigated as further research ideas. Those have been mentioned in the following section.
- This study has provided the school with scientific feedback and insight into the children's reading development.
- THRASS advocates a 'new' approach to phonics teaching i.e. the phonographic approach. Further research could investigate merit for this approach and inspire research into this method of teaching phonics.

## 5.5 RECOMMENDATIONS FOR FURTHER RESEARCH

It is known that no one research study can be all encompassing of any given phenomena. It is, therefore, imperative that research inspires further investigation and enquiry. Elements of this research study have led to the creation of further possible research enquiry based on reading assessment and reading programs.

### 5.5.1 LONGITUDINAL STUDY

It has been mentioned that the complete THRASS Program comprises a four-year curriculum. It is suggested that, to enable a comprehensive and holistic evaluation of the overall impact of the THRASS Program, it would be advisable to conduct longitudinal research to ascertain that. Consistent assessment of a sample from the introduction phase (i.e. the reception year of schooling) until the completion of the program four years later would enable such a longitudinal study.

### 5.5.2 RESEARCH USING CONTROL GROUPS

To improve the external validity of this study, and perhaps its generalisability, it is felt that using a control group may be useful in future research. That will result in a true experimental study, not only a descriptive study as is this study which will further intensify the research on the THRASS Program.

### 5.5.3 ERROR ANALYSIS

From the data gathered within this study differences in error analysis from pretest to posttest have been observed. Specific error analysis does not form part of the scope of this study and, consequently, such analysis is not discussed in detail. An analysis of changes in error

patterns may provide valuable information in forming assumptions with regard to developments and regressions within certain categories of errors. That analysis may provide valuable information in assessing the development of reading progression and confidence in subjects. As the NARA does not provide norm-based information with regard to errors obtained in the assessment, the exploration of other available norm-based assessment may prove helpful in that respect.

#### **5.5.4 ASSESSING THE DEVELOPMENT IN READING ACCURACY IN DIFFERENT CONTEXTS**

This research, which was conducted at a private school on the East Rand, provides a valuable description of reading accuracy within that context. The opportunity exists for a replication of this study in diverse contexts. Public schooling systems provide opportunities for studies regarding reading accuracy following exposure to the THRASS Program as well as other reading programs. The influence of culture in the development of reading accuracy and responses to varying programs may provide a valuable research opportunity.

#### **5.5.5 INTENSITY OF EXPOSURE TO THE THRASS PROGRAM**

This study was conducted in the context of the THRASS Program being implemented within school hours and within the limit of certain teaching times. The intensity of the exposure of the subjects to the THRASS Program may play a significant role in the degree of development attained in assessment. Evaluating optimal exposure times in relation to the development of reading accuracy may indeed create an opportunity for further enquiry.

#### **5.5.6 ASSESSING DIFFERENT ASPECTS OF READING DEVELOPMENT FOLLOWING EXPOSURE TO THRASS**

This study utilized the NARA in the assessment of reading accuracy following exposure to the THRASS Program. It is possible to extend research opportunities to the utilization of different assessment measures in assessing different aspects of reading, such as reading fluency, reading rate or comprehension.

### **5.6 DEVELOPMENTS IMPACTING ON FUTURE THRASS RELATED STUDIES**

As with any research, to be aware of all variables affecting the outcome of a study to an effective degree, it is essential to establish awareness of development in the field from one research study to the next. It is, therefore, important to include in future studies on the topic the latest developments which may impact that research but were not available at the time of conducting this study.

Two areas of development which came to my attention after the study was complete relate to the development of teaching mediums and teacher confidence in applying the THRASS Program.

#### **5.6.1 DEVELOPMENT OF TEACHING MEDIUMS**

Music has been used by Jeanine Plunkett (personal communication with Mrs. J. Plunkett, 5 August 2008), in the form of the “Sing-a-long CD and DVD”, to provide added teaching material utilizing different media to assist in the teaching the THRASS curriculum. Exposure to differing mediums throughout the teaching process may significantly impact on the degree of learning and the assimilation of knowledge (Vaughn, Bos & Schumm, 1997). That aspect is to be considered in future research conducted in this field.

