# THE INFLUENCE OF READABILITY OF EXAMINATION QUESTIONS ON ACHIEVEMENT IN SENIOR SECONDARY SCHOOL MATHEMATICS

**(A STUDY ON VERBAL PROBLEMS WITH SPECIAL REFERENCE TO SECOND LANGUAGE READERS)**

**ELIZABETH DIANA PRINS**

**Dissertation presented for the Degree of Doctor of Education,**

**at the University of Stellenbosch**

**Promoter: Prof PG Human**

**Co-Promoter: Prof JM Ulijn, Technische Universiteit, Eindhoven, The Netherlands**

**March 1995**

I the undersigned hereby declare that the work contained in this dissertation is my own original work and has not previously in its entirety or in part been submitted at any university for a degree.

Signature: Date:

Financial assistance of the Centre for Science Development (HSRC, South Africa) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at are those of the author and are not necessarily to be attributed to the Centre for Science Development.

# ACKNOWLEDGEMENTS

I would like to thank:

- My promoter Prof PG Human and my co-promoter Prof JM Ulijn without whose painstaking guidance and sustained assistance this study would not have been completed.

- The Tecnichal University of Eindhoven in the Netherlands for receiving me as a guest student for three months during this study. A special word of thanks to the following *assistente in opleiding* (AIO's): Margie Li, Yvonne de Kort, Elwin Savelsbergh and Marc van Kempen for their assistance, encouragement and friendship.

- The colleagues at the Freudenthal Institute in the Netherlands for their interest and help.

- Angelika Alberts for her willing and able assistance with the editing.

- Riana Els for doing the efficient final formatting of the manuscript.

- Luc Kuijk and his colleagues for their valuable advice.

- My husband, Rian, for his wonderful support.

- My Heavenly Father for His infinite love. Soli Deo Gloria.

*To my husband, children & parents*

# ABSTRACT

This study investigates the influence of readability of mathematics examination questions on achievement. The aim of any mathematics examination is to assess whether the aims of a specific mathematics programme have been realized. Readability factors that unnecessarily prevent a clear understanding of questions could jeopardize this aim. The important issue is, therefore, whether there are indeed readability factors in mathematics examination questions that cause comprehension problems for students and, if there are, do they have any effect on test scores?

The issue of readability is of even greater importance for second language readers. In the South African context, the reading problems of second language readers are of particular importance as most students at school are second language learners. An important question would therefore be: What readability factors cause comprehension difficulties for second language students, especially those whose mother tongue is not kindred to English? Furthermore, what is the influence of cultural factors on readability? This study provides answers to these and other related questions for mathematics text at senior secondary school level.

Protocol analysis was used to ascertain what readability problems are experienced by students when reading examination questions in mathematics. Three different language groups, comprising 17-18-year-old students, were used in the study: English First Language students and two groups who had English as a second language. One second language group had Afrikaans as first language while the other group comprised African students whose mother tongue is unrelated to English. A framework was developed to analyse the protocols and it comprised five categories:

- unfamiliar vocabulary
- structural problems
- obscure information
- visualization difficulties
- non-verbal factors

After the protocol study, students were asked to adapt the examination questions to a more comprehensible form. Students' adaptations addressed lexical, syntactical, discourse

and non-verbal factors. Most of the readability problems identified in the literature study were verified in the empirical study. However, the empirical study generated additional readability problems that are mainly restricted to mathematics text and relate to non-verbal factors like mathematical expressions. During the last phase of the empirical study a composite test was used to test the hypothesis that *improved readability of the common language used in mathematics examination questions will improve achievement.* Nine so-called "word problems" from previous examination papers were set in three different versions: original, adapted and non-verbal.

The hypothesis was confirmed in a number of important cases. A significant finding of the study was, therefore, that readability factors not only influence the comprehension of mathematics examination questions, but also have a marked influence on students' achievement levels. The results of the empirical study are reported quantitatively as well as qualitatively. Other results include the following:

- Not only second language students, but also first language students experienced a variety of readability problems.
- All three language groups demonstrated the same level of competency on the non-verbal versions. When comparing test scores of the verbal versions, differences in achievement levels between the different language groups were often caused by linguistic and cultural factors.
- Cultural thought patterns, typical of a mother tongue but absent in a second language, were often a source of comprehension difficulties for second language readers.

This study has led to certain conclusions for teaching and examination practice. For example, factors influencing the readability of ordinary English should be considered with other factors when writing mathematics examination questions. Furthermore, the distinctly different reading needs of second language students suggest that examination papers be set, so that the language needs of second language learners are accommodated.

Guidelines for writing more readable examination questions were developed and are presented as a readability checklist. Suggestions for further research include the investigation of the influence of readability on achievement in authentic examination conditions.

# OPSOMMING

Hierdie studie ondersoek die invloed van leesbaarheid van wiskunde eksamenvrae op prestasie. Die doel van enige wiskunde eksamen is om vas te stel of die doelwitte van 'n spesifieke wiskunde program bereik is. Leesbaarheidsfaktore wat die volledige begryp van vraestelle onnodig belemmer, kan die bereiking van hierdie doel verhinder. Dit is dus belangrik om vas te stel of daar wel leesbaarheidsfaktore in wiskunde eksamenvrae bestaan wat vir leerlinge begripsprobleme veroorsaak en, indien wel, of hulle enige effek op toetspunte het.

Vir tweedetaal lesers is die kwessie van leesbaarheid van nog groter belang. In die Suid-Afrikaanse opset is die leesprobleme van tweedetaal lesers 'n uiters aktuele saak aangesien die meeste skoolleerlinge tweedetaal leerders is. Belangrike vrae is dus: Watter faktore veroorsaak leesbaarheidsprobleme vir tweedetaal leerlinge, veral diegene met 'n nie-verwante moedertaal en, watter invloed het kultuur op leesbaarheid? Hierdie ondersoek bied antwoorde op hierdie en ander verwante vrae ten opsigte van wiskunde eksamenvrae op senior sekondêre vlak.

Protokol analise is gebruik om vas te stel watter leesbaarheidsprobleme ondervind word wanneer leerlinge wiskunde eksamenvrae lees. Drie verskillende taalgroepe, bestaande uit 17-18-jarige leerlinge, het aan die ondersoek deelgeneem: Engels Eerstetaal leerlinge en twee groepe wat Engels as tweede taal gehad het. Een van die tweedetaal groepe het Afrikaans as eerste taal gehad terwyl die ander groep bestaan het uit Afrikane wie se moedertaal nie aan Engels verwant is nie. 'n Raamwerk is ontwikkel om die protokolle te analiseer en het uit die volgende vyf kategorieë bestaan:

- onbekende woordeskat
- strukturele probleme
- onduidelike inligting
- visualiseringsprobleme
- nie-verbale probleme

Gedurende die tweede fase van die ondersoek is leerlinge gevra om die vrae tot 'n meer verstaanbare vorm aan te pas. Leerlinge se aanpassings het leksikale, sintaktiese, diskoers en nie-verbale faktore aangespreek. Sommige leesbaarheidsprobleme wat in die literatuurstudie geïdentifiseer is, is in die empiriese ondersoek geverifieer. Die empiriese

ondersoek het egter addisionele leesbaarheidsprobleme uitgelig wat meerendeels wiskundig van aard is en verband hou met nie-verbale faktore soos wiskunde uitdrukkings. Gedurende die laaste deel van die empiriese ondersoek is 'n samegestelde toets gebruik om die volgende hipotese te toets: *Verbeterde leesbaarheid van die gewone taal wat in wiskunde eksamenvrae gebruik word, sal die prestasie van leerlinge verbeter.* Nege sogenaamde woordsomme is op verskillende maniere gestel: oorspronklik, aangepas en nie-verbaal.

Die hipotese is in 'n hele aantal belangrike gevalle bevestig. Een van die bevindinge van die ondersoek was dus dat leesbaarheidsfaktore nie slegs begrip ten opsigte van wiskunde eksamenvrae beïnvloed nie, maar ook 'n beduidende invloed op prestasie het. Die resultate van die empiriese ondersoek word kwalitatief sowel as kwantitatief weergegee. Ander resultate sluit die volgende in:

- Nie slegs tweedetaal leerlinge nie, maar ook eerstetaal leerlinge het 'n verskeidenhied van leesbaarheidsprobleme ondervind
- Al drie taalgroepe het op die nie-verbale weergawe dieselfde bekwaamheidsvlak getoon. Verskille in die prestasievlakke tussen die verskillende taalgroepe op die verbale weergawes is baiekeer deur taalkundige en kulturele faktore veroorsaak
- Kulturele denkpatrone wat tipies is van 'n leerling se moedertaal, maar nie in die tweedetaal voorkom nie, het dikwels tot begripsprobleme by tweedetaal lesers gelei.

Hierdie ondersoek het sekere gevolgtrekkings vir onderrig- en eksamenpraktyk. Byvoorbeeld, faktore wat die leesbaarheid van gewone Engels beïnvloed, behoort saam met ander faktore in ag geneem te word wanneer wiskunde eksamenvraestelle opgestel word. Verder is daar duidelike aanduidings dat aparte vraestelle vir eerste- en tweedetaal leerlinge opgestel moet word sodat die taalbehoeftes van tweedetaal leerlinge in ag geneem kan word.

Riglyne vir die skryf van meer leesbare eksamenvrae is saamgestel en as 'n oorsiglys aangebied. Voorstelle vir verdere navorsing sluit in die ondersoek na die invloed van leesbaarheid op prestasie in ware eksamen omstandighede.

# CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# ACRONYM LIST

| | |
|---|---|
| E1 readers: | English first language readers of European descent |
| E2 readers: | English second language readers of European descent |
| E3 readers: | English second language readers of African descent |
| HSRC: | Human Sciences Research Council |
| Std 10 HG: | Standard Ten Higher Grade - a final secondary school examination in South Africa |
| CKS: | Contingent knowledge structure |
| STM: | Short term memory |
| LTM: | Long term memory |
| EFL: | English first language |
| ESL: | English second language |
| IBM: | International business machines |
| SGR: | Self governing regions |
| TBVC: | Transkei, Bophuthatswana, Venda, Ciskei |

The following list was used to refer to some of the different South African examination departments operating at the time of this study:

| | |
|---|---|
| CED: | Cape Education Department |
| TED: | Transvaal Education Department |
| OFSED: | Orange Free State Education Department |
| JMB: | Joint Matriculation Board |
| NED: | Natal Education Department |
| DET: | Department of Education and Training |

# INTRODUCTION

In this study the influence of readability on achievement in mathematics is researched. At school level the readability of word problems in mathematics is drawing the concern of more and more mathematics educators (Lagerwerf, 1992; Mulder, 1992). **Chapter 1** explains this concern by indicating that one of the consequences of the greater emphasis on problem solving, as in the "realistic approach" in the Netherlands, is the significant increase in ordinary language in mathematics text. Former mathematics text contained a relatively small amount of verbal language. The emphasis was on abstract mathematics, and in algebra, on symbol manipulation. Although real-life problems have numerous advantages there is a surmise that the corresponding increase in verbal language often causes unnecessary comprehension problems that could affect achievement.

A number of research projects have investigated the effect of linguistic factors on the solving of verbal problems in mathematics (Treacy, 1944; Jerman et al. 1972/ 1973/ 1974; Arter & Clinton, 1974; Cohen & Stover, 1980; De Corte et al. 1985; Lepik, 1990). However, almost all these studies have been restricted to primary or junior secondary school level. There is therefore a need to know what the effect of readability is on problem solving at senior secondary level where non-verbal expressions play an important role.

Researchers have called for studies to investigate the effect of readability in *authentic* writing. Unlike most other studies where questions were constructed for the purpose of the experiment, this study applies basic findings of readability research to authentic questions of previous std 10 HG examinations, a final senior secondary school examination in South Africa. These questions are adapted and then tested for improved readability. But before one is able to improve readability, one first has to ascertain what linguistic factors cause readability problems, and why. For this purpose, a literature as well as an empirical study is supplied. The literature study reported in Chapters 2 and 3 generated research questions which guided the rest of the study while the empirical research discussed in Chapters 4 and 5 was used to explore and formulate an hypothesis regarding readability and achievement in mathematics.

**Chapter 2** approaches readability from a cognitive and psycholinguistic point of view. The cognitive approach to readability investigates the mind of the reader whereas the psycholinguistic approach emphasizes the interaction between reader, writer and text. The question arises what readability problems are experienced by second language readers whose home and school cultures are very far apart. This question is of particular interest to the South African situation because most students at school fall in this category. Therefore, linguistic and cultural factors that influence readability when reading is done in a second language are researched in **Chapter 3**. The term, culture, is used in a much wider sense than the common meaning of "civilization". In this study culture refers to patterns of thinking, feeling and acting - patterns which are learnt throughout a person's lifetime. Bilingual education and the South African situation are discussed in more detail and the effect of linguistic factors on readability is researched by referring to different theories of bilingualism. Cultural thought patterns and cultural experiences are linked to language and readability while certain dimensions of culture are compared with mathematical culture.

**Chapter 4** discusses an empirical study. Think-aloud protocols were used to identify comprehension difficulties encountered by students as they read and solved nine algebraic word problems. Three groups of six std 10 HG students were used. White students of European descent and with English as a first language formed one group (the E1 group). The second group consisted of white students of European descent with Afrikaans as a first language and English as a second language (the E2 group). Black students of African descent with English as a second language represented the third group (the E3 group). The black students had isiXhosa, isiZulu or Setswana as a first language. Two groups of second language readers were included to ascertain the effect of readability factors on two different kinds of second language readers: To the E2 group, English is related to their mother tongue whereas to the E3 group, English is an unrelated language.

After the think-aloud experiment the same students were asked to adapt the word problems in such a way that the text becomes more accessible to their fellow students. Knowledge gained by the protocol study lead to the following main hypothesis: *Improved readability of the common language used in mathematics examination questions will improve achievement.* **Chapter 5** reports on the adaptations and testing of the hypothesis. Three groups of 36 students each wrote a composite test containing the nine questions. Questions were posed in three different versions: the original, an adapted and a non-verbal version. The non-verbal version was not used to test the hypothesis, but was included in the composite test to ascertain in what way the verbal version of a mathematics problem prevents students from demonstrating other mathematical skills. The three groups of students who wrote the composite test once again represented white first and second language learners as well as black second language learners. The results are reported qualitatively as well as quantitatively.

Conclusions and implications of the research results are discussed in **Chapter 6** while suggestions for examination practice and further research are made in **Chapter 7.**

# CHAPTER 1

# THE RESEARCH PROBLEM

**Rationale.**

*Verbal communication will always have the potential of being wrongly understood by someone. In mathematics, miscomprehension has more far-reaching consequences than in most other subjects. Exact reasoning necessitates exact understanding. Whereas a non-verbal mathematics problem can be posed in an international, precise mathematics language - a language students are expected to learn - a verbal problem has to be set in a language that takes the linguistic and cultural aspects of the reading audience into consideration. When writing for large audiences this is a difficult commission. There is a surmise that readability factors in the common language of mathematics examination text unnecessarily cause comprehension difficulties which subsequently affect students' achievement. If this is true, the impediment experienced by second language readers would be even greater. (In this study second language readers refers to students reading a language other than their mother tongue).*

*This chapter will point to the importance as well as the difficulties of reading mathematics text and how the issue of readability is related to the greater emphasis on problem solving in mathematics education, the so-called realistic approach. Once the term readability has been defined one could ask whether readability of ordinary language has any significant influence on students' mathematical performance. Research findings will be reported on linguistic factors affecting students' performance on mathematics verbal problems as well as the role that comprehension plays in the whole problem solving process. Research on these issues automatically lead to the role comprehension plays in the problem solving activities of second language learners. In this respect the situation in South Africa is very actual as the majority of students at school are second language learners.*

*A preliminary analysis of various South African examination papers was executed to form an idea of possible readability problems that could likely affect the performance of first and second language readers. The focus was on verbal problems posed at std 10 HG level.*

*Students writing this senior secondary school examination are usually 17 to 18 years of age. Enough evidence of comprehension difficulties was generated to pursue answers to questions related to readability and achievement.*

**Chapter contents**

## 1.1 Importance and difficulty of reading mathematics text

At school a great deal of assessment is done by means of *written* tests or examinations. One of the implications of this practice is the important role that reading plays in the communication process between examiner and examinee. If there are factors that cause problems in the reading process, comprehension could be inhibited and students' performance during assessment could be affected accordingly.

Mathematics, with its own specific type of language, is also dependent on reading for communicating knowledge and information. A great problem, however, is that many students do not understand what they read. The reading problem in mathematics is increased by the fact that reading in mathematics is more difficult than in most other subjects. Words and symbols have definite, precise meanings. Often the syntax is different to that commonly used in ordinary language. In order to obtain a clear understanding of a mathematics problem, students frequently have to reread a passage a number of times. This is partly due to the special vocabulary of mathematics (Harris & van Deventer, 1990:19).

Mathematical English is not the ordinary English students speak at home or use in the school playground. Reading ability in English is therefore no guarantee for the same level of reading proficiency in mathematics. Mathematics educators correctly argue that reading proficiency in mathematics should be taught (Gullat, 1986; Rakow & Gee, 1988). This is because mathematical English is a highly sophisticated, hybrid language, composed of ordinary English mixed with parts of one or more additional languages. Moreover, *the natural language component is often replete with translations from one of the other symbol systems. For example, phrases such as* ***if and only if, if...then,*** *are direct translations from the sentential calculus. Their precise meanings must be comprehended by the reader before he can hope to assign the author's intended meaning to the passage in which they occur* (Kane, 1970: 579). It is therefore no surprise that a study by Joubert (1985) comes to the important conclusion that reading proficiency in mathematics is an essential prerequisite for achievement in mathematics at tertiary level. Joubert's research investigated the mathematical reading ability of first year university students studying engineering. The findings of the research show that the achievement level of these students in university mathematics displays a higher correlation with their mathematical reading ability than with their achievement level in school mathematics the previous year. School mathematics at that time did not necessitate such a high level of mathematical reading ability as did the mathematics programmes of the university.

But the reading ability of a student is not the only factor that plays an important role when he reads for meaning. Another, equally important factor in written communication is the readability of text. Reading ease and comprehension are heavily dependent on the interaction between the reading ability of a reader and the readability of a text. Comprehension of a text, again, is crucial for the ongoing process of problem solving. If a student has difficulty in understanding the writer's information, mathematical thinking will be impaired.

During the last few years, mathematics educators have become increasingly aware of the important role reading and language play in the successful accomplishment of mathematics tasks (Mulder, 1991:70; Lagerwerf, 1992:36; Weerman, 1994:167). This is mainly due to the fact that world-wide there has been a call for a more realistic approach to mathematics education. In this approach problems are posed in the context of real-life situations. But these contexts have to be described verbally. One could ask whether these contexts, with the corresponding increase in common language, do not make mathematics unnecessarily difficult for students, especially those who receive their education in a language other than their mother tongue. Answers to questions like these will be pursued during the rest of this study.

## 1.2 Realistic mathematics education

One of the underlying principles of realistic mathematics education is that one should develop *and* apply a new concept in the context of the real world if one wishes to convince a child that the concept is worth attending to. The term *realistic* refers to real-life contexts which are familiar or *realistic* to the students concerned. Realistic contexts are carefully chosen to help a student relate his mathematical education to his environment. Real-life contexts are believed to have a further advantage in that they trigger mathematical thinking in a special way (Van den

Heuvel-Panhuizen, 1993). Often the method of the solution arises from the structure of the problem. A memorandum published as early as 1962 and signed by renowned mathematics educators such as Polya, Pollack and Kline, makes a strong appeal to relate mathematics to the real world. These educators argue that mathematics, separated from the other sciences, loses one of its most important sources of interest and motivation. An urgent call is made to precede the introduction of new terms and concepts by sufficient concrete preparation and follow them genuine, challenging applications and not thin and pointless material (Ahlfors et al. 1962).

The Netherlands, one of the leading countries in mathematics education, has implemented a new pre-tertiary mathematics programme, the so-called mathematics A. This curriculum is fully based on the "realistic" approach (Kindt, 1987). In this regard the work of Freudenthal needs to be mentioned. To him mathematics is a *human* experience (De Lange, 1987:98). In Freudenthal's own words: *What is worth being taught? In order to be taught, it must e applicable... what does this mean?... The right perspective is primarily from environment to mathematics rather than the other way round. Not: first mathematics and then back to the real world, but first the real world and then mathematizing* (Freudenthal, 1983:5).

In South Africa the answer to the call for more realistic mathematics education can be detected in the curricula of both primary and secondary schools. One of the main objectives of these curricula is to apply mathematical knowledge in real-life situations. In the senior secondary section, linear programming and differential calculus were specifically included in the curriculum to serve this purpose. However, the curriculum and type of questions used in the final std. 10 HG examination are not comparable with those of the Mathematics A programme in the Netherlands.

It has already been indicated that one of the important consequences of the realistic approach is the significant increase in the use of ordinary verbal language. For example, a textbook from the Netherlands following the new approach, uses *six* pages of *verbal* text to develop the concept of exponential growth from real-life situations *before* giving students the non-verbal rules (Dijkhuis et al. 1988). On the other hand, a South African textbook following a pre-dominantly non-realistic approach, starts with the non-verbal rules on the very first page with virtually no reference to real-life situations (De Jager et al. 1985). This has far-reaching implications for both reader and writer. On the one hand the increase in verbal text demands a higher level of reading ability from the learner. He has to *read* himself into the problem - not simply manipulate numbers of an expression or equation. On the other hand, the realistic approach demands a greater responsibility from the writer - a call for a higher level of readability.

It would be important to know what readability factors in ordinary English interfere with the comprehension of verbal problems and how these factors influence achievement. The identification of such factors would not imply that one should refrain from testing mathematical knowledge in real-life contexts. It would however point to the need for guidelines to improve the readability of verbal problems in mathematics. This is not only necessary for achievement purposes, but also for encouraging skills like interpretation, reflection and mathematization. Development of these skills is an important goal of realistic mathematics education.