Workbooks are in the process of being compiled (personal communication with Mr. A. Davies, 7 September 2008) in order to assist in the teaching of the THRASS Program. Exposure to that added medium may create an opportunity for new learning experience among subjects. Any further research in the field of applying the THRASS Program should take that teaching aid into account.

#### **5.6.2 CONFIDENCE IN THE THRASS PROGRAM**

It is likely that the degree of teacher confidence in presenting the THRASS Program may result in changes in the reception and assimilation of this information by subjects. Although no teacher effects were established in this study, it is acknowledged that the study was compiled in the first year of presentation of that program at the school.

There is a responsibility on the part of the researcher to make the results of this study available to the school and the people responsible for the continued implementation of the THRASS Program. The results of this study are important, firstly, to establish any effect after one year of implementation. Further research will, hopefully, be inspired as to the effectiveness of the program after a longer period of application, as well as into other variables that may affect its success, for example, teaching methods, time of exposure, etc.

### **5.7 CONCLUSION**

The hypothesis stated for this study was:

- Ho: Exposing a group of Grade 2 learners to the THRASS Program for a period of one year will have no statistically significant influence on their word level accuracy skills.
- Ha: Exposing a group of Grade 2 learners to the THRASS Program for a period of one year will have a statistically significant influence on their word level accuracy skills.

This research study investigates the influence of the THRASS Program on word level accuracy of a group of Grade 2 learners. There was a difference of approximately 4 months observed between the average accuracy scores on the pretest and the average accuracy scores on the posttest taking age into account; however, the difference shown between the pretest and posttest accuracy average scores on the NARA was not statistically significant. The Null Hypothesis has therefore been tested, and cannot be rejected in favour of the alternate hypothesis. The result has stimulated enquiry and creates the opportunity for further research in the field of reading accuracy and the general development of reading.

It is hoped that this study will inspire a host of other research studies which will collectively contribute to the use and development of effective reading programs across Southern Africa.

---oOo---

## References

---

- Adams, M.A. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.
- Balnaves, M. (2001). *Introduction to quantitative research methods: an investigative approach*. London: SAGE.
- Barone, D.M. & Morrow, L.M. (Eds.). (2003). *Literacy and young children: Research based practices*. New York: The Guilford Press.
- Bouwer, A.C & Jordaan, V (2002). Reading and Writing. In: Eloff, I. & Ebersohn, L. (2004) *Keys to Educational Psychology* (pp.84-118).Cape Town: UCT Press.
- Bouwer, AC. (2000), Reading and Writing. In: Eloff, I. & Ebersohn, L. (2004) *Keys to Educational Psychology* (pp.84-118). Cape Town: UCT Press.
- Burns, P.C., Roe, B.D. & Ross, E.P. (1992). *Teaching reading in today's elementary schools* (5<sup>th</sup> ed.). Boston MA: Houghton Mifflin.
- Byrne, B. & Fielding-Barnsley, R. (1989). Phonemic awareness and letter knowledge in the child's acquisition of the alphabetic principle. *Journal of Educational Psychology*, 81:313-321.
- Chall, J.S. (1996). *Learning to read: The great debate* (3<sup>rd</sup> ed.). Orlando, FL: Harcourt Brace.
- Clay, M.M. (1984). *Observing the young reader*. Auckland, New Zealand: Heinemann.
- Clay, M.M. (1993). *Reading Recovery: A guidebook for teachers in training*. Portsmouth, NH: Heinemann.
- Cockcroft, K. & Balckburn, M. (2008). The relationship between the Senior South African Individual Scale – Revised (SSAIS-R) subtests and reading ability. Retrieved March 11, 2009 from World Wide Web: [http://www.sabinet.co.za/abstracts/sapsyc/sapsyc\\_v38\\_n2\\_a9.xml](http://www.sabinet.co.za/abstracts/sapsyc/sapsyc_v38_n2_a9.xml).
- Creswell, J.W. (2003). *Research design: qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, California: Sage Publications.
- Davies, A. (2006). *Teaching THRASS Whole-Picture Keyword Phonics: The essential guide to progression and assessment for all teachers of English*. Chester: THRASS (UK) Ltd.
- Davies, A. (2008). *THRASS: Theoretical underpinnings*. Retrieved April 24, 2007 from <http://www.THRASS.co.uk>.