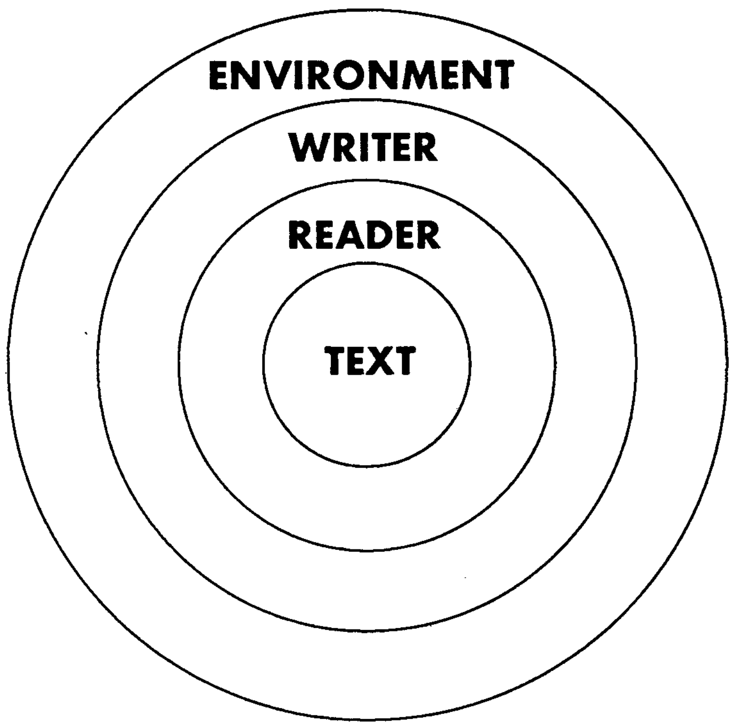
## 1.3 **Defining readability**

Readability needs to be defined because the term is used in several senses. According to Klare (1963:1) readability is referred to in three different ways:

1. To indicate legibility of handwriting or printed matter.
2. To indicate *ease of reading* due to either interest value or the pleasantness of writing.
3. To indicate *ease of understanding or comprehension* due to the style of writing.

Some researchers try to distinguish between readability and comprehensibility. In this case, readability is viewed as a text-based concept that indicates that which can be measured by readability formulas. Comprehensibility, on the other hand, is seen as a reader-based notion that includes elements that are not addressed by the formulas. The majority of researchers have, however, identified readability with comprehensibility (Selzer, 1983:73). This study will consider readability according to the last point of view. Readability in this sense relates mainly to the third definition mentioned above by Klare, although the other two uses are not excluded. According to this approach, factors other than those reflected in the superficial formulation of a text could contribute to readability. Figure 1.1 illustrates the concept of readability in a wider context.

**Figure 1.1 Schematic representation of variables that could contribute to the readibility of text**




Referring to Figure 1.1, *text* variables refer to factors like word or sentence length whereas *reader* variables include elements like knowledge of the context. *Writer* variables could be the writer's subject proficiency while an important *environmental* factor could be the available reading time. Different readability factors interact to generate a certain level of comprehensibility. Readability problems make text low in accessibility and could require extensive mental processing on the part of the reader. Highly accessible text on the other hand, does not require extensive mental processing. Readability and accessibility are closely linked (Kirschner et al. 1992:540). For the purpose of this research, readability is understood as *the ability of the text to communicate the intention of the writer to the reader in such a way that the intended message is comprehended by the reader.*

## 1.4 Verbal problems, assessment and language

A more realistic approach to mathematics education has not only increased the amount of verbal language in textbooks, but there is also an increase in the number of verbal problems used during assessment - either in ordinary class tests or in more formal examinations. There has been a shift from the ordinary manipulation of numbers towards the important area of problem solving. When comparing South African std 10 HG examination papers of the previous syllabus with that of the new one, it appears that the use of natural language has virtually doubled. One of the reasons for this study is the surmise that elements of natural or verbal language might cause comprehension problems that could have a significant effect on achievement. Consider the following question as an example:

> *A particle moves along a straight line so that t seconds after observations have commenced, its distance, s metres, from a fixed point O, is given by*
>
> $s = \frac{1}{2}t^3 - \frac{11}{2}t^2 + 19t - 20 \quad (t \geq 0)$
>
> *Calculate how far the particle will be from the fixed point O at the moment when observations commence.*
>
> ( CED, 1990)

In a class test, students were baffled because they did not understand what the examiner meant by the phrase, *at the moment when observations commence*.

The fact that students in America perform 10% to 30% worse on arithmetic word problems than on comparable problems presented in numeric format, encouraged a study on the role of understanding in the solving of word problems (Cummins et al. 1988). The researchers report that *all* first grade students had one type of arithmetic problem correct when it was presented in numerical form. When the same type of problem was presented as a word problem, only 29% of the students could do the question. The findings suggest that much of the difficulty students experience with word problems can be attributed to difficulty in comprehending abstract and ambiguous language.

However, one is aware that there could be other reasons why students experience so-called word problems as more difficult than ordinary non-verbal ones. Factors like learning experience, the attitude and aptitude of students as well as language proficiency all contribute to the accomplishment of a mathematics task. For the purpose of this study, the focus will be on factors related to ordinary natural language.

Research addressing the effect of linguistic factors on students' performance on word problems has been done by researchers like Treacy (1944), Jerman et al. (1972; 1973; 1974), Arter & Clinton (1974) and more recently by Cohen & Stover (1980), De Corte et al. (1985), and Lepik (1990). The following table gives an overall idea of the reported research. All studies were focused on verbal problems in mathematics. As far as one can follow the literature reports, it seems that the language of the tests was the same as the mother tongue of the students.

**Table 1.1 Research reports on the influence of linguistic factors on students' performance**

| Researchers | Grade level | Important factors |
|---|---|---|
| Treacy (1944) | Junior secondary | Reading skills; Vocabulary |
| Johnson (1944) | Junior secondary | Familiar vocabulary |
| Jerman et al.(1972; 1973; 1974) | Senior primary | Problem length |
| Arter & Clinton (1974) | Fourth grade | Irrelevant data; Problem solving time. |
| Cohen & Stover (1980) | Sixth grade | Diagrams; Extraneous information; Order of information input |
| De Corte et al. (1985) | First and second grade | Semantic relations |
| Lepik (1990) | Eighth grade | Information density<br>Problem solving time |

As indicated in the above table much of the reported research was restricted to word problems at primary or junior secondary level. At this level word problems require little or no reference to mathematical expressions where letter symbols are used as variables. On the other hand, algebraic word problems at senior secondary level demand the use of letter symbols for translating the problem situation from natural language to the corresponding non-verbal mathematical form - a process referred to as mathematization. (Generally speaking mathematization is the representation of a real-life situation in terms of mathematical ideas. For the purpose of this study, the translation process from ordinary English to the non-verbal mathematical form will be referred to as *mathematization)*. This activity is considered to be one of the most difficult stages in the problem solving process. There is a definite need to know more about the influence of linguistic variables at senior secondary level where mathematical models play a crucial role. The present study intends to address the issue.

In spite of the fact that much of the reported research on word problems was confined to primary and junior secondary level, the results are important for senior secondary mathematics as well. It gives an indication of what factors could possibly be determiners of readability for older students. Therefore some of these studies will be discussed in more detail.

Treacy (1944) studied the relationship between reading skills and the ability to solve word problems with junior secondary students. Four of the reading skills in which good and bad achievers differed significantly were associated with vocabulary. This fact suggests the importance of meaning, both general and mathematical. A study by Johnson (1944) confirms the importance of familiar vocabulary. His research with junior secondary students indicates that instruction in vocabulary not only leads to an improvement in understanding, but also to an improvement in the solution of problems based on this vocabulary.

Jerman and his colleagues (1972; 1973; 1974) endeavoured to set up a list of variables other than the aptitude of the student, that could predict the relative difficulty of verbal problems in arithmetic. They considered variables in the format of the problem statement itself. In this way they hoped to formulate a clear set of rules for generating word problems of a specified difficulty level. The aim was to control the difficulty level when preparing instruction materials. In all the studies the negative influence of linguistic factors like unfamiliar words and long sentences were verified. When looking at structural variables, the length of a problem had a greater effect on upper grades than lower grades. A further important observation was that not only problem length, but problem length in *interaction* with other factors like syntax and the number of operations influenced the difficulty level of a verbal problem. One of the profound closing remarks of these studies is that the role of language seems so important that it may outweigh the computational variables as predictors of error rate (Jerman & Mirman, 1974:360).

The influence of irrelevant data as well as the position of the question in the given information was researched by Arter & Clinton (1974). Multiple-choice questions were used to test the performance on word problems in arithmetic. The results show that question placement and irrelevant data did not affect the achievement level *provided* that students had enough time to solve the problems. If, however, the time was limited, irrelevant information had a significant effect on test scores. Question placement seemed to have no significant effect, even when time was limited. Fourth graders were used in this project. One could ask what the effect would be if the subjects were older. The reported research by Jerman et al. (op. cit.) shows that age plays a role in certain linguistic variables like problem length.

The effects of sixth graders modifying format variables was researched by Cohen and Stover (1980). Three experiments investigated format variables that seemed to interfere with average achievement of sixth grade students. In the first experiment gifted students were asked to rewrite word problems to make them easier for their peers who achieved less well in mathematics. Twelve types of changes were suggested by these youngsters. The simplification of vocabulary and the shortening of sentences occurred most frequently. This indicates how much these students perceived word problem difficulty as more of a reading than a mathematics challenge. The researchers very aptly remarked that these young students considered *simplifying word problems* in terms of *improving readability levels* just as reading experts would do.

In the second experiment by Cohen and Stover the effect of three format variables on comprehension and achievement were investigated i.e. *absence of a diagram, the presence of extraneous information* and *the presentation of a number in an order other than the order required for the appropriate algorithmic solution*. Diagrams were added, extraneous information removed and the order of the information changed to the appropriate algorithm order. Large differences in successful problem solution rates were observed in favour of the easier formatted word problems. The more difficult formats appear to interfere with question comprehensibility.

In the third experiment, students were taught to modify the three format variables in word problems to enhance comprehension. In all three cases the students' performance in solving word problems was enhanced. One of the important findings of this study is that reading comprehension involves something more than subject knowledge of the content being read. Readers need what Cohen & Stover call a *comprehension set and specific strategies for teasing out the messages imbedded in deep structures* (1980:198). The research by Cohen & Stover has guided the empirical research of this study profoundly (cf. Chapters 4 and 5).

The influence of rewording verbal problems on students' problem representations and solutions was researched by De Corte, Verschaffel & De Win (1985). First and second grade students were used as subjects. Quantitative and qualitative analysis of data confirmed the hypothesis that the rewording of a problem in such a way that the semantic relationships are made more explicit facilitates the construction of an appropriate mental representation which in turn leads to more successful solutions. This investigation clearly demonstrates that students often fail to solve word problems not because they lack the necessary arithmetical skills, but because they do not succeed in constructing an appropriate mental representation of the problem. They are unable to correctly understand the condensed and often ambiguous information that intends to communicate the context and content. The mental verbal representation mentioned in this study corresponds closely to the *comprehension set* referred to by Cohen & Stover (1980). A part of the methodology followed by de Corte and his colleagues can be recognized in the methodology reported in Chapters 4 and 5 of this study. Subjects were only much older (grade 12 school level).

A study by Lepik (1990) investigated the role of linguistic and structural variables on the solution of *algebraic* word problems at junior secondary level. When solving *algebraic* word problems, letter symbols play an important role and students are expected to mathematize *and/or* interpret a non-verbal mathematical expression. The question is, how do linguistic factors influence the process of mathematization or the interpretation of a mathematical model? The linguistic factors investigated by Lepik were very much the same as those mentioned by Jerman and Rees (1972) and Jerman and Mirman (1974). The influence of word, sentence and problem length were examined as well as other factors like the number of sentences in a problem. Structural variables included issues like the number of relations between quantities and the number of equations required to solve the problem. Two measures of performance were used to assess the solution: the proportion of correct strategies and the average solving time.

The linguistic variables in Lepik's study did not prove to be good predictors of problem difficulty. Of the 16 *structural* variables, 14 reached significance at the 0,01 level. The only linguistic variable that reached significance was the average number of words for each relation. This variable describes the information density in the problem statement. In the study by Jerman and Rees (1972) the linguistic *length* variables (words, sentences, paragraphs) were significant indicators of the proportion of correct arithmetic word problems. Surprisingly, none of the length variables reached the significant level in Lepik's study. However, length variables did prove to be significantly correlated to problem solving time. The best predictor of solving time was the number of words in the problem statement. A significant correlation between length variables and problem solving time is an important factor to keep in mind when investigating the readability of mathematics examination text. In South Africa, all the final school examinations are time-limited.

Although the above-mentioned study by Lepik has generated important knowledge, it must be kept in mind that the subjects were 13-15 years old. It may be that older students have had enough experience to sufficiently master the semantic schemata underlying real-life problems and that factors related to language do not play such an important role at senior secondary level. On the other hand, *understanding the problem* forms such a basic part of the problem solving process, that it seems self-evident that linguistic variables play an important role in the successful outcome of a solution. The present study intends to generate more information on this issue. Whereas this paragraph reported on research focused on specific linguistic factors the following paragraph will point to the role of comprehension within the problem solving framework. Fortunately the important role of comprehension in problem solving has been described extensively in literature and will be used to support further investigations.

## 1.5 The role of comprehension in problem solving

In general, a problem can be considered as any situation where a person wants information, but does not immediately have a method to find it. Many of the verbal problems in the South African std 10 HG mathematics examination fall in this category. In this regard Polya's classic conception of problem solving behaviour is presented as a heuristic process consisting of four phases (Polya, 1946):

- understanding the problem
- devising a plan
- carrying out the plan
- looking back

It is obvious that first of all one has to *understand* the problem. To be able to devise a plan, one has to see clearly what is required and how the various items are connected. Questions like the following have to be asked: What is the unknown? What are the data? What is the condition? If one does not understand the problem correctly these questions will probably be answered incorrectly as well.

Schoenfeld (1985:183) set up a model for analysing problem solving moves. The model was based upon Polya's four phases and contained the following five phases which Schoenfeld calls episodes:

- reading
- analysis
- exploration
- planning/implementation
- verification.

Garofalo & Lester (1985) extended Polya's and Schoenfeld's models by developing a framework for analysing metacognitive aspects on a wide range of mathematical tasks. This framework consists of four broad processes: orientation, organization, execution and verification. These processes are closely related to Polya's four phases.

In all the models used to describe or analyze problem solving, the fundamental starting point is comprehension. The first three categories mentioned by Schoenfeld are all related to comprehension of the given data. Orientation, the first process described by Garofalo & Lester, is once again heavily dependent on the understanding of the problem. As the beginning of the problem solving process, comprehension of a problem is a logical starting point. It would be foolish to answer a question one does not understand.

More recently Artzt & Armour-Thomas (1992) developed a framework for describing the roles of metacognition and cognition within small-group heuristic problem solving in mathematics. The think-aloud behaviours of 27 seventh grade students, selected from three average-ability mathematics classes, were videotaped while they were attempting a verbal problem. Eight types of problem solving behaviour formed the elements of the framework. Each of the problem solving behaviours or episodes was characterized as cognitive or metacognitive. (Metacognitive behaviour is to think about what you are thinking or to monitor your cognition). Initially the episodes of Schoenfeld's framework were used to categorize the students' behaviour. It soon became apparent that two more episodes had to be included, that of *understanding* and that of *watching and listening*. The frequent comments students made regarding the conditions of the problem forced the researchers to include *understanding* as a separate, distinct episode - the same episode recognized by Polya as so important in the problem solving process. Understanding was classified as metacognitive behaviour while reading was taken as cognitive. The episodes were listed as:

- read
- understand
- analyze
- explore
- plan
- implement
- verify
- watch and listen

The 27 students were divided into six groups, two groups of each of the three classes. Table 1.2 illustrates the percentage incidences of cognitive and metacognitive behaviours coded by category for each of the six groups.

**Table 1.2** **Percentage distribution of incidences of cognitive, metacognitive and watch-and-listen behaviours by problem solving groups** *(Artzt & Armour-Thomas, 1992)*

| Behaviour category | Group | | | | | |
|---|---|---|---|---|---|---|
| | 1A | 1B | 2A | 2B | 3A | 3B |
| *Metacognitive* | | | | | | |
| Understand problem | *8,2* | *17,9* | *21,0* | *11,4* | *11,6* | *8,2* |
| Analyze | *4,5* | *8,9* | *8,1* | *2,3* | *1,2* | *4,1* |
| Explore | *3,6* | *11,9* | *16,1* | *15,9* | *25,6* | *15,1* |
| Plan | *4,5* | *4,5* | *4,8* | *0,0* | *2,3* | *0,0* |
| Implement | *4,5* | *3,0* | *0,0* | *0,0* | *2,3* | *1,4* |
| Verify | *0,9* | *1,5* | *1,6* | *9,1* | *3,5* | *0,0* |
| *Cognitive* | | | | | | |
| Read | *8,2* | *6,0* | *12,9* | *18,2* | *4,7* | *6,8* |
| Explore | *44,5* | *13,4* | *14,5* | *18,2* | *15,1* | *11,0* |
| Implement | *5,5* | *0,0* | *0,0* | *0,0* | *3,5* | *2,7* |
| Verify | *0,0* | *4,5* | *4,8* | *4,5* | *4,7* | *2,7* |
| *Watch-and-Listen* | *15,5* | *28,4* | *16,1* | *20,5* | *25,6* | *47,9* |

The results of this research illustrate the importance of metacognitive processes in mathematical problem solving. A continuous interplay of cognitive and metacognitive behaviour, not just a linear movement through the different phases, appears to be necessary for successful problem solving. A typical behavioural pattern could be: read, understand, analyze, read, understand , explore, plan, implement, read, understand, verify. The study by Artzt & Armour-Thomas revealed that a successful problem solver is continually monitoring his work against the backdrop of comprehended information.

The framework was also a useful tool for investigating the *frequency* of different types of problem solving behaviour. Across all groups there were 171 behavioural incidences coded as metacognitive. Of these, 62 were in the category of exploring and 55 were in the category of understanding. The researchers therefore found that 32,2% of the observed metacognitive incidences were attempts to understand the problem which ranks "understanding" second highest among the metacognitive episodes. It is also important to note that this research showed that the only group that could not solve the problem was the group that had the lowest frequency score in the understanding episode. Of course the type of problem used in this research was an important variable. Different frequency patterns for the specific problem solving episodes might have been observed had a different problem been used. However, the general assumption of the important role of understanding was clearly demonstrated.

During the episodes of reading and understanding, readers are forming a comprehension set or a mental representation of the given problem. Bobrow & Brown (1975) refer to this stage as the synthesis of a contingent knowledge structure (CKS). Although the organization and content of the CKS are substantially richer than the collection of raw input data, the synthesis of this structure is heavily dependent on initial information.

By using the concept of a contingent knowledge structure, a problem solving model has been designed to describe the problem solving process of verbal problems in mathematics. Basically the suggested model contains the same elements as those described by Polya, only the *understanding* and *planning* phases have been extended to describe the role of comprehension more pertinently. The first three phases of the model could be linked up with what Polya calls *understanding the problem*, whereas phases 4 and 5 relate to both *understanding the problem* as well as *devising a plan*. The suggested model contains the following phases:

1. reading
2. analysis of information
3. synthesis of a contingent knowledge structure
4. analysis of the contingent knowledge structure with the purpose of extracting the relevant mathematics
5. synthesis of a mathematical model
6. manipulation of the mathematical model
7. verification

When moving through these phases, different types of problem solving strategies may be observed. The reported study by Artzt & Armour-Thomas has indicated that a large percentage of these strategies are related to comprehension. Comprehension strategies could be activities like rereading, interpretation or reformulation.

Moving from the analysis of a contingent knowledge structure or CKS (phase 4) to the synthesis of an appropriate mathematical model (phase 5) is often experienced as a difficult process and commonly referred to as mathematization. Students may have constructed the correct CKS (phase 4) and have all the necessary manipulation skills at their disposal to execute the necessary calculations (phase 6) and yet fail to solve the problem because of their inability to extract the relevant mathematics from the synthesized CKS (phase 5). It is clear therefore that *understanding the problem* is no guarantee that students are able to solve it. Comprehension is a necessary prerequisite, but not a sufficient one for successful problem solving.

Although a variety of problem solving behaviours interact during the problem solving process, the present study will focus on comprehension and ascertain what the ultimate effect of readability factors could be on the successful solution of a mathematical word problem.

### 1.6 Verbal problems and second language learners

If verbal problems cause comprehension difficulties to first language readers as pointed out in 1.4, the difficulties experienced by second language readers may be even greater. Second language readers usually cannot cope with the same vocabulary or difficult syntax as their first language peers. There has been a large amount of research on the difficulties experienced by students who learn mathematics in a language other than their mother tongue (Bishop, 1979; Austin and Howson, 1979; Dawe, 1983; Brodie, 1989). The question is how the language factor affects these second language students' performance on word problems.

Research by Lean et al. (1990) compared the performance of two groups of students, aged 5 to 15 years, with each other. They were asked to do a similar set of 22 arithmetic word problems. One group had English as a first language while the other students were from Papua New Guinea (PNG). To the PNG group English was a second, third or even a fourth language. The test was posed in English, the medium of instruction for both groups. Results show that the two groups used similar strategies, made similar errors and revealed the same order of relative difficulty regarding the 22 problems. However, there was a difference in achievement levels. The conclusion was that the differences in performance between the two groups had to be attributed to differences in the degree of English language proficiency rather than to numerical facility. The researchers warn that, due to the language factor, many young students are simply not ready for the type of verbal arithmetic problems posed at school level.

In another study, Adetula (1990) did extensive research with grade 4 level Nigerian students. There were two groups of 24 students. One group represented students from public schools where the mother tongue is used only during the first three years. After that, both English and mother tongue are used as medium of instruction. The other group came from private schools where the mother tongue is taught as a subject, but English is used in teaching all other subjects, beginning from primary 1. This means that at the year of the experiment, all students were receiving mathematics instruction in English. Students from the private schools had however been subjected to this condition from the beginning of their schooling, whereas the public school students had had, at most, one year of English as language of instruction. Ten arithmetic word problems were presented, in English and in their native language.

Results revealed that both private and public school students performed better when problems were presented in their native language, than when presented in English. The result was only significant in the case of the public schools. These findings are not surprising. Apart from other variables, language is also an important factor. Mathematics learning requires a variety of linguistic skills that second language learners may not have mastered. The researcher reports that stumbling blocks created by the language of the questions often led to faulty translations from words to the appropriate mathematical model.

Adetula argues that there are at least two ways in which language can be an obstacle to a second language reader. These are referred to as inter-problem and intra-problem (Adetula, 1990:352). *Inter-problem* obstacles concern problems confronting a reader who cannot grasp the full meaning of word problems because the language of presentation is foreign to him. This may be regarded as an influence of a second language. *Intra-problem* difficulties concern the wording of the question. A student's performance not only depends on the language of presentation but also on the language of the problem, that is, the way in which the language is used. For example, a misleading meaning of a key word can lead to a wrong operation. Consider the following examples from research by Nesher & Teubal (1975:51). Done with elementary school children in Jerusalem. Here, *more* and *less* are verbal cues for addition problems:

*The milkman brought on Sunday 4 bottles of milk more than on Monday. On Monday he brought 7 bottles. How many bottles did he bring on Sunday?*

Compare this with another version of the same problem:

*The milkman brought on Monday 7 bottles of milk. That was 4 bottles less than he brought on Sunday. How many bottles did he bring on Sunday?*

Nesher & Teubal report that although most students did the addition operation in the first problem, many resorted to subtraction in the second. The word *less* naively suggests a subtraction problem. This type of problem distinguishes between the rote translation of key words and the ability to comprehend the meaning of the problem. Of course, the intra-problem obstacle is experienced by first as well as second language readers.

The students in Nigeria are subjected to a situation which is typical of the rest of Africa or other previously colonized countries: Education is offered in a language other than the mother tongue. The African student therefore suffers a double language obstacle. First he has to understand the foreign language and then he has to do his thinking in a language the intricacies of which he is yet to master.

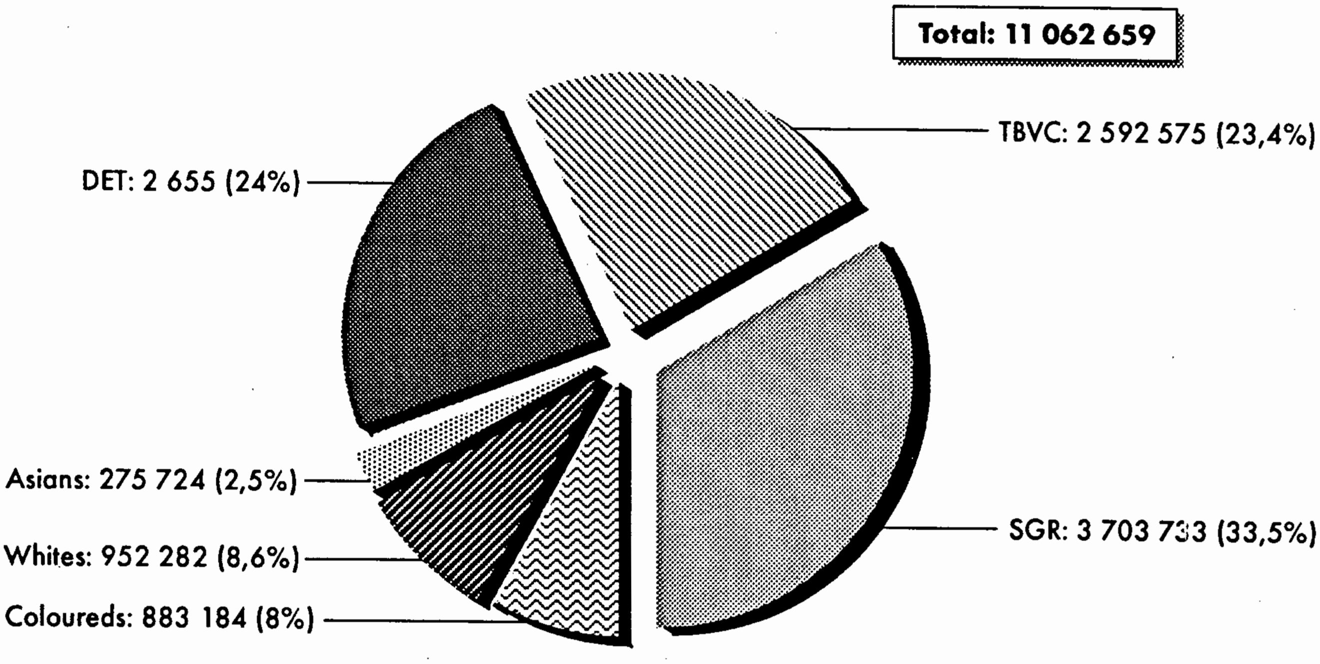
The above research regarding second language readers was limited to primary and junior secondary students. Subjects of the experiments were between 5 - 15 years old. Once again the question arises how the language issue would affect older students. Has the language proficiency of an older second language student developed in such a way that he is able to overcome the difficulties related to second language reading? Is he able to overcome what Adetula calls the inter-problem obstacle? And what about the influence of culture on the comprehension of verbal problems? Language is a manifestation of culture (Whorf, 1956; Wierzbicka, 1991). Moreover, language is communication and communication cannot be separated from culture (Hall & Hall, 1990:3). Discrepancies in language abilities may disappear as students become older, but culture is deeply rooted. Therefore it seems that the question of culture should also be addressed when looking into the language problems of second language readers. What is more, mathematics is a cultural phenomenon with certain cultural attributes (Bishop, 1988). Mathematics text therefore cannot be culture free. In order to investigate these issues in more detail, cross-linguistic and cross-cultural aspects of reading will be reported in Chapter 3.

### 1.7 Second language learners in South Africa

The issue of second language learners in mathematics is especially relevant to a country like South Africa where the highest percentage of students at school are Africans. The following figure illustrates the distribution of std 10 students across all population groups.

**Figure 1.2 Student enrolment (std 10) of all population groups in South Africa, 1993**

*(Strauss et al. 1993)*




The statistics belonging to the DET, SGR and TBVC represent black students. From the above information it is clear that less than 20% of the std 10 school population are Whites, Coloureds and Asians. Although separate education departments for different races do not exist in the new political dispensation, the large number of second language learners still remains. Present statistics will not be significantly different from that of 1993. This means that at least 80% of all std 10 students are Africans. As in the rest of Africa they receive their secondary school education in a language which is not their mother tongue. Formal education in English starts from the fifth school year (std 3), but reports on the increase of students' language proficiency in English are not very encouraging. *The senior primary phase in black education is rather like a bed of sand. As students veer off into English as medium of instruction at the beginning of std 3, all learning drags to a halt* (Southey, 1992:19).

The main cause of this threatening stand-still and subsequent high drop-out after std 3 seems to be the difficulties experienced by the "deep end" language-medium change. By this is meant that students make a complete changeover from their mother tongue to English on a total of *ten* subjects in English. The Threshold Project reports that students simply cannot cope with the massive range of new vocabulary, structure and concepts (MacDonald, 1992:162). According to this project the vocabulary requirements alone increase from about 800 words to approximately 7000 . In this regard Walters (1993:21) reports on a study which reveals that African students at std 3 level could generally not do tasks depending on:

- reading their textbooks with understanding
- reading skills such as skimming or scanning for specific information
- interpreting maps and diagrams.

The result of these and other language related problems leads to a learning culture which is very much teacher-centred and memory-dependent. Although the above-mentioned study focused on Geography learning at std 3 level, a group of post-graduate African students assured the researcher that this was the essence of their *entire* school experience. This learning style is detrimental to mathematics. To do mathematics, a good deal of independent thought and logical reasoning is needed. A dominant memory-dependent and teacher-centred learning style cannot develop mathematical thinking effectively.

In a newspaper report (Beeld, 18 January 1989), a spokesman from the Department of Education and Culture isolated language as one of the chief reasons for the poor scholastic performances of African matriculants (std 10 students). According to him, most black students enter their final school examination without an adequate language proficiency in English - the medium of the examination. The result is that many students take recourse to memorization with very little understanding. This has detrimental effects on subjects like mathematics and physical science where understanding and logical development of arguments are crucial for problem solving. Swanepoel (1989:1) also points out that in all studies of poor scholastic performance of black students (compared to students in white schools) the language problem is singled out as one of the contributing causes.

For various reasons the African community is vehemently opposed to solutions that extend the phase of mother tongue education. To many of these parents the African languages do not enjoy much economic or social status. English, to them, is the key to higher education and job opportunities. Until parents and students have reason to believe that there would be no disadvantage in having secondary schooling in the home language, this negative attitude will prevail (National Education Policy Investigation, NEPI-report, 1992:77).

The above remarks are sufficient to indicate that the language problem in black education is paramount, multi-faceted and deserves profound research. As regards this study, the described situation clearly shows that linguistic factors in the text format of examination questions could very likely have an important influence on the achievement level of African students when they solve word problems in mathematics. The language issue in black education is even more reason to investigate the readability of mathematics text.

## 1.8 Readability problems and assessment

More than twenty years of teaching experience have convinced the author that difficulties in text comprehension are experienced more dramatically during examination conditions. During examinations, students are not allowed to ask for any clarification of text. When they experience comprehension difficulties, they often lose much time in their search for meaning. Loss of time increases the possibility of not being able to complete the examination paper. This realisation almost always adds to examination tension. Research by Horn & Dollinger (1988) found a relationship between test anxiety and performance. They claim that the higher one's test anxiety, the worse one's performance in an examination. A large group of students also suffer from mathematics anxiety. Although they may be well skilled in a mathematics task, the anticipation of possible incompetence often blocks the operation of necessary skills (Skiba, 1990:188). Readability problems - especially during examination conditions - seem to aggravate this condition.

When addressing the comprehensibility of final school year examination papers, factors like language proficiency and reading ability of students are beyond the control of the examiner. On the other hand, readability factors like ambiguity or unfamiliar words are within his power to change. The examiner has a definite responsibility to make the text as accessible as possible to *all* students. Obstacles related to comprehension difficulties must be avoided. In Israel, Kirschner and her colleagues conducted a research project on the comprehension problems that university students encounter during language testing (Kirschner, et al. 1992). In no uncertain terms they emphasize the truism that it is the test writer's task to define, identify and subsequently remove any potential comprehension difficulties inherent in test questions.

As has been indicated, this study will focus on exactly this task, namely the defining, identification and removal of comprehension difficulties in examination questions. Furthermore, the study intends to determine the *effect* of these comprehension difficulties on achievement. In a way the research relates to what Selzer calls the need to observe the effects of readability in real writing (1983:83). According to him, more research opportunities should be created where basic research in readability can be applied and tested in the real world of on-the-job writing. Research is often done under highly controlled and artificial laboratory conditions. Readability factors that affect comprehension are studied in isolation from other words, sentences and paragraphs. Will they have the same effect in the context of real discourse? Answers to questions like these will be pursued in the following chapters.

In an examination question, which is an example of real writing, information is given and a question is asked. The question asked is referred to as metatext whereas the conditions of the problem form the text. Readability problems could be related to text, metatext or to generated answers. Furthermore, text, metatext and answers should complement each other. If there is any misfit between these elements, students experience difficulties in comprehending the situation. For example, if a question on the initial speed of a cyclist gives an answer of 36 km/h, an examinee could think that he had read or understood the information wrongly and revert to the episodes of reading and understanding. Examinees usually expect cyclists to start at zero. Problems like this make the text low in accessibility and could therefore require extensive mental processing on the part of the reader. As indicated before, readability and accessibility of text are closely linked (cf. 1.3).

Before doing the empirical research reported in Chapters 4 and 5, a preliminary analysis of examination papers was done from a readability point of view to identify readability factors that could possibly influence comprehension and achievement. Many years of classroom practice supported the more or less intuitive investigation. The purpose of the analysis was to gain some indication of the range of the problem. Papers representing five examining boards were used. The following problem areas, all related to readability, were identified:

- a time problem (cf. 1.8.1)
- poor formulation, often causing ambiguity (cf. 1.8.2)
- senseless solutions (cf. 1.8.3)
- non-verbal interference (cf. 1.8.4)
- visualization difficulties (cf. 1.8.5)
- unfamiliar contexts (cf. 1.8.6)

### *1.8.1 A time problem*

Time as a readability factor means comprehension difficulties are not caused by the language or style of the text, but by insufficient time allowed for questions. It has already been indicated how important the continuous interaction of cognitive and metacognitive processes are for successful problem solving (cf. Artzt & Armour-Thomas, 1992). Students repeatedly reread a passage to monitor their understanding and other problem solving activities. If examination time starts running out, anxiety sets in and the interaction between reading, understanding and other cognitive processes is impaired. Smith (1982:12) explains this situation by saying that anxiety increases the demand for visual information and has the paradoxical consequence of making it harder to see the text. The more one reads the less one sees. The result is that the text becomes less accessible.

The emphasis on problem solving suggested by the 1985 senior secondary syllabus necessarily results in the testing of processing skills like interpretation, mathematization and visualization. All these skills are time consuming. A student cannot just read superficially. Another type of reading style is needed to evaluate information critically and detect underlying relationships. Research done by the Hewet Project (Kindt, 1987) identified time-limited tests as being a hindrance to the goals set by realistic mathematics education. It seems that process orientated activities are not well suited to time-restricted tests (De Lange 1987:47).

During the above-mentioned analysis of question papers, it appeared that the time allowed for the execution of necessary skills during examinations was not always realistic. Consider the following question:

> *A cyclist sets out to travel a distance of 20 kilometres. After an hour his bicycle breaks down and it takes him two hours more to walk the rest of the way. If the bicycle had broken down after he had travelled 6 kilometres, it would have taken him 4 hours to travel the entire distance.*
>
> *Find his average speed in kilometres per hour when cycling and walking.*
>
> (National Examinations, 1990)

Although at first sight the above question does not seem to have readability problems, the accessibility of the text is very likely impaired by insufficient problem solving time. In std 10 HG mathematics examinations, students have three hours to answer an examination paper of 200 marks. The number of marks therefore gives the student a rough indication of the amount of time he could spend on a question. The above question counted 11 marks which means students had about 10 minutes to answer the question. For many students this was not enough.

Not just the quality of the text but also the *quantity* has to be taken into consideration when allocating examination time. The increase in the verbal mode demands more reading time for the obvious reason that there is more printed matter. While analysing the examination papers had the feeling that this fact had not always been taken into consideration. Many times it seemed as if time had only been allowed for the execution of non-verbal manipulative skills without providing for other processes like reading, understanding or planning. The following question illustrates the problem:

> *A man's salary for 1991 is R20 000 and his expenses are R16 000. His salary increases by R2 000 per annum while his expenses increase by R2 500 per annum. Each year he manages to save the excess of income over expenditure.*
>
> *- Represent his accumulated savings in an arithmetic series.*
>
> *- After how many years will his savings be totally depleted?*
>
> *- To what amount must he reduce the rate at which his annual expenses increase so as to let his savings last for 21 years?*
>
> (OFSED, 1993).

A total of 7 marks was allocated to the question as a whole. This means students had only about six minutes to move through all the problem solving phases efficiently.

A time problem is also experienced by students who write examination papers that include a large number of multiple-choice questions. For example, in 1991 one of the examination papers included 35 multiple-choice questions in the std 10 HG first paper. Each question counted 2 marks. This meant that in just more than an hour students had to read themselves into 35 different kinds of problems. Many of these questions required a high level of intelligent reading and understanding. The specific examination paper was prepared for amongst others, visually disabled students as well as second language learners. Referring just to second language learners, Cummins (1984) argues that these students take about twice as long to read as their first language counterparts.

Consider two questions from the same examination paper, one from the multiple-choice section of the paper.

> *P(x;y) is a variable point such that its distance from the point Q(1;2) is twice its distance from the line x = -2.*
>
> *What is the equation of the locus of P?*
>
> *(A)* $4y + 3x - 4x - 16y + 16 = 0$ *(C)* $y - x - 4y - 8x - 3 = 0$
>
> *(B)* $y - 3x - 18x - 4y - 11 = 0$ *(D)* $2y + x - 8x - 8y + 6 = 0$

A similar question in the non-multiple-choice section reads as follows:

> *Given the points A(-2;4) and P(x;y)*
>
> *Determine the equation of the locus of point P which is equidistant from the x-axis and the point A.*

(National Examinations, 1991)

The first question was worth 2 marks while the second counted 7. More time was therefore allowed for the second question but it was quicker to do this question because of the following reasons:

- The words *variable point* in the first question are more difficult to understand than for example, the point P(x;y).

- The line x = -2 is more difficult to understand than the x-axis

- The mathematization of the first question is more difficult. Many students have a problem with the translation of the expression, *twice its distance.*

- The answer to the first question is prescribed in a certain form. A possible manipulation mistake can cause the student to lose much time when "looking" for the correct answer.

Time allocated for the second question was not excessive - the time problem is situated within the multiple-choice question. It seems as if time *allocated* does not always correspond to time *needed.*

More reading time is also necessary when verbal data and corresponding sketches are not on the same page. The student has to page back and forth to link the two sets of data. Because all the information is not visually available at the same time, the text is less accessible. (cf. Question 9, CED, 1993)

All of the above are examples of questions that most probably cause a time problem. Classroom practice has confirmed that shortage of time contributes directly or indirectly to readability problems. If the immediate question is not influenced, a following question is likely to be.

*1.8.2 Poor formulation*

Poor or ill-planned formulation of an examination question can cause a variety of comprehension difficulties. In many cases it leaves the reader in a state of uncertainty or confusion. Consider the following question:

> *It is estimated that t years from now the circulation of a magazine will be given by the function*
>
> $m(t) = 50t^2 - 200t + 3000$

*- At the time of first publication, what was the circulation of the magazine?*

*- Derive an expression for the rate at which the circulation will be changing 3 years from now*

(House of Delegates, 1993)

The information concerning the circulation of the magazine is confusing. It is not clear when the magazine was published for the first time. The following questions illustrate various degrees of uncertainty that could occur:

- How often is the magazine published?
- What is the unit of circulation?
- Must $t$ be substituted by $\frac{1}{365}$ or by O in the first question?
- Will a second language reader know what is meant by *circulation of a magazine*?
- Can students relate, *the rate at which the circulation of a magazine will be changing 3 years from now*, to their real-life experiences?
- Does the examiner mean *average* rate or *immediate* rate?

During the intuitive analysis, questions were encountered where it was clear that the examiner had wrongly assumed that he and the examinees shared the same knowledge of a situation. Important detail was missing and often the language used was not the language in which students experience the described context.

### *1.8.3 Senseless solutions*

Senseless solutions interfere with readability in that they cause a reader to think that he has understood the information incorrectly. Although the text might be formulated perfectly correctly, the senseless solution forces the reader to revert to unnecessary comprehension strategies. To check whether the solution of a problem makes any sense is one of the important phases in the problem solving process. Polya (1946) calls this phase "looking back". By doing so, the problem solver is in a position to judge whether he has solved the problem correctly or not. If the answer does not fit the given situation, the student usually first checks for manipulation mistakes. If nothing is found he will probably think that he has misunderstood the problem and start reading all over again. It is clear that "correct" solutions which are senseless can cause unnecessary comprehension problems. The student can lose much time by re-reading in an attempt to understand the problem correctly so that he can reach a sensible answer. The following question illustrates this problem.

*A farmer has 500 boxes of peaches for sale at R6 per box. If he keeps the peaches with the intention of selling them later, the price increases by 10c per week. Every week 5 boxes become polluted and therefore cannot be sold.*

*- After how many weeks did he sell the peaches if he received R 3 182 for them?*
(JMB, 1991)

The "correct" answer to the above problem is 14 or 26 weeks which of course is senseless. Peaches do not last that long.

*1.8.4 Non-verbal interference*

This refers to the use of the letter symbol in such a way that it interferes with the processing of information. Research has identified many misconceptions of the letter symbol (Hart, 1981) and examination text should not add to this confusion. Note the different ways in which the letter symbol R and F are used in the following question:

*A company manufactures fertilizers for the industry. On delivery each bag must contain at least 108 kg of a mixture made up of compounds R and F, where at least 12 kg of F are used. The mass of R in the mixture must not exceed twice that of F. 1 kg of R costs R5 and 1 kg of F costs R8. The cost of a bag of the mixture must not exceed R720.*

*- Set up a system of inequalities to represent the constraints involved.*

*- If the pollution ceiling given by the index $P = 5R + 4F$ is exceeded, the pollution*

*levels on the farms will be unacceptable. Calculate the maximum mass of R and F respectively which respect the constraints above and at the same time fall within the pollution restrictions.*

(JMB, 1991)

In one and the same examination question the letter symbol R was used in three different ways, i.e. as an abbreviation for rands, as the name of a fertilizer and as an undefined variable. This confusing practice could interfere with reading ease and therefore also with the processing of information.

### *1.8.5 Visualization difficulties*

Difficulties in this category are often caused because of abstract information. In the following question students did not know to what particle or what situation the examiner was referring.

*A particle moves along a straight line so that, t seconds after observations have commenced, its distance, s metres, from a fixed point 0, is given by*

$$s = \tfrac{1}{2}t^3 - \tfrac{11}{2}t^2 + 19t - 20 \qquad (t \geq 0)$$

*Calculate when and how many times the particle will pass the fixed point 0.*

(CED, 1990).

After this specific examination, students had many questions to ask about the above information. For example: *"How can a particle move in a straight line and still be able to pass a fixed point a few times?"* or *"I couldn't understand what was going on. What is the particle and what is the fixed point 0 supposed to be?"* It was clear that *one* of the comprehension problems was due to the fact that students had difficulty in visualizing the situation. A graph would have been helpful but was only introduced at the end of the 34 mark question. Many students did not get that far.

### *1.8.6 Unfamiliar contexts*

One of the reasons for the realistic approach in mathematics education is the fact that students can relate their education to the world in which they live. Bishop (1988) supports this view by strongly arguing against impersonal education. If, however, the contexts used in real-life questions are not part of students' real-life experiences, the goals of realistic mathematics education will not be achieved. Mathematics will still be removed from the personal lives of students. Moreover, where the context is intended to aid comprehension, an unfamiliar context does exactly the opposite. The question from the CED 1990 examination referring to a moving particle is a good example (cf. 1.8.5). Students who do not take physical science as a subject are unfamiliar with contexts like these. Even physical science students do not experience their real world in terms of abstract particles and points.