De Vos, A.S. (2000). *Research at grass roots: a primer for caring professionals*. Pretoria: Van Schaik Publishers.

Dechant, E (1993). Reading Instruction. Chapter 2. In: Schumm, J.S. (Eds.). (2006). *Reading Assessment and Instruction for all Learners* (pp.27-42). New York: The Guilford Press.

Department of Education. (2005). *Revised National Curriculum Statement: Grade R-3 (Schools) Foundation Phase*. Johannesburg: Department of Education.

Donald, D., Lazarus, S. & Lolwana, P. (2002). *Educational Psychology in Social Context* (2<sup>nd</sup> ed.). Cape Town. Oxford University Press Southern Africa.

Ehri, L.C. & McCormick, S. (1998). Phases of word learning: Implications for instruction with delayed and disabled readers. *Reading and Writing Quarterly: Overcoming Learning Difficulties*, 14(2):135-164.

Fox, B. (2000). *Word identification Strategies: Phonics from a new perspective*. New Jersey: Prentice-Hall Inc.

Foxcroft, C. & Roodt, G. (Eds.). (2001). *An Introduction to Psychological Assessment: In the South African Context*. Cape Town: Oxford University Press.

Gillon, G. (2004). *Phonological awareness: from research to practice*. New York: Guilford Press.

Gorard, S. (2003). *Quantitative Method in Social Science*. London: Continuum Books

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

Huberman, A.M. & Miles, M.B. (1997). Quantitative Data Collection Methods. Section C, Chapter 12. In: De Vos, A.S. (2000). *Research at grass roots: a primer for caring professionals* (pp.159-190). Pretoria: Van Schaik Publishers.

Hudson, R., Lane, H. & Pullen, P. (2005). Reading fluency assessment and instruction: What, why, and how? *The Reading Teacher*, 58(8):23-37.

International Reading Association (1996) *Code of Ethics*, 2008. Retrieved June 5, 2008 from World Wide Web:<http://www.reading.org/association/about/code.html>

Leong, F.T.L. & Austin, J.T. (2006). *The Psychology Research Handbook*. Thousands Oaks: Sage Publications.

Mabotja, S. (2008). August. *THRASS ABSA Talk together project*. Retrieved Aug 11, 2008 from <http://www.thrass.co.uk/absa.htm>.

Manyak, P.C. (2007). A Framework for Robust Literacy Instruction for English Learners. *The Reading Teacher*, 61(2):197-199.

McGuinness, D. (2004). *Early reading instruction*. Cambridge, MA: The MIT Press.

McMillan, J.H. & Schumacher, S. (2001) *Research in Education: A conceptual introduction* (5<sup>th</sup> ed.). New York: Addison Wesley Longman Inc.

Mertens, D.M. (2004). *Research and evaluation methods in special education*. Thousand Oaks, California: Corwin Press.

Neale, M.D. (1997). *Neale Analysis or Reading Ability; Revised British Edition*. London: NFER. Nelson Publishers.

Neuman, L.W. (1997). *Social Research methods. Qualitative and Quantitative approaches*. (3<sup>rd</sup> ed.). New York: Allyn and Bacon.

Perfetti, C. (1985). *Reading Ability*. New York: Oxford University Press.