## 1.9 Research questions

The readability problems reported in 1.8 have mainly been identified intuitively. An important next step is to do a literature survey and ascertain what factors have been reported in literature as having an influence on the readability of common language text. Primarily this study is concerned with the influence of readability factors related to *common language*. This is because algebraic word problems are, to a large extent, described verbally. For a writer it would be important to know what makes one text more readable than another. What readability factors inhibit written communication? Another important aspect is the second language reader. Is there enough evidence that leads one to believe that readability can be influenced by factors like mother-tongue interference and culture? With these issues in mind the ongoing research will be guided by the following research questions:

1. What type of readability factors in the common language of mathematics text prevents a clear understanding of examination questions?

2. Are these readability factors restricted to second language readers only or do first language readers experience similar problems?

3. If there are readability problems in the text, how do they influence mathematical achievement?

4. What is the influence on the African pupil whose home and school cultures are completely different?

Psycholinguists, cognitive theorists and anthropologists have done much to find answers to these questions. Although more research is still needed, a large body of knowledge is available at present and is utilized in the following chapters. Whereas the experiments reported in this Chapter refer to linguistic factors in mathematics questions, the research reported in Chapters 2 and 3 will focus on linguistic factors in ordinary English.

# CHAPTER TWO

# READABILITY: A COGNITIVE AND PSYCHOLINGUISTIC APPROACH TO VERBAL LANGUAGE IN MATHEMATICS TEXT

## Rationale

*In this Chapter an answer is sought to the following question: What readability factors in ordinary English could possibly prevent a clear understanding of a mathematics examination question? Apart from difficulties experienced in the mathematics register, students also have to cope with a great deal of ordinary, common language. If this common language were to have readability factors that inhibit comprehension, the reliability of test scores could be affected. This will influence the validity of inferences made from these scores.*

*Readability formulae do not seem to be able to guide writers in the production of more readable mathematics text. If readability formulae cannot help, what can? Contemporary research on readability calls for more than just an investigation on the level of word and sentence length. A closer look at literature suggests that readability be approached from a cognitive and psycholinguistic point of view.*

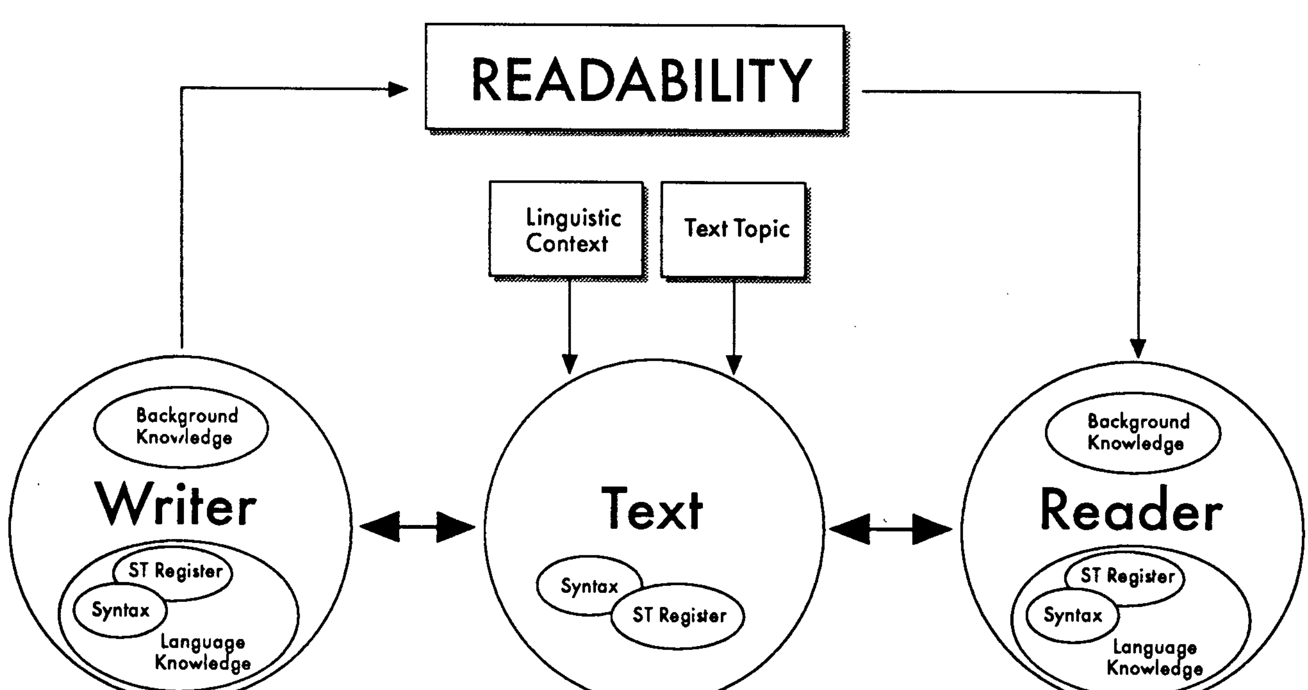
*Cognitive psychology investigates the way information is perceived and processed in the mind of the reader as well as the way information is expressed and verified. Psycholinguistic theory on the other hand takes the language area of the brain as the starting point to understand how people read and why some texts are more readable than others. There is a strong emphasis on communication by means of language. Since language is an integral part of comprehension, readability will also be considered from this point of view.*

## Chapter contents

## 2.1 The influence of readability on reliability and validity of test scores

A key link between the reader and the writer is readability. Communication in written discourse takes place when the reader and writer meet via the text. Although one usually refers to readability of text, many researchers agree that readability is influenced by many factors other than the superficial formulation of the text (Huckin, 1983; Samuels, 1988). Characteristics of the reader as well as the writer are as important to readability as the properties of the text. For example, information written by a computer expert may be highly accessible to a specialist in the field. To a computer novice, the text may be hardly readable. This implies that during reading, there is a constant interaction between reader, writer and text. Strother (1990) presents the following model to show the interactive nature of readability.

**Figure 2.1 An interactive reader-writer readability model**



As has been mentioned before, the focus of this study is on the readability of *common* language in mathematics text. The aim is to identify readability problems that cause comprehension problems during assessment with a view to producing guidelines for writing more accessible text. The question could arise whether the removal of external readability interferences would not prevent the development of a student's reading ability. The answer to this question is closely related to the issue of reliability and validity.

Reliability and validity are two terms related to test theory. Both terms give an indication of the *value* of the observations made during testing. Reliability is the extent to which test scores are consistent, accurate and reproducible. According to Slotboom (1987), reliability is always related to test scores and indicates the precision of the measuring instrument. A reliable instrument will always give the same result when measuring the same subject under the same conditions.

Reliability of test scores can easily be jeopardized if the measuring instrument is affected by factors other than the abilities one wants to measure (Bachman, 1990:160). In mathematics, a commonly used measuring instrument is a written test question. To determine the reliability of this instrument, it is impractical to repeat the same question with the same students under the same circumstances because students may have learnt from their first experience or thought about the question afterwards. The reliability of a question is therefore measured by determining how consistently the test score of a question correlates with the test scores of other, similar questions. A high correlation between scores or a small standard error of measurement indicates reliability. The argument is based on the assumption that to be reliable, the test score of one question must consistently show a high correlation with scores of other, similar questions.

An important step towards reliable test scores is adhering to the principles of testing. Gronlund (1968:4), Cockroft (1982:159) and De Lange (1987:179) consider the following principles to be of paramount importance for the construction of achievement tests:

- Achievement tests should measure clearly defined learning outcomes that are in line with the instructional objectives.

- Achievement tests should measure an adequate sample of the learning outcomes and subject matter included in the instruction.

- Achievement tests should give students the opportunity to demonstrate what they know rather than what they do not.

- Achievement tests should be made as reliable as possible and should then be interpreted with caution.

Factors that violate the principles of testing can affect the quality of the measuring instrument. The error in the measuring instrument may distort the test score. For example, if a question contains ambiguous information this could prevent students from demonstrating what they know. The ambiguous information could cause comprehension difficulties in varying degrees which in turn affect the reliability of test scores.

Whereas reliability refers to test scores, validity refers to the *interpretation* of these scores. Bachman (1990:237) points out that validity always refers to the degree to which evidence supports the inferences that are made from test scores. The inferences are validated, not the test itself. These inferences are closely related to the aim of the question. To interpret a test score, one should ask: *Have I tested what I had in mind?* The measuring instrument could contain a disturbing element - like redundant information. If students are expected to cope with redundant information in an examination question, but they were not taught to do so in ordinary classroom practice, the aim of assessment is not in harmony with the instructional objectives pursued in the teaching situation. This violates one of the principles of testing that could prevent the aim of the test from being achieved. In this case no valid interpretation of test scores can be made. If the test score for a question with redundant information were low, one would not be able to interpret the test score as an indication that the specific mathematical knowledge has not yet been mastered. A primary concern in testing is therefore not only that scores should be reliable, but that the interpretations and uses made of scores be valid.

In a mathematics test, the prime objective is the testing of certain mathematical abilities and not the ability to decipher the language of a test question. The removal of factors causing readability problems is not intended to jeopardize the development of reading ability, but rather to promote reliability and validity.

## 2.2 Readability and readability formulae

To understand the factors that influence readability is no easy matter. During the past decades researchers have developed various formulae to measure the readability of text. Two popular readability formulae are the ones by Flesch and Gunning. Flesch's formula was in use as early as 1943 (Flesch, 1948:229).

FLESCH'S READING EASE SCALE

Reading ease = 206.835 - 0.846wl - 1.015sl

where wl = number of syllables per 100 words
sl = average number of words per sentence

(Reading ease corresponds to a number between 0 and 100, with 100 representing the easiest reading.)

Gunning's "Fog Index" works in the following way:

GUNNING'S FOG INDEX

- Start by using a 100 word passage. Find the average number of words per sentence. Then find the number of "difficult words" in that same passage. A "difficult word" has three or more syllables. These words do not include proper names, combinations of simple words (such as *horsepower*) or verbs whose third syllable is *-es* or *-ed* (such as *contracted*). Add the average sentence length and the number of difficult words. Multiply this sum by 0,4.

Example:

| | |
|---|---|
| Average number of words per sentence | 13.9 |
| Number of difficult words: | 16 |
| | 29,9 |
| | 29.9 x 0.4 = 12 |

The result represents an approximate grade level, therefore in the above example a twelfth grade student could understand the writing (Markel, 1992:158).

The main factors focused on in these formulae are sentence and word length. These factors have indeed proved to have a significant influence on readability (Coleman, 1962). But these formulae unfortunately have grave shortcomings. They have been developed to measure the readability of ordinary language. As soon as the formulae are applied to technical, scientific or mathematics material, many problems arise. Factors other than word and sentence length start playing an important role in the readability of the text. A short word like *speed* could give more readability problems than a longer word like *objective*. In mathematics other factors like a high level of precision as well as the logical structure of a sentence are important issues. These are but two factors not addressed in readability formulae. For example, readability formulae are not able to deal with the overall structure of text. By focusing on the readability of single sentences the larger problem of comprehensibility is ignored (Flower, et al. 1983:41). When reading mathematics text, clarity is more important than brevity.

There are two other characteristics of a mathematics examination paper that distinguish it still further from ordinary expository prose. One is the difficulties caused by the various aspects of the mathematics register. An example would be the logical connector, *if and only if*. The words in this expression are easy, common language words. However, in the context of mathematics, this word combination is heavily loaded with meaning. A readability formula is not able to capture the comprehension difficulty communicated in this way. Another important issue is the fact that a mathematics examination paper is a functional document - people who read it, do not merely *gain* information. They have *to do* something with it. These two properties of mathematics text cause readability problems that cannot be quantified by readability formulae.

Selzer (1983:76) argues that unqualified suggestions that technical writers should use short sentences are misguided. Contrary to often quoted opinion, there is no evidence that shortening sentences will always make writing more comprehensible. The same applies to the call for shorter words, as has been illustrated above. Reading is far too complex for any formula to predict readability with perfect accuracy.

Essentially, readability is concerned with comprehension difficulty. The readability of a text indicates the degree of comprehension excited in the mind of the reader. To understand the factors underlying the comprehension difficulty of a text, Huckin (1983) suggests a more cognitive approach to readability. In the following section, a cognitive view of readability will be given. The theoretical framework of Huckin forms the basis of this discussion.

## 2.3 Reading and a cognitive approach to readability

Reading is an intricate and concealed working of the human brain - a highly sophisticated cognitive skill. Researchers are still far from knowing what exactly occurs in the brain of a reader while reading. But more and more researchers realize that reading is a multi-faceted activity and cannot be understood without considering factors like perception, cognition, linguistics and motivation. Smith (1982) emphasizes this by stating that an understanding of reading requires acquaintance with research in a variety of disciplines. Fortunately, a considerable body of scientific knowledge based on experimental research in cognitive psychology has become available in recent years. Instead of focusing on the written product as traditional readability research has done, this newer approach focuses on the reading *process*. It investigates the mind of the reader. As a result of these studies, a number of theoretical concepts have been found to be particularly relevant to the readability of technical documents. These concepts can be considered to be of equal importance to the readability of mathematics text. Huckin (1983) considers the following theoretical concepts to be the most useful for producing readable text: "schema theory," "activated semantic contexts," "the levels effect" and the "leading-edge strategy."

### *2.3.1 Schema theory*

In schema theory, the word "schema" is not used in the ordinary sense of a static written or graphical representation of knowledge, but to indicate a mental construct in the mind of the reader. Cognitive theorists like Bartlett (1964) and Piaget (cf. Tagatz, 1976:55) refer to schemata as higher order cognitive structures that serve a crucial role in providing an account of how old knowledge interacts with new in perception, language, thought and memory.

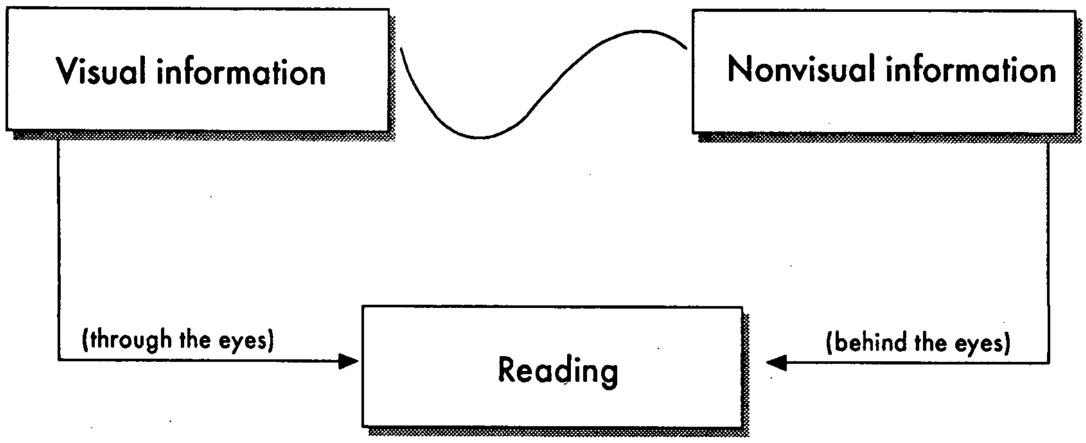
Schema theory explains why knowing something about a subject makes it easier to absorb new knowledge about that subject. New knowledge is not only processed more easily, but it can also be retrieved in more detail. Studies by Bransford & Johnson (1972) confirmed this phenomenon by showing that relevant contextual knowledge is a prerequisite for comprehending a piece of prose. Students supplied with appropriate information before they heard test passages not only demonstrated increased comprehension ratings, but also showed an increase in recall scores.

According to schema theory the human mind constructs abstract concepts or schemata on the basis of experience. These schemata or mental constructs of concepts are considered to have fixed place holders for the constituent features of the concepts (cf. Huckin, 1983). The more experience a person has of a concept, the richer the schema because of more place holders for the different features. In the communication process these schemata have the powerful ability to induce inferences in the reader. If a writer and reader share the same schemata, it is not necessary for the writer to mention all the characteristics of the concept under discussion. The reader will simply generate missing data by inference.

Schema theory links up with what Smith (1982) calls non-visual information. Access to visual information is a necessary part of reading, but it is not sufficient. One may have a wealth of visual information in a text and still not be able to understand what is being read. The language of the text may be completely unknown or the exact meaning of crucial words may be lacking. Non-visual information also includes prior knowledge of the subject matter. Many people will not understand an article on differential calculus, not because they are unable to read or because there is some inadequacy in the text, but because they lack appropriate non-visual information. *Non-visual information is easily distinguished from visual information - it is carried around by the reader all the time; it does not go away when the lights go out...reading always involves a combination of visual and non-visual information. It is an interaction between a reader and a text* (Smith, 1982:11). The sharing of the same schemata or non-visual information plays an important role in any communication process. The following model illustrates the two sources of information when reading:

**Figure 2.2 Two sources of information in reading**

*(Model by Smith, 1982)*

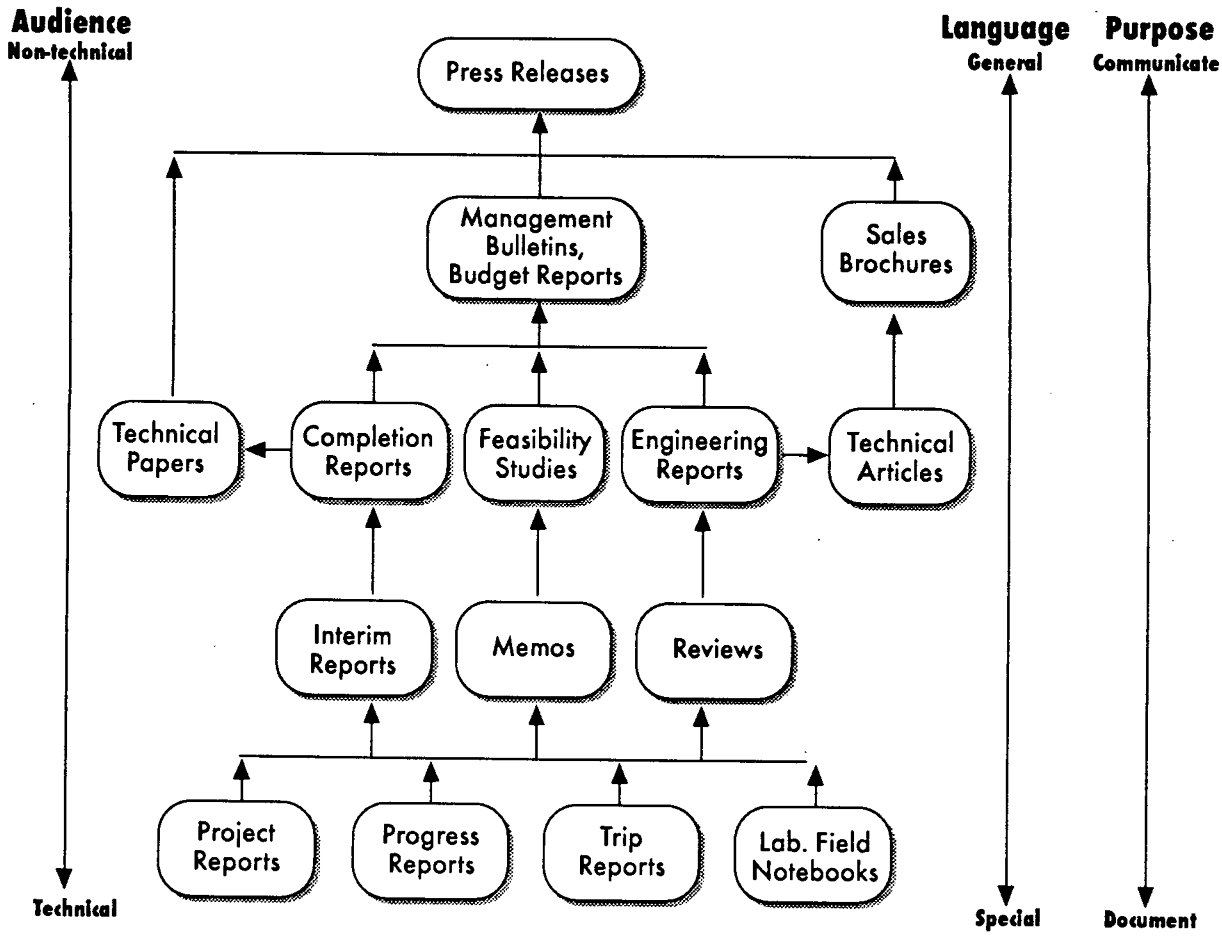



Schema theory also explains why writing for a specialist (i.e. a native reader of the language and competent in the subject matter), differs from writing for a non-specialist. If writer and reader share the same schema or non-visual information, not all detail needs to be explicitly mentioned in the visual information. By means of schema based inferences, the reader is able to deduce information and give meaning to what is being read. But when writing for non-specialists, the call for more explicit detail is imperative. The more novel the situation is to the reader, the greater the responsibility of the writer to keep the poor schemata of the reader in mind. The language should be less technical, the detail more abundant and the structure of the discourse more informative. The communication will still be based on shared schemata but then on those available to the non-specialist reader. If care is not taken here there can be a complete breakdown in the communication process. Perhaps the best examples of such misunderstandings are to be found in cross-cultural communication. During a recent visit to an African school, someone asked an African student to go out and look for a bottle top on the playground. She proudly came back with the upper *neck* of a glass bottle!

The content and style of a text meant for a specialist reader must therefore necessarily differ from that meant for a non-specialist reader. The model by Ulijn & Strother (1994) illustrates how the difference in audience, language, use and purpose of a text places written discourse somewhere in a hierarchy of documents.

*Figure 2.3* **The documentation hierarchy**

*(Model by Uljin and Strother, 1995)*




In a mathematics examination, the schemata of examiners and examinees differ substantially. Referring to language use only, the examiner is fluent in mathematical English - he *speaks Mathematics like a native* (Pimm, 1987:1). Most examinees do not share this competency. English first language examinees might be fluent in ordinary English, but as regards mathematical English they do not operate on the level of a "native speaker". Compared to E2 and E3 examinees, the examiner is also much more competent in the use of ordinary English. This situation is illustrated by the following table:

**Table 2.1 A schematic comparison of language competency of examiners and examinees**

| | Ordinary English | | Mathematical English | |
|---|---|---|---|---|
| | E1 | E2/E3 | M1 | M2 |
| Examiner | x | | x | |
| Examinees (E1) | x | | | x |
| Examinees (E2 and E3) | | x | | x |

E1: First language competency in ordinary English

E2/E3: Second language competency in ordinary English

M1: First language competency in mathematical English

M2: Second language competency in mathematical English

From the above, it is clear that the language of the mathematics paper can be a source of misunderstanding. Even though the language is not the object of measurement, it is the instrument used to communicate the information. For testees this is crucial. The instrument of measurement is handled by a specialist. If he does not select more commonplace schemata, miscommunication could be the result.

But it is not only the language of the examination paper that could give rise to readability problems. The described situation or concept could be so unfamiliar to the reader that the sharing of schemata is virtually impossible, unless the situation or concept is fully described. This is especially important when writing for readers in a pre-technical or pre-commercial culture. In the South African context, much of senior secondary mathematics finds its application in the field of physics and commerce. However, many students have very little technical or commercial background. When visiting an African school, a group of std 10 HG students were asked if they could explain the concept of profit. It soon became clear that many of them were unacquainted with the concept as well as the formula for profit. Without this knowledge it was impossible to solve a specific examination question. The examiner had wrongly assumed that he and the examinees had these schemata in common. Cultural experiences provide valuable schemata, but they could differ from one cultural group to another (Kintsch & Greene, 1978).

One of the important implications of schema theory is that examiners should take care that questions are culture free. If students' schemata of their known world have not been expanded by their learning experiences, it is unfair to suppose these schemata to be present in the minds of students during examination conditions. This could privilege one group above another.

*2.3.2 Activated semantic contexts*

Although prior knowledge of subject matter makes it easier to comprehend text about that subject, prior knowledge does not automatically guarantee comprehension. A schema will only facilitate comprehension if it is consciously in a reader's mind. *In order for prior knowledge to aid comprehension, it must become an activated semantic context* (Bransford & Johnson, 1972:724). In this regard one could refer to the work of Paivio (1971) who distinguishes between verbal and visual conceptions of schemata. He views images and verbal processes as alternative coding systems or modes of symbolic representation. According to Paivio the more concrete a situation may be the more effective or functional a visual representation may be to arouse the necessary schemata in the mind of a reader (Paivio, 1971:9). Zimmler & Keenan (1983) have challenged this view by reporting on research where congenitally blind subjects' performance on tasks presumed to involve visual imagery in memory were remarkably similar to the sighted. The theoretical approach of Paivio should nevertheless be kept in mind when investigating the readability of mathematics text. Mathematics text communicating real-life problems includes abstract as well as concrete information.

A crucial question would be: What activates semantic contexts in the mind of the reader? What causes information in long term memory (LTM) to be activated in the working memory of the brain? A number of studies suggest that schemata are activated whenever they are perceived as important (Huckin, 1983:94). But, one could ask, how does the brain decide whether a certain schema is important or not? That which is important to one person may not be important to another. Here again familiarity with subject matter is of paramount importance. A specialist will read a text and immediately identify key words or phrases that are important for the comprehension of the situation or the solving of the problem.

In mathematics, numerous examples illustrate this phenomenon. For example, when addressing a certain geometry problem, the words *equidistant points, parallel lines, the centre of a circle* or *tangents* will immediately be perceived by specialists as important for the solving of the problem. These words activate rich schemata and evoke many inferences not mentioned in the text. These activated mental contexts are then consciously present in the working memory of the brain and available during the ongoing process of problem solving.

A reader who is less familiar with the subject matter must necessarily depend more heavily on the structure of the text to decide what information is important and what not. If important information is not placed in prominent positions, it will attract less attention and will subsequently not be identified as important. No activation of relevant schemata takes place. Comprehension is therefore poor. This leads to the next theoretical concept relating to the structure of a text.

### *2.3.3 The levels effect*

Research in cognitive psychology has shown that readers generally process information hierarchically, paying more attention to seemingly more important information than information experienced as less meaningful (Thorndyke, 1977). Main ideas placed in prominent positions were also found to be recalled better than propositions which were structurally placed in subordinate positions in the text (Kintsch & Keenan, 1973). This research links up with that of other researchers who have demonstrated that the structure of a whole passage, and not just that of single sentences, affects comprehension, reading time and how much readers are able to remember what they have read (Bormuth, 1967; Kieras, 1978 and Selzer, 1983). Other empirical research shows that most readers select the first idea in a passage as the main idea, unless that idea proves to be inconsistent with the rest of the passage (Kieras, 1980). When it comes to readability, the organisation of discourse variables cannot be ignored.

Although a specialist reader is not as dependent on the structure of discourse, well organized information is much more accessible - it increases the reading ease of text. Highly accessible information is of great importance to the non-specialist reader who does not always have the ability to distinguish high from low level detail. He is heavily dependent on well-organized discourse.

An analysis of the manner in which journal articles on physics and molecular biology are structured shows that writers believe that scientists read journal articles by searching for most newsworthy information - a behaviour similar to that of newspaper readers (Huckin, 1987). For this reason, scientific journals are gradually adopting text features that promote "surprise value", much like that found in news reports: titles have become more informative, section headings and subheadings are more frequent, abstracts are becoming standard and visual aids are impressively focused and attention getting. Drawing the attention more specifically to different levels of interrelated information is one way of helping to demystify academic discourse.

All this research has important implications for the structuring of written text. *To put it simply, the important points in a text should be placed in superior positions hierarchically: in headings, in sub-headings, in topic sentences at the beginning of paragraphs, etc.* (Huckin, 1983). Referring to mathematics discourse written for examination purposes this would mean that if the aim of a question is not to disentangle information, text should be structured in such a way that readers are able to recognize crucial detail with ease.

### *2.3.4 The leading-edge strategy*

The leading-edge strategy was developed by Kintsch and his co-workers (Kintsch & Keenan 1973; Kintsch & Van Dijk, 1978). It is a theoretical model of text comprehension and relates to an information processing view of cognition where the computer is used as a metaphor for man. Psychologists see the reading process as mainly controlled by two psychological constructs, long term memory (LTM) and short term memory (STM). The essential elements of comprehension, the schemata, are said to be stored in LTM. The capacity of LTM is thought to be limitless (one can always learn something more). Although LTM contains a vast amount of relatively permanent information, it has a low level of accessibility (Frederiksen, 1990). STM is seen as the working memory of the brain. Although accessibility to STM is high, it has a limited capacity. According to the work by George Miller (1956), STM can hold no more than about seven units of information. Therefore to keep information in STM it must be "re-entered", usually by explicit repetition.

According to the leading-edge strategy, a reader processes information top-down (cf. 2.4.1.1) while simultaneously attending to high level information from LTM and low level detail from STM . A text is coherent to the degree that these two parameters are well integrated. Kintsch & Van Dijk call this referential coherence or argument overlap. *One of the linguistic criteria for the semantic coherence of a text base is referential coherence...If a text base is found to be reverentially coherent, that is, if there is some argument overlap among all of its propositions, it is accepted for further processing* (Kintsch & Van Dijk, 1978:367).

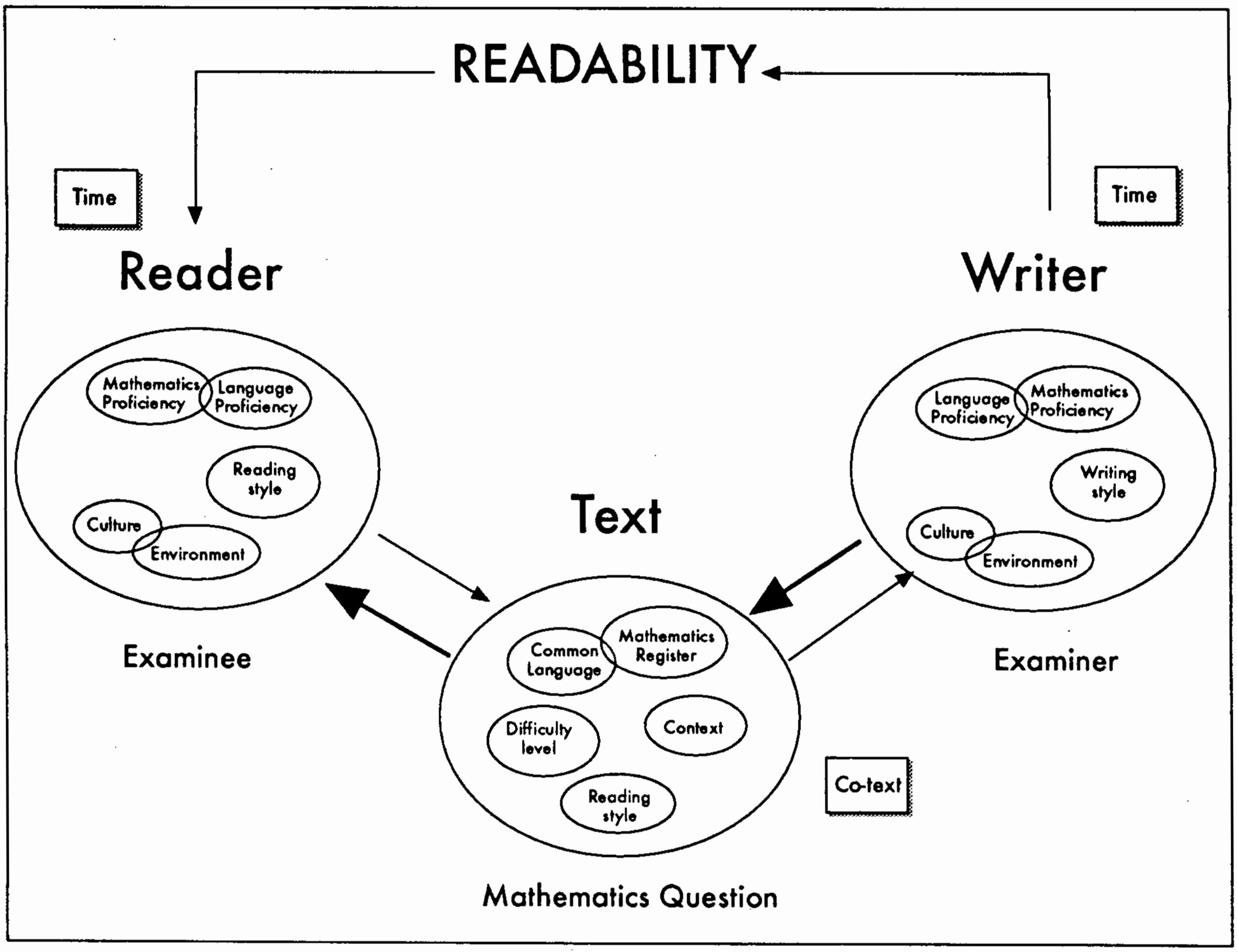
The leading-edge strategy is a call for writing coherent text. Even specialist readers have trouble in successfully bridging the inferential gaps caused by incoherent text. The writer must supply the "argument overlaps". Apart from activating high level schemata, as discussed in the levels effect, propositions must be joined by leading sentences or phrases. An explicit connection or argument overlap between schemata must be realized by references or details from lower levels of the text hierarchy. This is a leading-edge or most important strategy. Without referential coherence the propositions cannot be accepted for further processing. The low level details act as transitional phrases or leading sentences that evoke understanding. Without these "logical connectors" there is no link between abstract mathematical ideas and concepts (cf. 2.4.3.2).

When considering readability from a cognitive point of view, the importance of the reader audience becomes obvious. It reinforces the call for an increasing emphasis on audience analysis as suggested by contemporary psycholinguistic research.

## 2.4 Reading and psycholinguistic theory

Psycholinguistics is a science that studies how people learn and use language. It views readability somewhat differently to cognitive psychology. Although the readability framework of each discipline has its own fundamentals, there are some overlaps. This is not surprising because both are viewing the same concept. In psycholinguistics, *communication* between reader and writer plays an important role. Readability always implies an interaction between the reader, writer and text. When discussing readability from a psycholinguistic point of view, these three elements will form the basic frame of reference. The following sketch is an adaptation of the model presented in Figure 2.1.

***Figure 2.4*** **An interactive readability model for mathematics text**




From Figure 2.4 it is clear that in a mathematics examination, the text refers to the examination question whereas the reader and writer are respectively the examinee and examiner. Very much the same factors influence examiners and examinees - the first in their production of text and the second in the comprehension thereof. For example, *time* is an important factor- not only to the writer in order to produce readable text, but also to the reader if he is to comprehend the communicated message correctly. The *writing style* of the examiner as well as the nature of the examination question predict the *reading style* demanded by the text. Usually mathematics examination text demands a receptive or critical *reading style* from the examinee (cf. 2.4.1.2). Another contributing factor is considered to be the linguistic *co-text*. The term co-text is described as the formal syntactic structure of those sentences surrounding the immediate text (Strother, 1990:25-26), whereas *context* is the situation in which the text operates. If the syntactic structure of a previous question were to compare well with that of a given question it could aid the readability of that question. *Mathematics* and *language proficiency* show a certain degree of overlap because certain elements, like vocabulary, are common to both sets. Likewise there is an interaction between *environment* and *culture*.

### *2.4.1 The reader*

The reader is often considered as the "passive" partner in the communication process. Some debate does however exist about the definition of *passive* in this context. In a special way, reading is a very active skill and certainly the reading of mathematics text requires active attention. The perceived information must be processed for exact comprehension because exact comprehension is needed for an exact response.

The psycholinguistics of reading, studying how people read and what factors influence reading, is an area of common concern in psychology and linguistics, The assumption that a better understanding of *how* people read will lead to a better understanding of *why* some texts are more readable than others, seems to be a fair premise.

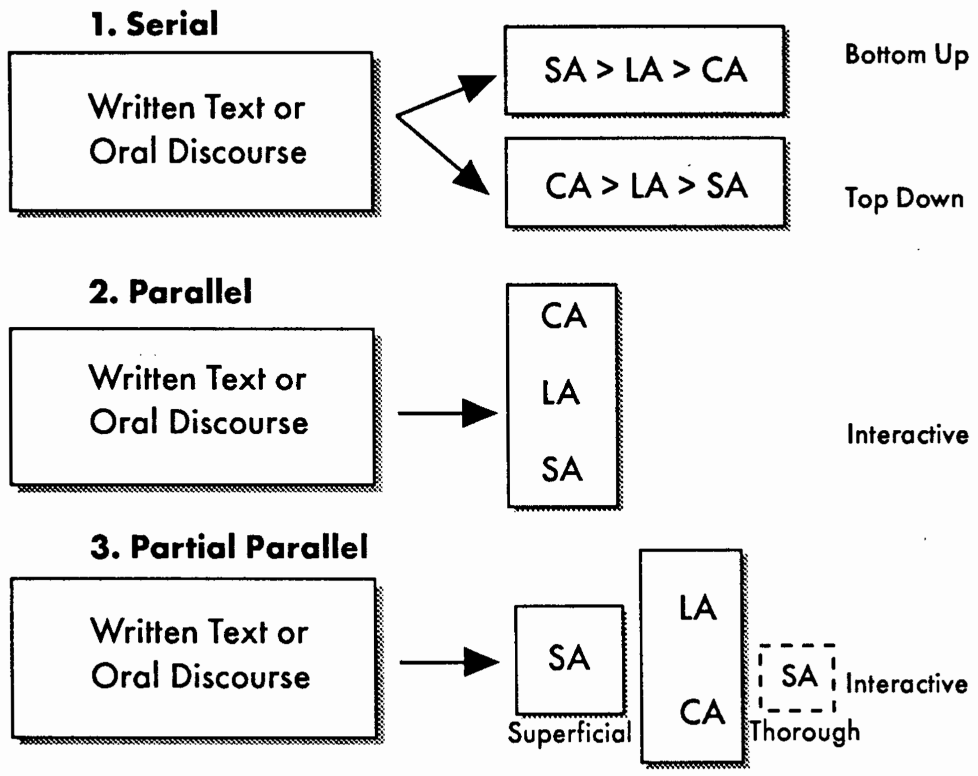
#### 2.4.1.1 Basic theories of reading

Two main theories attempt to describe this intricate human operation in the mind of the reader. The cognitive theorists consider the reading process to be controlled by a central mechanism in the brain called the monitor or central processing unit (Kempen & Hoenkamp, 1987). The monitor is deemed essential for the serial processing of information. This traditional approach uses the digital computer as metaphor to theorize about reading. The connectionist theorists suggest that the reading process is a connection of neural networks, interacting in a parallel way, without any controlling mechanism (Bechtel & Abrahamson, 1991). The parallel processes allow more processing to occur per unit of time. A more classic approach to reading is a combination of the cognitive and connectionist models. The following model by Ulijn and

Strother (1994) illustrates serial, parallel and partial parallel processing of information. The presence of a monitor is implied in every type of staging.

**Figure 2.5 Information processing model**

*(Model by Ulijn & Strother, 1995)*




When a person reads, he integrates single words into higher order linguistic units to form a hierarchical organization. By *bottom-up* staging is meant the serial decoding of words beginning with a syntactical analysis followed by the analysis of lexical and conceptual features. *Top-down* is the reverse process, starting with a global conceptual analysis. Psycholinguistics considers reading as a combination of serial, parallel and partial parallel processing. Partial parallel processing means that a reader starts reading by making a superficial syntactic analysis - often only a quick identification of the verb. This is followed by an interactive parallel analysis of lexical and overall conceptual features. Only in the case of comprehension difficulty does a reader return for a more thorough analysis of the syntactical, lexical and conceptual features.

A highly accessible text does not require extensive processing and can easily be retrieved. Low accessibility entails more processing before comprehension is complete. A high level of accessibility implies a high level of readability. The reader, writer and text all have their own particular set of characteristics. The interaction between these characteristics determines the flow of information or the accessibility of the text. The question is: What characteristics of the reader interact with the text to facilitate meaning?

2.4.1.2 Characteristics of the reader

Three important characteristics of a reader that influence readability are background knowledge, language and subject proficiency and reading style.

- *Background knowledge*

The source of a reader's background knowledge is his exposure to a variety of experiences. His *communal* environment gives him background knowledge of everyday experiences, like factories or cars. His *school* environment gives him background knowledge of academic experiences, like mathematical and commercial concepts. His *cultural* environment (cf. Figure 2.4) gives him background knowledge of historical experiences, like heroes or rituals.

A reader's background knowledge is the linguistic equivalent for schemata discussed in 2.3.1. Background knowledge can therefore be considered to be represented in a reader's long term memory by mental constructs or schemata. They guide the way a reader perceives and remembers things. The richer the experience, the richer the schemata. A teacher of mathematics will have much richer schemata on the subject of calculus than his students who are introduced to the concept for the first time.

The power of schemata resides mainly in the ability they give the reader to induce inferences from written information. Readers with rich schemata do not need so much explicit detail. A single word can call up a large amount of unwritten information. In this case too much information can even be experienced as unwanted or irritating. The schemata of background knowledge act as the given or known information that is needed to absorb new information.

According to the given-new contract of Haviland & Clark (1974; 1977), a writer agrees to present information in accordance with a certain format rule: "new" or unfamiliar information is linked to a reader's "given" or previously presented information. Communication between reader and writer is therefore based on shared schemata - on those available to the non-specialist partner of the communication process. If a student has had no academic experience of vectors, a question referring to velocity and displacement might cause many comprehension problems due to unshared schemata between examiner and examinee.

The presence of schemata in long term memory (LTM) does not guarantee comprehension. The schemata must be activated in LTM to form a semantic context in the working memory of the reader. This activation of schemata largely depends on the language and subject proficiency of the reader. Proficiency at these levels helps the brain to decide what words are important for comprehension and ultimately for problem solving.

- *Language and subject proficiency*

Language proficiency includes reading and lexical knowledge. Both have important effects on how a reader copes with text. There is an important relationship between a reader's vocabulary and his level of comprehension. Subject proficiency in mathematics refers to knowledge, insight, manipulation techniques as well as reading proficiency in the mathematics register. Elements of language and subject proficiency interact to identify key words or phrases necessary for comprehension. Consider the following question:

> *A ball is dropped from a height of 15 metres and constantly rebounds three fifths of its previous fall. Assume that this process continues indefinitely. Calculate the total distance that the ball covers.*

A reader with a high language and subject proficiency will recognize the words, *rebounds, constantly* and *indefinitely* as core concepts of a geometric sequence. This makes the text highly accessible and at the same time triggers mathematical thinking.

Language and subject proficiency also influence the effect of improved readability on comprehension (Klare, 1976 ; Tudor, 1988). For example, an increase in readability will have little effect on a reader's comprehension if he already knows a great deal about what the text has to say. Readers with a very low language and subject proficiency might also not benefit from improved readability. The key words remain inaccessible. So, improved readability can improve reading ease and reading rate without improving comprehension. Language and subject proficiency are crucial characteristics.

- *Reading style*

The way in which information is processed i.e. top-down, bottom-up or partially parallel, plays an important role in comprehension. Much of this is regulated by the reading style of the reader. Reading theorists have identified at least five different reading styles (Samuels, 1988).

SKIMMING - Reading for the *general* drift of a passage.

SCANNING - Reading quickly to find *specific* information.

SEARCH READING - Scanning with attention to the *meaning* of *specific* items.

RECEPTIVE READING - Reading for *thorough comprehension.*

CRITICAL READING - Reading for *evaluation.*

Of course, reading styles differ in the speed at which they are generally performed. A person's reading style depends on the particular purpose of his reading. Examination text for the assessment of realistic mathematics usually necessitates receptive and/or critical reading. These reading styles are slower and more analytical than the ones needed for the reading of ordinary expository prose. The slower reading rate aids the monitor in the mind of the reader in the processing of crucial data. Receptive and critical reading also supports reflection and interpretation of information, important goals of realistic mathematics education.

If a reader does not adapt his reading style to the reading style demanded by the text, neither comprehension or mathematical thinking can be facilitated efficiently. Consider the following question:

> *A manufacturer has the capacity to produce 8 000 calculators of a certain type per week. The cost of producing n calculators is $C = 100n + 200$ and n calculators can be sold per week at a price $P = 400 - 0,02n$ per calculator.*

Determine the value of n which maximizes the profit.