Phillips M., Clancy-Menchetti J., & Lonigan C. J. (2008). Successful Phonological Awareness Instruction with Preschool Children. *Topics in Early Childhood Special Education Journal*, 28 (1):3-17.

Pike, K., Compain, R. & Mumper, J. (1997). *New connections an integrated approach to literacy* (2<sup>nd</sup> ed.). New York: Longman.

Robeck, M.C. & Wallace, R.R. (1990). *The Psychology of Reading: An Interdisciplinary Approach* (2<sup>nd</sup> ed.). New Jersey: Lawrence Erlbaum Associates, Inc., Publishers.

Rose, J. (2006). *Independent Review of the teaching of early reading*. Nottingham: DfES Publications.

Schumm, J.S. (Ed.). (2006). *Reading Assessment and Instruction for all Learners*. New York: The Guilford Press.

Scott, D. & Usher, R. (Ed). (2001). *Understanding Educational research*. London: Routledge.

Share, D.L. & Stanovich, K. E. (1995). Cognitive processes in early reading development: Accommodating individual differences into a model of acquisition. *Issues in Education*, 1:1-57.

Snowling, M.J. (1998). Reading development and its difficulties. *Educational and Child Psychology*, 15(2):44. The British Psychological Society.

Stanovich, K., Nathan, R. & Zolman, J. (1988). The Developmental Lag Hypothesis in Reading: Longitudinal and matched-reading level comparisons. *Child Development*, 59:71-86.

Steyn, A.G.W., Smit, C.F., Du Toit, S.H.C. & Strasheim, C. (1994). *Modern Statistics in Practice*. Pretoria: J.L.Van Schaik Publishers.

Sulzby, E. & Teal, W.H. (1991). Early Reading. In: Schumm, J.S. (Ed.). (2006). *Reading Assessment and Instruction for all Learners* (pp.7-26). New York: The Guilford Press.

Tashakkori, A. (2008). *Foundations of mixed methods research: integrating quantitative and qualitative techniques in the social and behavioral sciences*. London : SAGE.

Terre Blanche, M. & Durrheim, K. (Eds.). (1999). *Research in Practice: Applied methods for Social Sciences*. Cape Town. University of Cape Town Press (Pty) Ltd.

Thompson, C.B. & Panacek, E. (2006). *Research study designs: Experimental and Quasi-experimental*. Retrieved Dec 19, 2008 from <http://sciencedirect.com/science?ArticleURL&udi=B75B6-4M6ROCK-9>.

THRASS. 2008, August. *THRASS Keyfacts*. Retrieved Aug 11, 2008 from <http://www.thrass.co.uk/absa.htm>).

Tredoux, C. Sound conclusions; judging research designs. In: Terre Blanche, M. & Durrheim, K. (Eds.). (2002). *Research in Practice: Applied methods for Social Sciences* (pp. 309-330). Cape Town. University of Cape Town Press (Pty) Ltd.

Trochim, W. (2006). *Research methods knowledge base: Positivism and Post-Positivism*. Retrieved Feb. 1, 2009 from <http://www.socialresearchmethods.net/kb/positvsmm.php>.

Truxler, J. & O'Keefe, B. (2007). The effects of phonological awareness instruction on beginning word recognition and spelling. *Augmentative and Alternative Communication Journal*, 23 (2):164-176.

Vaughn, S., Bos, C. & Schumm, J. 1997. Teaching exceptional, diverse and at-risk students in the general classroom. Boston: Allyn and Bacon.

Wyse, D. & Styles, M. (2007). Synthetic phonics and the teaching of reading: the debate surrounding England's 'Rose Report'. *UKLA, Literacy Journal*, 41(1):37-38.

Yopp, H. (1988) The Validity and Reliability of Phonemic Awareness Tests. *Reading Research Quarterly*, 23:159-177.

--ooOoo--