If the reading style is receptive, but the rate too quick, the words *per calculator* in the phrase *$P = 400 - 0,02n$ per calculator*, may not be seen. On the other hand, if the reading rate is too slow, the global message of the text can be lost. Local and global understanding are necessary to understand the exact relationships between variables. Without this, mathematization is not possible. Mathematization is mainly the translation of verbally described *relationships* into non-verbal mathematical language, a type of recasting of *related* information.

Experimental research suggests that the variation in reading rate can best be explained in terms of two distinct types of processing, macrostructural (meaning) and microstructural (form). Skimming is fast because it uses fast macrostructural processing and hardly any microstructural processing. Receptive reading is slower because it requires a great deal of both macro- and microstructural processing (Huckin, 1983:100). First and second language readers of mathematics text often experience comprehension problems because of incorrect reading styles, the one because of a too fast and the other because of a too slow reading style.

*2.4.2 The writer*

Whereas the reader is often considered as the "passive" partner in the communication process, the writer is the "active" creator and decisive determinant of text. Readability of text depends to a large extent on the writer's choices. The characteristics of the writer that have a definite influence on the communication process are very much the same as those of the reader. Reading processes depend on written text. However, examination situations differ from most other written communications - the reader and writer are not equal. This inequality holds true for most of the reader/writer characteristics influencing readability.

Due to more experience in practically all walks of life, the background knowledge of the writer, here the examiner, is far more extensive than that of the reader or examinee. His schemata are therefore much richer. This gives him more inferencing power. He needs much less detail to understand written information. To the examiner, all the information in the text is given or known and he could very easily violate the given-new contract if he does not keep the reading audience in mind.

Mathematics examiners are mainly selected on the basis of their educative abilities. Their subject proficiency has been enriched by tertiary education and teaching experience. Furthermore they can be considered to be native speakers of the language used in the examination text. This gives them the ability to distinguish between relevant and irrelevant information. The correct semantic contexts are activated practically automatically in their minds. Less difficult words, well structured discourse or enough time to assemble cognitive structures on which to operate are not crucial to examiners when they read text. They are also far more able to adapt their reading style to the demands of the specific passage than their less proficient exmaminees. Because they themselves experience little difficulty with mathematics text, examiners may not be so aware of the reading needs of their examinees. Unless they make an analysis of their reading audience, they could very easily produce text with equal partners in mind. The language of the examination paper is the vehicle for carrying the mathematical information. Examiners can become so preoccupied with the mathematics that they forget the role of language. If the vehicle is faulty, the information may never reach its destination.

Writing should always be seen as both product and process, (Flower et al. 1983). *Planning, drafting, reading, and revising* are all part of the writing process. Much care must be taken to consciously approach the writing process recursively. A great deal of reflection and consideration are needed. Only then will the product reflect a writing style that is more compatible with the reading needs of the reader.

*2.4.3 The mathematics text*

The std 10 HG examination paper contains large portions of common language together with portions of the mathematics register (cf. 2.4.3.1). Furthermore, mathematical knowledge is often applied in the field of physics, commerce or other sciences. Consequently, scientific language, which is typical of these disciplines, may also find its way into mathematics text. For example, words like *velocity* or *displacement* could be used in problems relating to a physics context. All this adds to the complexity of mathematics text (cf. 1.1).

One of the important reasons for teaching mathematics at school level is that it is a powerful tool for communication (Cockroft, 1982). A mathematical formula is a good example of a collection or "chunk" of interrelated data carrying much information. Furthermore, the universal nature of the mathematics language makes it relatively easy to communicate mathematical detail - even across language borders. However, this strength of the mathematics language can become a stumbling block to many because of the range of comprehension difficulties that arise when operating within this language field. Pimm argues that he interprets "mathematics *is* a language", as a metaphor (Pimm, 1987:20). Mathematics is *like* a language - to speak mathematics fluently one must successfully give meaning to it when one *writes* it, *reads* it, *speaks* it or *hears* it. One must have sufficient knowledge of the pronunciation, spelling, syntax, vocabulary and structure.

To understand what makes "the language of mathematics" so unique, a closer look is needed. This should illustrate why readability formulae, intended for expository prose, are unsuitable for mathematics text. It should also support the call for the removal of any readability factor that *unnecessarily* prevents a clear understanding of the written information. The mathematics register has enough difficulties of its own - difficulties students should be taught to master.

#### 2.4.3.1 The concept of a mathematics register

Any language can be considered to include a universe of language skills as certain areas of language are used for specific purposes. *Natural language, the language used in everyday communication, is one of the components or subsets of this universe. The language used to discuss computer technology or that used for scientific topics are two other subsets. Linguistically, these subsets of language are referred to as* **registers** (Dale & Cuevas , 1987:12). Strother explains a register by saying that *a register incorporates a subset of the total language, depending on the activity or context in which the language is functioning* (Strother, 1990).

The mathematics register can be divided into three modes: the verbal, non-verbal and iconic mode. Students experience comprehension difficulties in each of these modes. For the purpose of this study, more attention will be given to problems in the verbal mode. The ongoing call for more realistic mathematics education has automatically caused an increase in verbal problems and subsequently a corresponding increase in the use of this mode.

#### 2.4.3.2 Readability problems in the verbal mode of the mathematics register

Comprehension difficulties in this mode refer to elements such as vocabulary and syntax. The given examples are by no means exhaustive. They are, however, sufficient to illustrate the many problems experienced by students as they to try to negotiate meaning. It is therefore not surprising that when *unnecessary* readability problems are added to the already existing difficulties, a complete communication breakdown can result.

***Vocabulary***

Mathematics vocabulary includes many words that are specific to the mathematics register. Words like *coefficient, quotient, logarithm, index, quadratic* are new to most students and have no analogy in the common language of the playground. To link an unknown word to an unfamiliar, abstract concept is one of the difficulties when teaching and learning mathematics.

Another, perhaps more confusing problem, is the use of words that have other meanings in ordinary English. Words like *root, function, curve,* and *real* are only a few examples of this phenomenon. A study by Kouba (1989) illustrates this problem very effectively by pointing to the great confusion caused by words having different meanings within at least three registers: ordinary English, mathematics and science. The following are examples of but a few: *line, curve, inverse, expression, variable and constant* (Kouba, 1989:606).

Apart from isolated words, the mathematics register *combines* words to form phrases with a completely new meaning, like: *the inverse of a function; lowest common multiple; additive inverse; negative exponent; absolute value; linear programming.*

When giving meaning to words in mathematics, scholars must also take into account the context in which they are used. Students cannot just learn lists of words. The same word can have different meanings in different contexts. Non-native readers of English in particular experience difficulties in this field. Consider the following two meanings of the word *factor*:

- If ( x - 7 ) is a *factor* of f(x), find the other *factor*.

- What *factor* in the following linear programming problem can be ignored?

Students also need to keep in mind that different words can be used to signal the same operation - they have the same meaning. *A minimum of 60 000* or *at least* 60 000 both signal $\geq$ 60 000. *At the most* or *not more than* both signal $\leq$. Words like *less than, decreased by, minus, less, differ,* all signal subtraction.

Some words in the mathematics register have a broader meaning than the meaning experienced by many students in everyday life. A fraction of the cost always means an amount less than the cost, but in mathematics a fraction is not *always* between one and zero. In real life, if your problems multiply, they definitely become more. Multiplication in mathematics does not *always* make bigger or more. When doing verbal problems in particular, the exact meaning of certain words can be crucial for solving the problem. Certain words signal definite meaning. The correct solution often hinges on the ability to identify key words. *Speed* is not *velocity* and *distance from a point* is very often not the same as *the distance travelled.* Pimm is convinced that *meaning* rather than lack of rigour is the central problem facing mathematics education (Pimm, 1987:7).

### *Syntax*

Syntax refers to the manner in which words are organized in sentences to create meaningful language. The mathematics register has its own unique way of structuring words for meaning. Dale & Cuevas (1987) point to four syntactical factors that cause comprehension difficulties to readers. These will be discussed briefly.

#### *1. Parentheses*

A parenthesis is an additional word, phrase or sentence inserted into a passage which would be complete without it. Usually a parenthesis is separated from the rest of the text by commas, brackets or dashes. Parentheses abound in mathematics text. They are often inserted to keep the given information valid and not always *actively used* in the understanding or solving of the problem. Just this causes a readability or comprehension problem. Consider the following question:

> *If m and n are real numbers, determine, with reasons, the minimum value of*
> $7m^2 + (n - 3)^2 + 5$ (NED, 1991)

The parenthesis, *if m and n are real numbers* has been inserted to ensure that the answer students are able to generate at std 10 HG level is mathematically correct. However, very few students understand the insertion of this clause because the std 10 HG syllabus only operates within the set of real numbers. Instead of aiding the readability of the information, it seems that parentheses often make mathematics text less accessible.

Readability problems are increased when the parenthesis is not clearly separated from the rest of the text. Even a separation by commas causes comprehension difficulty. Non-specialist readers of mathematics in particular tend to read parentheses as indispensable information that must be actively used to solve the problem. Students are used to mathematics text being concise, very seldom including redundant information. Information not for active use causes doubt and subsequently comprehension difficulties. Even if the parenthesis *is* necessary for the *active* part of the problem solving process, it upsets the reading rhythm and therefore necessitates a higher level of processing. This makes the text less accessible. It causes uncertainty, especially to non-specialist readers.

2. *Comparative structures*

Mathematics is a study of relationships, therefore comparative structures are an essential part of the mathematics register. They are often difficult for students to master. The following example illustrates this difficulty.

> *The number of passengers, x, transported by buses must be a minimum of 400 per day but must not exceed three times the number of passengers, y, transported by mini-buses.*
>
> (TED, 1992).

Examples like these abound in the linear programming part of std 10 HG examination papers.

3. *One-to-one correspondence*

Dale & Cuevas (1987:15) indicate how the lack of a one-to-one correspondence between the mathematical symbols and the words they represent can cause students considerable comprehension difficulties. Students tend to write mathematical sentences in the same order as they have read the verbal form. Consider the following two examples:

- "*A number **a** is five less than a number **b***" is often written as: **a = 5 - b**

- "*P is twice as far from Q as from R. Determine the locus of P*" is translated by many as, **2PQ = PR.**

4 *Logical connectors*

These are words or phrases used in mathematics text to help develop and link ideas. Logical connectors like *because; if...then; if and only if,* help to write cohesive text. They supply the argument overlaps between statements as indicated by the leading strategy in the cognitive approach to readability (2.3.4). Students must recognize these words and know what relationships are being signalled. Very often these connectors are used in complex statements. Consider the following example from the std 10 HG syllabus:

$$y = \log_a x \quad \textbf{\textit{if and only if}}$$
$$x = a^y\ ;\ a > 0\ ;\ a \neq 1;\ x \varepsilon R;\ y \varepsilon R$$

Many E1 students have comprehension problems with the above. Clearly. the difficulties experienced by E2 and E3 students will be even greater.

2.4.3.3 Common language difficulties: vocabulary and syntactical issues

The information of contexts used in realistic problems must be communicated by natural or common language. Readability problems in this field cut across vocabulary, syntax and discourse.

*Difficult vocabulary*

Difficult vocabulary is any word or phrase which is not immediately accessible to the student. Words like *yield* or *momentarily* are apt to cause readability problems (cf. CED examination paper, 1990). They are not words commonly used by most students.

*Syntax*

Syntactic structures that seem to be most commonly used in the scientific/technical or ST-register are passives, nominalizations (verbs transformed to nouns), infinitives and participles

(Strother, 1990:7). Since mathematics is often applied in a scientific or technical context, the language of verbal problems in mathematics will likely contain these same syntactic structures.

The passive voice is frequently used when it is necessary to emphasize the action more than the actor or in the writer's attempt to remain impersonal. There are researchers who claim that the active voice is comprehended and recalled more easily than the passive voice. Others believe the problem is not so simple, especially if the effects of actives and passives are evaluated in rhetorical contexts (Selzer, 1983:80). Strother found no significant effect of the passive voice on comprehension in the ST-register (Strother, 1990). It seems as if no consensus has yet been reached about the readability of passives and actives in functional documents like mathematics text. Readability formulae do not address this issue.

Another syntactic structure that seems to be problematic to students, especially to second language readers, is the hypothetical construction in English (Au, 1983). In many traditional African languages, this type of sentence structure does not exist. The ability to cope with hypothetical constructions is a skill usually acquired at a later stage of language acquisition in both first and second language students (Kirschner et al. 1992). Yet hypothetical constructions appear quite frequently in mathematics examination text. Consider the following verbal problem:

> *There are exactly enough exercise books in a class to supply each pupil with the same amount of books. If there had been five more students each would have received 2 books less. If however there had been 3 students less each would have received 1 book more and there would have been 11 books left over. How many students are in the class and how many books are available?*
>
> (JMB, 1992)

Apart from the incorrect use of the words "amount" and "less", the mathematical difficulty of the above question is increased by the hypothetical construction of the sentences. Not only the hypothetical construction but also the hypothetical situation adds to the difficulty. Research indicates that problems are experienced as more difficult when using a hypothetical situation

(Dale & Cuevas, 1987). A question like: *If 9 apples cost R4,50, how much will 2 apples cost?* is experienced as more difficult than, *9 apples cost R4,50. How much do 2 apples cost?*

The hypothetical situation can result in a sentence construction where the hypothetical *condition* and the relevant *question* are entangled in the same sentence. This often happens in mathematics, making the text even less accessible. Consider the following example from an examination paper:

> *If* $3\,{}^{5}/_{9}$ *and* $40\,{}^{1}/_{2}$ *are the first and last terms of a geometric sequence, determine the fourth term(s) if there are seven terms in the sequence.*
>
> ( CED, 1991)

2.4.3.4 Common language difficulties: textual features

Many studies on readability have limited their research to factors concerning words and sentences. More recently however, psycholinguists, cognitive psychologists and discourse analysts have directed their interest towards factors that affect the readability of connected discourse. The overall organisation of text is believed to be the major determinant of readability. This is true especially for adult readers (Coleman, 1962).

Factors beyond words and sentences seem to have a still greater impact on the readability of so-called "functional" documents (Faigley & Witte, 1983). Functional documents, like contracts, manuals and regulations, have goals beyond those of ordinary expository prose. People read them in order to act. A mathematics examination paper is a typical functional document written in a special type of technical language for a selected audience and specific purpose (cf. Figure 2.3). Comprehension of the text must trigger mathematical thinking and lead to the execution of a variety of actions or mathematical processes.

In this research the following discourse features were found to be important for the readability of ordinary English. This knowledge seems to have important consequences for the writing of mathematics text.

1. *Elucidation of text*

Examination papers consist of text and metatext - information and questions set on the information. Although the questions form part of the text, they seem to be less cohesive in relation to one another than regular continuous discourse. Cohesion of text has been identified as an important readability factor (Kintsch & Keenan, 1973). To work towards more cohesive text, questions should elucidate the text (Kirschner et al. 1992). In this way students are helped to work through the hierarchy of ideas presented in the information. For realistic mathematics, this implies that initial questions on given text should be concrete, real-life questions - compatible with the real-life experiences of the students. Questions could bring the context to life, especially to readers who have little experience of the context in their own cultural setting. Elucidation of text could help students to understand what variables influence the situation *before* they are expected to start with the process of abstraction or mathematization.

In this regard examiners should keep in mind that students ask their own questions when trying to elucidate mathematics text. This leads to comprehension problems if mathematical functions are used in an artificial way. Consider the following extract from a mathematics question:

> *A manufacturer of hi-fi sets determines that in order to sell x units of a new hi-fi, its price per unit must be* $p = (1000 - x)$ *Rand.*
>
> (CED, 1990)

This implies that if the selling price is 1000 rands, no units will be sold. To sell 1 unit, the price must be 999 rands. To sell 2 units, the price must be 998 rands and so on. This is not true to real life. Attempts to elucidate the text lead to meaningless answers and subsequent constraints in the further processing of information. Pimm warns that students can become conditioned to these senseless messages conveyed by mathematical sentences. He illustrates his warning by giving an example. A youngster was asked to translate $6 + 2 = 8$ into a real-life situation. His answer was: "*My mother has six irons. I gave her two. Now she has eight*" ( Pimm, 1987:12).

2. *Multiple-choice questions*

When testing reading comprehension, various testing methods have shown to have a significant effect on achievement levels (Shohamy, 1984). A testing method that is systematically becoming part of the std 10 HG examination is multiple-choice questions (National Examinations, 1990-1993; DET, 1992-1993). An important question is whether any factors in this type of test format affect comprehension.

One of the main problems seems to be caused by the sentence-completion format (Statman, 1988). Students were found to perform better if the stem of a multiple-choice question was an interrogative, rather than an incomplete sentence. The reason seems to be a processing problem. Students attempt to reformulate the question as an interrogative one (Kirschner et al. 1992). The double cognitive operation involves more processing, which indicates low accessibility. Time-wise this influences the whole examination. According to Just & Carpenter (1980) more processing necessitates more reading time which means less time remains for the rest of the examination. It has already been indicated in 1.8.1 how a lack of time could have detrimental effects on readability.

3. *Topic sentences*

The structure of individual sentences cannot be explained without referring to their textual setting. The described context as well as the co-text (cf. Figure 2.4 ) have an influence on the structure and position of sentences in text. Writers move from global planning to local detail. Apart from syntactical structures, text has an important topical structure - the way the topic is developed. Realistic mathematics is assessed in the context of real-life situations. The context and topic of the problem allow the students to grasp the intention of the questions more effectively (Van den Heuvel-Panhuizen, 1993). Research by Faigley & Witte (1983) shows that readers comprehend and recall passages better if the topical subject corresponds to the grammatical subject of a sentence. A topic sentence can therefore be considered as a sentence where the topic or theme of the text is carried by the subject position of the sentence.

Similar research by Kieras (1978) found that readers were much more likely to identify the main idea or topic of a passage when it appeared in the initial sentence of the paragraph. Readers tend to accept the initial information as the most important part of a passage. To a reader, this information signals the thematic content or purpose of a passage. The purpose of a passage guides the rest of the comprehension process (Kieras, 1980).

As suggested before (cf. 2.3.3), an important way of activating high level schemata would be by referring to them in topic sentences at the beginning of paragraphs. Text for the assessment of realistic mathematics could reserve first sentences for the topic of the verbal problem.

Similarly, sentence subjects could be utilized effectively when considering the limitations on short term memory. It may be recalled that short term memory is considered to be able to store only about seven items at a time (cf. Miller, 1956). Activated schemata must therefore constantly be reinstated in the mind of the reader. An effective way of reactivating these mental constructs would be to use the subject position of a sentence for reinstating important information.

4. *Proposition density*

Propositions are units of information. To illustrate the concept of propositions, consider the following sentence:

*A great rocket stood in the desert.*

The sentence contains three propositions:

- *The rocket was* ***great****.*
- *It* ***stood*** *somewhere.*
- *That somewhere was* ***in the desert****.*

Valuable research linking proposition density to readability was done by Kintsch and his colleagues (Kintsch & Keenan, 1973; Kintsch et al. 1975). They believe readers process propositions, not sentences. According to empirical evidence reading time increases when the number of propositions increases per sentence or per paragraph even though the length of the sentence or paragraph may remain the same. An increase in proposition density therefore increases the level of information processing (cf. Just and Carpenter, 1980). Proposition density seems to be an important indication of reading ease or readability.

Examination questions often reveal a high proposition density per sentence and paragraph. This implies :

- A high level of information processing which makes the text less accessible or less readable.

- A possible overload of the limited space in short term memory. In mathematics the cognitive space in STM is not just used for comprehension, but also for the rest of the problem solving process (Frederiksen, 1990).

- Loss of examination time due to the extra time required for reading. This is crucial for the readability of the rest of an examination paper.

*5. The given-new contract*

Communication is considered to be a co-operative enterprise. According to Grice (1967) the most important convention for successful communication is what he calls the "co-operative principle". This means speakers (or writers) should follow a simple rule: "Be co-operative." Grice presents four maxims for co-operative communication:

*Quantity*: Make your contribution no more or no less than is required.

*Quality:* Say only that which you and your audience have adequate evidence of, that is, try to make your contribution one that is true.

*Relation:* Be relevant.

*Manner:* Be clear. Make your contribution easy to understand. Avoid ambiguity, obscurity, and prolixity.

Haviland & Clark (1977) have used the work of Grice to describe what they call the given-new contract, a contract that has consequences for the process of comprehension. According to the given-new contract, writers (or speakers) agree to present new information (usually at the end of a sentence) by linking it to given or previously presented information (usually at the beginning of a sentence). If new information is not effectively tied to what is known or if too much new information comes at once, comprehension is impaired. The "given" information is related to the available schemata or background knowledge of a reader and therefore the given-new contract can be considered as the linguistic expression of cognitive schema theory (cf. 2.4.1.2).

For effective comprehension, mathematics text should reveal the characteristics of this contract. It could also be applied to the logical order in which the metatext (the question) is linked to the text (the given information) of a mathematics problem. If the question is asked before the information is given, there is a delay in the communication process. One cannot answer a question if one doesn't know what "given" information the question is referring to. Consider the following two examples from ordinary discourse:

- *Will you have tea with me on Saturday? I am having a birthday.*

- *I am having a birthday. Will you have tea with me on Saturday?*

In the first example there is a delay. Only after the fact of the birthday is known, does one go back to the question to consider it in the given context. In the second example the information prepares the reader for the question that is to follow. This makes the second example more accessible than the first.

The same is true for mathematics text. Consider the following question:

> *Calculate the value of k if three consecutive terms of a geometric sequence are $2k - 5$; $k - 4$ and $10 - 3k$*
>
> (NED, 1990)

In this piece of examination text, the question is asked before any information is given. If the given-new contract were to be applied, a more readable text could be:

*A geometric sequence is given. The first three consecutive terms are:*

*2k - 5; k - 4 and 10 - 3k.*

*Calculate the value of k.*

The cohesive and logical development of a theme, as suggested by the given-new contract, is important for the reader to construct a contingent knowledge structure of the communicated data (cf. 1.5; Bobrow & Brown, 1975). For highly accessible text, the order of data input should follow the same order as that which is necessary for the synthesis of a contingent knowledge structure (CKS). Only after the CKS has been constructed, can the reader begin with an analysis in an attempt to answer the questions. In a *verbal* mathematics problem, the formation of a visual image of the given data corresponds to the construction of a CKS. An examinee needs to first form a visual image of the problem situation before he can start operating on it - first synthesis, then analysis. Therefore an examiner should first present the information and then set the question.

*6. Reading style and textual features*

As mentioned before, mathematics text describing real-life problems usually demands a receptive or critical reading style (cf. 2.4.1.2). However, the text format is seldom adapted to activate these specific styles. Huckin (1983:101) urges writers to anticipate what reading style the reader is likely to need. If only skimming is needed, the writer can concentrate on italics, bold letters or headings. But if the reader needs to read for detailed comprehension, techniques must be followed that aid the reader in the step-by-step processing of information. The purpose of the text and the question difficulty will to a great extent determine the necessitated reading style.

## 2.5 Preliminary readability guidelines

The research on readability presented in this chapter is by no means complete. Nevertheless, the reported literature suggests a variety of guidelines that could be considered when writing mathematics for examination text. The underlying principle of these guidelines is reader analysis. Great differences exist between examiners and examinees. This has important consequences for writing. If the examiner writes according to the reading needs of his equals, the text will probably not communicate the intended message to the intended audience (i.e. the examinees) successfully.

Not all factors that inhibit comprehension can be removed by improved readability. The examiner does, however, have the power to manipulate the readability of text. The following suggestions could guide the writing process and refer to vocabulary as well as syntactical and textual issues.

### *2.5.1 Vocabulary*

- Use familiar, common language words.

- Continue to use the same word when referring to the same concept.

- Take care not to presuppose the familiarity of words belonging to other science registers.

### *2.5.2 Syntax*

- Avoid unnecessary parentheses. If they *are* needed, place them in subordinate positions - for instance, in a separate sentence at the end of a paragraph. Separate them clearly from the information needed for the actual problem solving process.

- Avoid hypothetical situations as well as hypothetical constructions.

- If possible, use interrogatives instead of sentence-completion questions in multiple-choice questions.

*2.5.3 Global guidelines - textual issues*

- Take care not to presuppose background knowledge of subject matter. Obey the principles of the given-new contract by linking new information to that which is known.

- Curb the proposition density of sentences and paragraphs.

- Take care that the order of data-input supports the order needed to construct a contingent knowledge structure (CKS).

- Design the hierarchy of the metatext in such a way that the questions elucidate the text. For example, start with concrete, real-life questions to explain the given information in more detail.

- Place text before metatext - first the information, then the question. Use a separate sentence for the question. Use one sentence per question.

- Take care that the mathematical model reflects the relevant real-life situation.

- Anticipate the reading style required of the reader and format the text accordingly. Some of the following guidelines describe a few of these formatting techniques.

- Make the purpose of the question clear. Use topic sentences or headings at the beginning of the discourse.

- Place important information in superior positions. One such a position is the subject position of a sentence.

- Depending on the length of the discourse, important information should be restated.

- Allow for enough reading time.

The above guidelines, however effective they may appear, should be seen as preliminary. There are quite a number of research questions that still have to be answered before one can settle for more definite recommendations.

## 2.6 Possible research questions

From the research reported in this chapter, the following issues emerge:

1. The preliminary readability guidelines were developed mainly from research done on ordinary, common language text. Will linguistic factors that influence the readability of common language prose have similar effects on mathematics text, especially text written for examination purposes? Every examination question consists of a relatively short piece of discourse. Will the non-verbal data, or other verbal cues, not have a strong enough influence to remove any serious comprehension difficulty that could bar the way to mathematical thinking?
2. Reported research has indicated that improved readability does not always improve comprehension. If this is the case, will improved readability have any substantial influence on achievement?
3. If readability *were* to affect achievement - who would gain the most by improved readability, first or second language readers?
4. What effect will improved readability have on achievement levels of second language readers whose school and home cultures do not correspond?
5. What special reading needs do students have whose home and school cultures are not related and what implications does this have for the readability of examination text?

The last two questions are of special interest in the South African context. Therefore linguistic and cultural issues influencing reading in a second language will be considered in the following chapter.

# CHAPTER THREE
# LINGUISTIC AND CULTURAL ASPECTS OF READING IN A SECOND LANGUAGE

## Rationale

*Second language reading is not only a matter of language acquisition but also a matter of learning another culture. There seems to be no culture without a language and no language without a culture.*

*In a certain sense the reading process in a second language is the same as in a first language. On the other hand, even if the underlying reading strategies are the same, important additional factors influence the process when reading is done in a second language. Factors like culture, second language proficiency and the possibility of mother tongue interference all come into play. An important issue is how these additional linguistic and cultural factors interact during the reading process and what influence this interaction has on the readability of text.*

*Second language readers and bilingual education are common in the South African context. Knowledge of linguistic and cultural effects on readability could enable examiners to address areas of possible concern when writing examination text. If one were to apply the advice of Ausubel (1963) to writing, one could say, "Ascertain the reading needs of your audience and write accordingly." This chapter will guide the audience analysis by looking at linguistic and cultural aspects of reading in a second language and the possible influences these factors could have on readability. The theoretical basis developed in the previous chapter will be expanded to direct the empirical studies reported in the following chapters.*

## Chapter contents

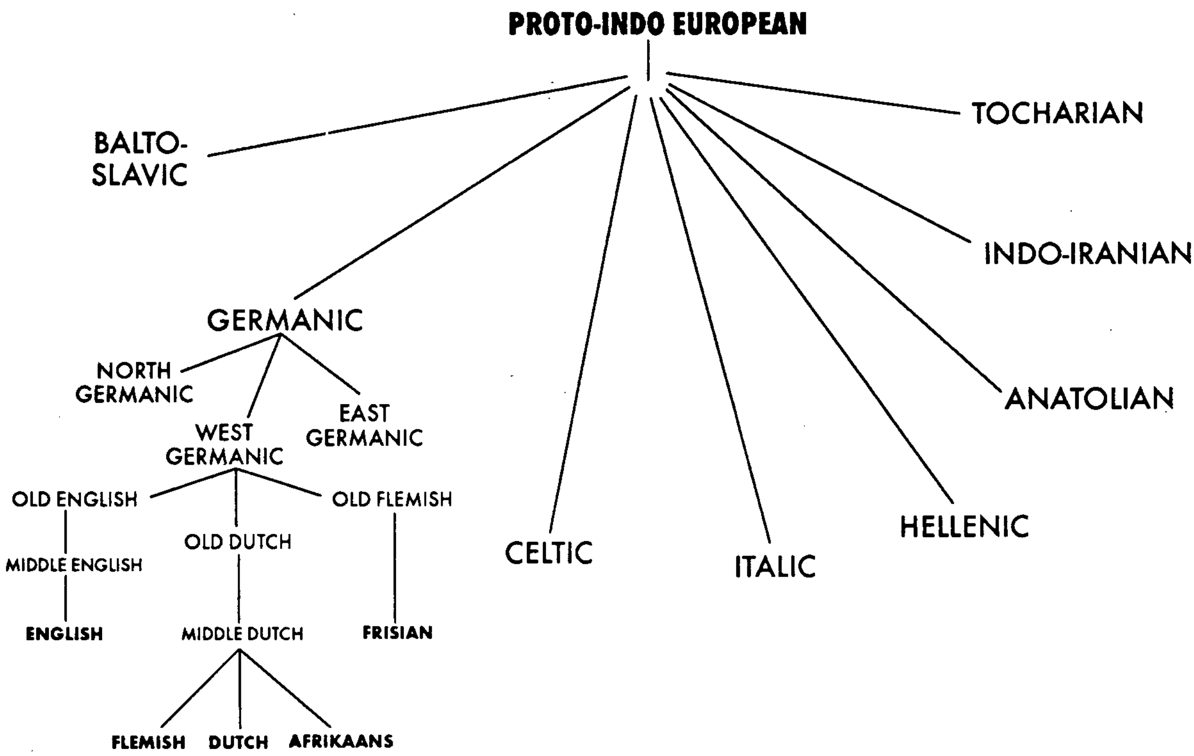
### 3.1 Bilingual education - the South African situation

By bilingual education is meant the presence of more than one language in the learning environment. In South Africa, bilingual education prevails in most schools, which makes the school population a complex reading audience. As indicated in 1.7, most black students receive their secondary school education in English. They are taught in English, use English textbooks and write their examinations in English. To most students however, English is a second language and to some even a third or fourth language as quite a number of black students speak two or more of the eleven official languages better than English. The eleven official languages are: Sesotho, saLebowa, Sesotho, Setswana, isiNdebele, isiXhosa, isiZulu, Tshivenda, siSwati, Xistonga, Afrikaans and English. (For the purpose of this study a second language will refer to a language other than the mother tongue.)

The languages of black students in South Africa can be divided into four main groups i.e. the Sotho-Tswana, the Nguni-Ndebele, the Venda and the Tsonga groups. All languages belonging to these groups are indigenous to Africa and developed from a prehistoric language known as Urbantu. These languages are in no way related to the two languages spoken by most white students in South Africa i.e. Afrikaans and English. The following sketch is a schematic representation of the Indo-European family of languages. According to this representation both English and Afrikaans have Germanic roots and are descended from a prehistoric language, Proto-Indo-European.

**Figure 3.1 A schematic representation of the Indo-European family of languages**

*(Adapted model from the American Heritage Dictionary of the English Language, 1981)*




It is clear that the majority of students in South Africa read their std 10 examination paper in a language which is in no way linked to their mother tongue. Although these students receive their education in English from their fifth school year, many of them are not fluent speakers of English. Most of them do not live in an English environment and most of their fellow students are also Africans. The *natural* medium of communication at most schools is one of the African languages. To many, therefore, conditions to acquire a high level of proficiency in English hardly exist.

Much research has been done on the needs and effectiveness of bilingual educational programmes (Cummins, 1980; 1984; Hamers & Blanc, 1989). Cummins points out that substantial differences exist between the language needed for social interaction at school and the academic language required to communicate within different academic fields. He refers to the first as basic interpersonal communication skills (BICS) and to the second as cognitive academic language proficiency (CALP). Educators are unanimous in their opinion that BICS is a necessary, but not a sufficient condition for progress at school level. To be able to progress efficiently, CALP is needed (Terlouw, 1991:115).

Research indicates that although face-to-face communicative skills are largely mastered by immigrant students within two years of arrival in the host country, it takes between five and seven years, under optimal conditions, for students to reach CALP parity with first language learners (Cummins & Swain, 1986:156). How does the lack of second language academic skills influence bilingual education in South Africa? Southey (1992) is convinced that the language problem in black education is not *a* problem, it is *the* problem.

According to statistics published in 1993, the achievement level of African students in mathematics is not very promising. Less than 30% of the African std 10 population present mathematics as a subject and less than 30% of these students eventually pass (DET-report, 1993). If one considers the cumulative way in which mathematical knowledge develops, one realizes the major role language could play in this regard.

To obtain a better idea of the problems experienced by second language readers, the issue of bilingualism and the effect of linguistic factors on the readability of text read by second language students will be explored in more detail. In this regard *inter*-linguistic factors refer to linguistic factors that cause comprehension difficulties in the mother tongue as well as in a second language, but because of a weaker second language proficiency they cause more problems in a second language. An example of an inter-linguistic factor would be unfamiliar vocabulary. *Cross*-linguistic factors on the other hand refer to linguistic factors influencing communication *across* languages. For example, the subjunctive mood in English could be experienced as a difficult sentence structure by a reader whose mother tongue does not use this type of sentence structure.

## 3.2 Bilingualism and the effect of linguistic factors on readability

The concept of bilingualism is a difficult one. Hamers & Blanc (1989) distinguish between the concept of bilingualism and bilinguality. According to them *bilinguality is the psychological* *state of an individual who has access to more than one linguistic code as a means of social communication.* (Hamers & Blanc, 1989:6). The degree of access will vary along a number of dimensions which have, amongst others, a psychological, cognitive, psycholinguistic and sociocultural basis. They regard bilingualism as the state of a linguistic community in which two languages are present. Bilingualism, to them, includes bilinguality (or individual bilingualism).

For the purpose of this study the concept, bilingualism, will be used and will refer to the ability of a person to speak and read two languages, the mother tongue and a second language. The mastering of the second language, compared to the first language (or mother tongue), could reveal different levels of competency.

### *3.2.1 Lexicon*

Although concepts are independent of language, language does play an important role in the formation of concepts in the mind of the reader. The intellectual growth of a child is contingent on his mastering of language (Vygotsky, 1962:51). In the process of concept formation (and concept retrieval), the lexical component is a crucial element. There is an inseparable relationship between a reader's conceptual system (memory store or long term memory) and his lexicon. If the meaning of a word is misunderstood, considerable comprehension problems may occur.

In mathematics, the bilingual person must not only learn the special terminology and syntax of the mathematics register but also has difficulty to understand the conversational language of the text. Imagine the following question being read by a non-native reader of English:

> *It is estimated that t years from now the circulation of a magazine will be given by the function,*
>
> $m(t) = 50t^2 - 1200t + 3000$
>
> *At the time of first publication, what was the circulation of the magazine?*

If a reader does not know what *circulation* or *magazine* means, he has no idea what information the function is communicating. The questions based on this function will make little sense and the required mathematical operations cannot be executed. In a non-English environment the word *magazine* might be completely unknown. A translation into Afrikaans might lead to a word like *magasyn* - a concept related to pistols or warfare. The word *circulation* might only be known to an E3 reader in the context of liquid movement - like that of blood or water. In this way the search for meaning could lead to confusion because of difficulties on the lexical level.

During a visit to an African school, a teacher shared the following experience which explains just how important lexical features are for problem solving in mathematics. Students were given a test where the word *deficit* was crucial for solving the problem. Although this was a commercial mathematics class and children had been taught the concept of deficit, the word *deficit* was unfamiliar to most students and did not activate the necessary schemata. In fact, the key word became a stumbling block, preventing a clear understanding of the question. Asked whether children had not been taught the meaning of the word, the appropriate answer was: "*Yes, but learning is seldom linear*".

When considering the importance of lexical features, questions like the following arise: What is the nature of the lexicon of a bilingual reader? How is access obtained to this important component of language communication? Current psycholinguistic research suggests a monolingual to have a single mental lexicon connected to one conceptual system. Access to this lexicon is realized in two different ways, by sound and visual features (Nas, 1983:527). But what about the bilingual or multilingual? Is every lexicon connected to a different conceptual system? How are the different lexicons linked to one another?

A better understanding of certain processes in the mind of the bilingual reader should help to identify the needs of these students as they read second language text. Awareness of reader needs could guide the examiner in his writing. For this reason the concept of bilingualism will be discussed in more depth. Referring to the work of Hamers & Blanc (1989), different dimensions of bilingualism can be identified according to different psychological features like language competence and cognitive organisation. The latter refers to the way in which the possible lexicons and conceptual systems are represented in the mind of the language user.

**Table 3.1** **Dimensions of bilingualism**

*(Table by Ulijn & Strother, 1995)*

| | Psychological Features | | | | |
|---|---|---|---|---|---|
| | Language Competence | | | Cognitive Organization | |
| **Types of Bilingualism:** | Balanced | Dominant | Subordinate | Compound | Co-ordinate |
| **Characteristics of Bilingualism:** | same competence in both languages | competence in one language superior to the other | second language not used in native environment | same conceptual system | separate conceptual systems with overlap |

In the South African school situation dominant bilingualism prevails amongst most students. Subordinate bilingualism describes the situation in the larger part of the black school population. Many African students hardly ever speak English except during school hours and even then the speaking is restricted mainly to communication with the teacher.

Hamers & Blanc (1989) use two models to describe the degree to which a bilingual's conceptual system (long term memory or memory store) is common to the relevant lexicons:

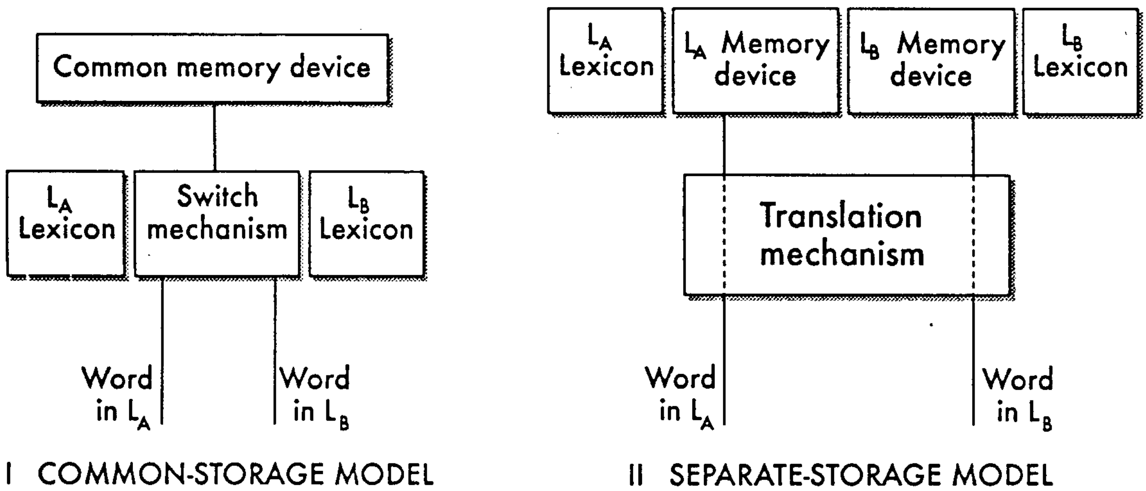
1. The *separate storage* model states that there are two independent or separate language-specific memory stores that are in contact with each other via a translation mechanism. This is comparable to the mental representation of a co-ordinate bilingual.

2. The *common storage* or interdependence model holds that bilingual memory is a single system. Language is a means of activating the common memory device which is in contact with the two lexical systems via a switching mechanism. This representation is comparable to that of compound bilingualism.

Both models advocate the existence of a mechanism which allows the bilingual to switch from one language to another.

**Figure 3.2 The two models of the bilingual's memory**

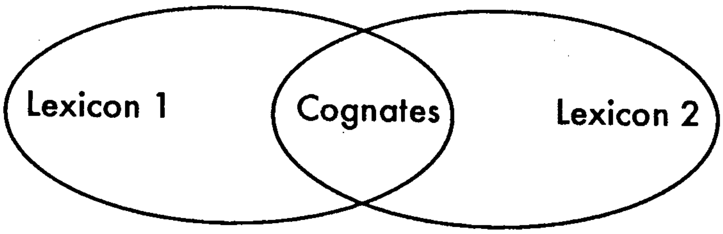
*(Model by Hamers & Blanc, 1989).*




Although both models can be used to describe different aspects of bilingualism, it seems feasible to suggest that bilingualism develops from a co-ordinate to a more compound cognitive organization. The more competent the bilingual becomes in the second language, the more the separate conceptual systems fuse into one. The on-line processing subsequently becomes quicker and more efficient. Bilinguality can therefore be described by referring to a continuum between separate and common storage although current psycholinguistic research favours the single-store model (Hamers & Blanc, 1989:101). It seems that most of the evidence from empirical studies supports the latter hypothesis.

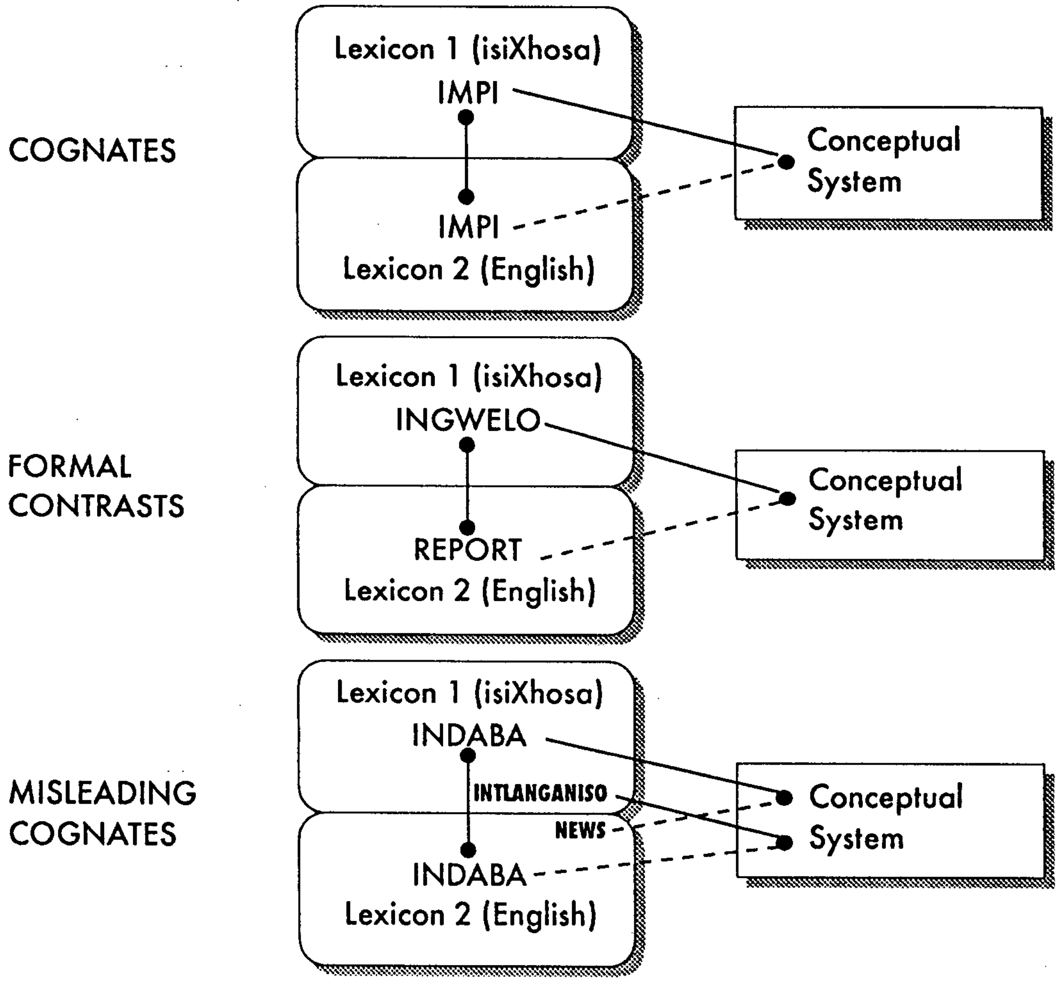
The theory of separate lexicons being connected to one conceptual system is extended by the subset hypothesis of Nas (1983:533). According to this theory, lexical items are linked together in an intricate network. The more frequently they are used, the stronger the links between the items become. So the network connecting words belonging to the mother tongue have a more stable and definite structure than those belonging to another language. Separate lexicons therefore exist for each language with empirical evidence of a common lexical store or a subset for words belonging to both languages (cognates).

**Figure 3.3 Venn diagram of two intersecting lexical sets**




The separate/common storage models of language representation as well as the subset hypothesis of Nas are important frames of references to aid the discussion of a variety of empirical data. For example, elements from the subset could be a source of positive or negative interference. Whereas true cognates could further comprehension, misleading cognates could be very confusing. The following figure is an adaptation of the model by Ulijn & Strother (1995) to illustrate the possible interference caused by true or misleading cognates.

**Figure 3.4 Three examples of two lexicons related to a conceptual system: cognates, formal contrasts and misleading cognates.**



The common lexical store (subset) of a subordinate, co-ordinate bilingual is not yet well established due to lack of usage. Elements belonging exclusively to one language can easily be mistaken for elements of the subset. This explains why bilinguals cannot always distinguish between true and false cognates even if they have been taught the difference before. The subset hypothesis also explains why bilinguals of languages with the same origin (Afrikaans and English) have the possibility of understanding the second language much more quickly than those bilinguals using unrelated languages (eg. English and Zulu). Unrelated languages not only have a smaller subset of cognates but the common linkage with the rest of the second language lexicon is less abundant and therefore weaker. The subset acts as an important thoroughfare to the rest of the second language or non-native lexicon.

The fact that access to lexical networks is based on visual and sound features explains why words looking (or sounding) the same could activate the wrong comprehension especially if the lexical network is not well founded. Words like *quality* and *equality* in English, *geruil* and *verruil* in Afrikaans and *búá* and *bùà* in Sesotho are examples of intra-lingual interference and could be a source of much confusion.(the last word-pair is the Sesotho for *speak* and *slaughter* respectively). This interference can become even worse if a word triggers the meaning of a cognate. For example, to a second language Afrikaans speaking reader the word *locust* could trigger the concept of a locus which is an English/Afrikaans cognate. The meaning of the abstract mathematical concept, *lokus* (Afrikaans) and that of an ordinary locust is very far apart.

The issue of intralingual and interlingual interference is discussed by Mägiste (1985). According to her the dominant language of a bilingual causes more interlingual interference than the weaker language. Second language readers often translate difficult text back into their mother tongue in an attempt to understand. A wrong translation due to factors like misleading cognates activates the wrong mental constructs with consequent incorrectly induced inferences.

Moreover, second language readers experience more intralingual interferences than first language readers. The instability of the lexical network can be used to explain this. Low frequency words i.e. words not so commonly used, have not been strongly integrated into the lexical network. Similarly sounding/spelled words with a higher frequency can activate the wrong concept and cause confusion. Even if the confusion is removed after the consideration of the context, the initial uncertainty retards comprehension.

The following two characteristics of second language readers emphasize their reading needs even more. Studies reveal that generally speaking a co-ordinate bilingual (cf. Table 3.1)

- has a stronger mental organization for *concrete* words than for *abstract* ones (cf. Hamers & Blanc, 1989:95).
- has more difficulty in recognizing *core* concepts than his first language counterparts (Lambert & Rawlings, 1969:607).

These findings have important consequences for second language readers who try to understand mathematics text. Many *core* concepts in mathematics are communicated by *abstract* words. If these core concepts cannot be recognized when reading a verbal problem in mathematics a student would not be able to attempt a solution even if he had the ability of doing so. This would necessarily affect achievement.

The weaker second language lexicon with all the possible interlingual and intralingual interferences emphasizes the need for careful text production. Readability guidelines referring to vocabulary in 2.5 are therefore even more applicable when writing for second language readers. A similar level of cognitive academic language proficiency (CALP) as that of first language learners cannot be assumed. Examiners should be sensitive to use high frequency words that are common to the real life of these students. High frequency words relate to the stronger part of the lexical network.

### *3.2.2 Syntax*

Both syntax and discourse refer to structure. Although most studies on second language communication focus on the lexical component of the text, the syntax and discourse levels play an important role in the logical development of an argument. The structure of sentences and discourse should therefore have an important influence on mathematical reasoning.

An inter-lingual syntactic feature that could cause more comprehension problems in a second language is high proposition density. Kintsch & Keenan (1973) have proved proposition density to be an important determinant of readability. This readability factor is equally important for the readability of paragraphs or longer pieces of discourse (cf. 2.4.3.4).

Dawe (1983) studied the effect of learning in a second language on mathematical reasoning. He found that knowledge of, and the ability to use the syntactical structures caused by logical connectors, were important indicators to distinguish between high and low achievers amongst bilinguals on a test of deductive reasoning. His extensive research on second language learners and mathematical reasoning led him to the following conclusion: *Mathematical reasoning in the deductive sense, is closely related to the ability to use language as a tool for thought. In the case of bilingual children this involves competence in* ***both*** *languages. It has been clearly shown that the ability of the child to make effective use of the cognitive functions of his first language is a good predictor of his ability to reason in a second language* (Dawe, 1983:349).

Bilinguals do not always have this "competence in both languages" to cope with difficult sentence structure in a second language. In particular, subordinate bilinguals (cf. Table 3.1) in South Africa often have a low level of English proficiency. It may be remembered that parentheses as well as hypothetical constructions cause processing constraints to first language readers (cf. 2.4.3.2; 2.4.3.3). Even more problems are experienced by second language readers especially if this type of construction is not present in the mother tongue. This is the case with many African languages. In examination text, a parenthesis and hypothetical situation are often included in the same sentence. Consider the following question:

*If $x \varepsilon R$ and $x \neq 6$ prove that k, where*

$$k = \frac{x^2 + 4x - 24}{x - 6}$$

*can have any real value except values between 4 and 28.*

(CED, 1993)

Apart from other readability problems caused by this question, the parenthesis, *if* $x \varepsilon R$ *and* $x \neq 6$, is at the same time the hypothetical condition. For second language readers this sentence structure should be even more problematic than for their first language peers. Apart from the entangled parenthesis and hypothetical condition, the parenthesis occupies the subject position of the sentence. Readers expect this position to contain important information and mathematics students are accustomed to actively use important information to guide future calculations. However, this phrase is not used in the solution of the problem and part of the parenthesis i.e. $x \varepsilon R$ is irrelevant to std 10 HG students. The examiner has inserted this information to keep the answer mathematically valid, but very few students realize this because, as mentioned previously, the std 10 HG syllabus only operates within the set of real numbers.

As indicated in 2.4.3.3, more research is needed to determine the influence of passives and actives on the readability of functional documents like mathematics text. Research by Ulijn & Strother (1990) addressed the effect of syntactical simplification on the readability of EST (English for Science and Technology) text. The subjects were first and second language readers of English. The syntactical features addressed were mainly passives and nominalizations. Using authentic and syntactically adapted versions of computer science text, this study determined that the complexity of syntax does not significantly affect the level of reading comprehension. While a complete conceptual and lexical analysis may be necessary for reading comprehension, this research suggests that a thorough syntactic analysis is not. Ulijn and Strother conclude that syntactic simplification does not seem to be any real simplification to a reader.

The above study could hold true for the specific type of text used in the research experiment. However, when considering mathematics text, the minor role of syntax in reading comprehension or the relatively unimportant influence of L1 (first language) syntax on L2 (second language) reading should not be over-generalized. There is a constant interaction between the lexical, syntactic and textual components during reading (cf. Figure 2.5). Most of the research done on the influence of syntax has manipulated this variable in isolation, keeping the other two variables constant. The question arises what the influence of syntax would be if all three language components were manipulated as variables simultaneously? A change in syntax could then have a catalytic effect as a result of interactions between lexicon, syntax and discourse. Often a whole is more than the sum of the different parts.

Another factor to take into account is the fact that syntactic simplification was researched using readers who had a certain level of language proficiency (Ulijn & Strother, op.cit.). Could this proficiency level have been enough to break through a type of ceiling effect caused by a possible readability factor like syntax? Would readers with a weaker language proficiency or severe lack of background knowledge not gain more by syntactical adaptations? According to the short circuit hypothesis of Clarke (1980), the transfer of first language reading strategies is dependent on second language proficiency. A level is reached where the weaker second language proficiency short circuits the reading strategies of the first language and readers revert to weaker strategies like serial and bottom-up reading. Syntactic simplification could guide these weaker strategies towards better comprehension. The subordinate bilingual in bilingual education in particular needs much support. Reading energy released by syntactic simplification could help a weak second language reader to cope with readability problems like an overloaded lexical network.

From the above it is clear that more research is needed to determine the influence of various syntactical structures on second language readers. It seems that the researcher would need a good knowledge of syntactical features of the second language readers' mother tongue.

### *3.2.3 Discourse*

On the discourse or textual level it can be expected that the presence or violation of any of the maxims needed for good readable text is experienced more acutely by second language readers. The following arguments should illustrate this fact:

1. The role of background knowledge can be more important with second language readers. Background knowledge forms the given or known information to which new information can be linked (cf. the given-new contract, 2.4.3.4). Comprehension difficulties due to a low language proficiency can be relieved by the reader's background knowledge of the context. A familiar context could help a reader decide

whether the meaning of an unknown word is important for the meaning of the text as a whole.

2. The readability of functional documents, like mathematics examination papers, are said to be more dependent on the overall structure of text than ordinary expository prose because readers do not only read the text, but also have *to do* something with the comprehended information. If this is true for first language readers, it should be even more so for second language readers. Simpler or more logical discourse structure could therefore help to overcome comprehension problems and support mathematization. Good structure of a logical, verbal argument describes important relationships more clearly. In verbal problems, these relationships often have to be translated into mathematical relationships by mathematical symbols. Good structure therefore not only aids comprehension, but also mathematization. In this regard one should remember that mathematical symbolism is a shorthand developed by a few closely related languages from the Indo-European group. The use of this symbolism can cause considerable difficulties to those whose native language has different structures (Austin & Howson, 1979).

3. Well structured discourse could also give the necessary support to second language readers to construct a correct contingent knowledge structure or CKS (cf. 1.5). Second language readers need this support more than their first language counterparts. If the structure of the text follows the same order that is needed for the construction of a CKS, the construction can be done more easily. In a certain sense we could say:

*structuring ease = reading ease = readability.*

The synthesis of a CKS followed by the analysis of the structure can be compared with the structuring and development of a problem space during problem solving (Newell & Simon, 1972:809). During the process of problem solving students are considered to be searching within a problem space - an imagery or hypothetical domain that includes all the available information of a problem. This problem space is forever changing as more becomes known about the problem and its solution. It is, however, important to remember that the augmented problem space evolves from the initial one. If the initial problem space has an error due to a wrongly constructed CKS, the further development of the problem space can be inhibited. Well structured discourse could give the necessary support to second language readers to operate within a specific problem space more effectively.

4. Interesting research was done by Flower and her colleagues to ascertain what kinds of strategies readers use when trying to comprehend difficult functional documents (Flower et al. 1983). Their research was guided by two questions:

a. What would a reader-based revision of a functional document look like? What do readers need?

b. What kinds of revisions do expert writers make when they revise a functional document? How do they meet the readers' needs?

They used the most significant results to coin what they call *the scenario principle.* This principle states that functional prose should be structured around a *human agent* performing *actions* in a particular *situation*. The scenario principle goes beyond the recommendations which urge writers to transform passive sentences. The research on passives pays little attention to whether the subject of the new active sentence is a person or a thing. However, the scenario principle suggests that expert revisers create a human agent or person to execute the necessary actions. This principle guides a writer to produce more explicit, concrete information when writing functional documents. In mathematics text this principle can help second language readers who often find it difficult to draw inferences from implicit, abstract information (cf. Hamers & Blanc, 1989:95).

When considering the factors reported above, one could ask what influence the restructuring of discourse would have on the readability of *mathematics text.* If comprehension could be improved by certain principles, who would gain most, first or second language readers, and would this improvement lead to better scores - also for second language readers?

But second language reading is not only a matter of language acquisition. It is also a matter of learning another culture. The culture of a people is communicated by the language they speak. The attempt of an African student to learn English is therefore also an attempt to incorporate the culture of the English language into his own. Information does not only consist of words. The words are always embedded in a certain cultural background. Therefore the following section will focus of cultural aspects of comprehension.

### 3.3 Readability and cultural factors

The school population of South Africa does not only represent a variety of languages but also a variety of cultures. The white cultures are by and large Western, while the black cultures are indigenous to Africa. The Western cultures very much reflect the characteristics of a technological society - computers and calculators are common objects in many homes. On the other hand, the African culture is to a great extent pre-technological. Often not only the language, but also the culture of the school is different to what black students experience in their homes and the environment in which they live (Berry, 1985).

Contemporary research on readability constantly stresses the importance of reader variables in text comprehension. Factors like the reader's background knowledge or available schemata have been found to play a crucial part in the interaction between reader, writer and text (cf. 2.4). This explains the ongoing call for reader analysis as fundamental to readability research. Guidelines for producing more readable text attempt to match the mind of the reader to the demands of the text. An inseparable characteristic of all readers is their culture. If reader variables play such an important role in readability as research suggests, how do cultural factors influence readability of mathematics examination text?

Someone could say, "Why should questions like these be asked? Are the truths of mathematics not universal? The theorem of Pythagoras is true to everybody, everywhere, every time. Or to put it differently: $(a + b)^2$ will still be $a^2 + 2ab + b^2$, whether a person's background is African or European. Mathematical truths just have to be learnt."

Bishop (1988:9) points out that although an attitude like the above might be reasonable, it leads to a situation where subject matter becomes of such prime importance that the needs of the learner are virtually ignored. It does not matter who comes to the learning experience, as long as everyone goes away with the same thing. Education of this kind becomes entirely impersonal and impersonal learning is far from being educational. Bishop emphasizes the fact that mathematics education cannot ignore the culture of the learner.

When considering cultural influences on the comprehension of mathematics text, three manifestations of culture will be considered. First of all, the focus will be on the relationship between *culture and language* (cf. 3.3.1). There seems to be no language without culture and no culture without language. Next, the influence of *cultural experiences* on readability will be discussed. Cultural experiences provide readers with a special type of background knowledge or cultural schemata that could influence reading ease (cf. 3.3.2). Finally there is an attempt to relate *cultural behaviour* to *mathematical culture* (cf. 3.3.3). Literature extensively reports on the influence of cultural behaviour on cross-cultural communication (Hall & Hall, 1989; Hofstede, 1992). Differences in cultural behaviour have proved to cause communication problems in written and oral communication (Wierzbicka, 1991). If, as regarded by many, mathematics is a cultural phenomenon, what aspects of cultural behaviour will influence the communication between the reader and mathematics text?

### *3.3.1 Culture and language*

Culture is a phenomenon that distinguishes one group of people from another. It is the way members of a group of people relate to one another and to their environment. *Culture* in most Western languages usually refers to *civilization*. In this study culture will be used in a wider sense and relates to what Hofstede calls *patterns of thinking, feeling and potential acting* which were learnt throughout a person's lifetime (Hofstede, 1991:4). Using the analogy of the way computers are programmed, Hofstede describes these patterns as *mental programmes* or *software of the mind*.

Culture and language seem to be inseparable entities. The Sapir-Whorf hypothesis, emerging from the work by Edward Sapir and Benjamin Whorf's writings (Sapir, 1949; Whorf, 1956), relates language to human thinking. This hypothesis suggests that the structure of a language strongly influences or even fully determines the way its native speakers perceive the world. *We cut nature up, organise it into concepts, and describe significances as we do, largely because we are parties to an agreement to organize it this way - an agreement that ... is codified in the patterns of our language* (Whorf, 1956:213). The Sapir-Whorf hypothesis has endured much criticism and has even been formulated into a weaker form. According to Au (1983:156) this form states that languages differ not so much in what can be said in them, but rather as to what is *relatively* easy to say. However, even the weaker form of the hypothesis predicts that culture and language influence each other mutually.

The language of a people reflects their culture. Cultures differ, therefore different cultures do not use language in the same way. For example, the British are considered to be a relatively individualistic group of people. They often use the subject, "I", whereas a more collective society like some of the African nations are inclined to rather refer to "we".

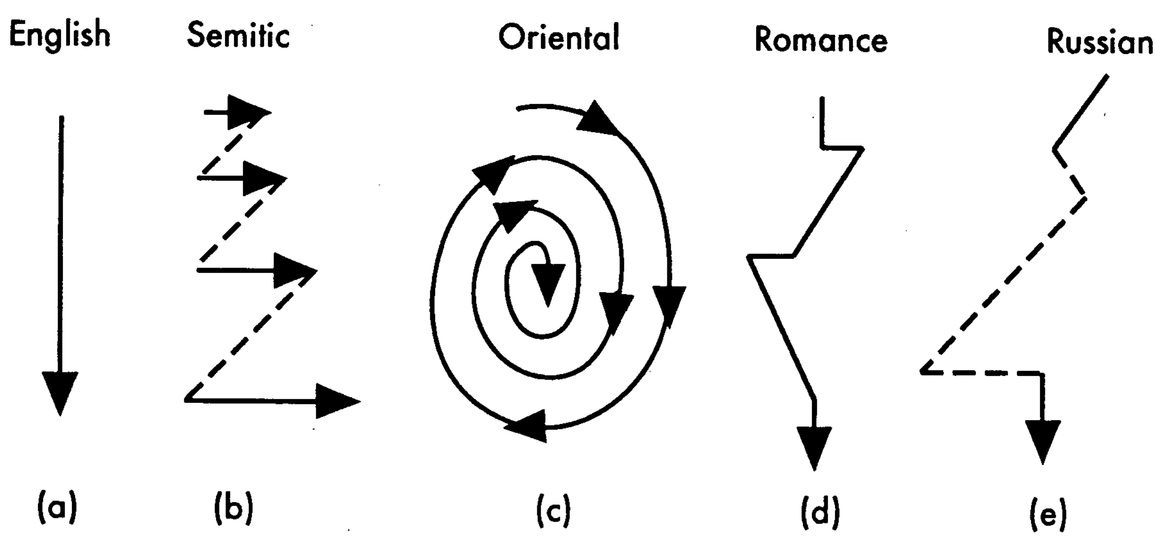
Since different cultures experience the world in different ways, not all cultures share the same concepts. For example, an Eskimo relates to snow quite differently than someone from Africa. Eskimos have numerous words for snow whereas many parts of Africa never experience snow. Therefore many African languages do not even have a word for snow. On the other hand, to Africans like the Sotho or Zulu people, it is important to distinguish between different kinds of goats, so they have a word for a black goat with white spots, a single word for a white goat with brown spots and so on. What is a void in one language can be a cultural extension in another.

Valuable studies by Kaplan (1980) have linked cultural thought patterns to language. Kaplan experienced that second language students who were brought to the level of proficiency necessary for the writing of text, wrote texts which differed in important ways from the texts written by native speakers of English. The important differences were to be found not on the level of discrete sentences, but on what he calls the rhetoric level, i.e., at the level of organization of the whole text. Referring to schema theory, Kaplan & Connor (1987) argue that it seems that schemata are not only reticulated sets of ideas, but also prefabricated sets of discourse structures specific to a language. Kaplan refers to these different rhetorical structures as *cultural thought patterns*.

The following sketch shows four kinds of rhetorical structures as contrasted to English linearity.

**Figure 3.5 Four kinds of discourse structures contrasted to English linearity**

*(Graphs by Kaplan, 1980)*




Clyne (1981) explains the cultural thought patterns as follows:

(a) Linear constructions, sequence of thought is direct

(b) Parallel constructions, with the first idea completed in the second part

(c) Circularity with the topic looked at from different angles - a more indirect approach

(d) Freedom to digress and introduce "extraneous" information

(e) Similar to (d), but with different lengths, and parenthetical amplifications of subordinate elements

If the cultural thought patterns of the traditional black South African student were to be contrasted to English linearity they seem to tend towards the patterns represented by graphs (c) and (d) in Figure 3.5.

Although the different rhetorical structures (or thought patterns) are possible in any language, certain languages have clear preferences as advocated by the weaker form of the Sapir-Whorf hypothesis (cf. Au, 1983:156). For example, the English language expresses hypothetical situations by means of hypothetical constructions more easily than Chinese or most African languages. It may be remembered that hypothetical situations are abundant in mathematics text.

Although all human populations have an *oral* language, not all cultures have developed a *written* language. The development of written language gradually created a total change in attitude towards fact and truth (Kaplan & Connor, 1987). In societies in which information is held in living memory, fact is inevitably somewhat flexible and truth mutable. This is due to the fact that information is retrieved variably. Much depends on the condition of the owner of the memory and the nature of the audience for whom the retrieval is accomplished. But once the ability to store information in invariable form exists, and once the capacity to retrieve information invariably across time and space exists, the nature of fact and truth changes. Fact becomes invariable and truth immutable. The whole development of science (and mathematics) is dependent upon invariable retrieved information.

Cultures with a written language therefore view fact and truth somewhat differently than cultures who have to depend on memory for the retrieval of information. In South Africa the European and African cultures have different histories with respect to written language. The African languages have no written records of language. Most of these languages were written for the first time by German, French or English missionaries not more than two hundred years ago. In those beginning years these manuscripts were very limited and not freely available. This could be the reason that Africans are a speaking rather than a reading people. Even in more modern times, many African students are not exposed to books in their homes. Reasons could be economic, but culture could also play an important role. This could mean that the European and African cultures of South Africa respond to communicated information in different ways.

For example, the reaction to written facts and truths could be different because of attitude differences. For the reading of mathematics text, this could have important implications. Mathematics text contains exact, invariable information.

Written information differs from oral language in a variety of ways. If one is mainly accustomed to oral communication, the ability to interact with the different nature of written information can cause subtle difficulties, especially if the language of the written information is not in the reader's mother tongue. It is a known fact that oral knowledge of a language does not in itself lead to the ability to compose written discourse. Therefore it seems likely that the grammar of oral discourse is different to the grammar of written modes. For example, written argumentation is rather close to oral argument but it differs in its elaboration, in the structures and lexicon it uses. The information is more planned and therefore more structured. Written communication differs in another important way. It offers no immediate opportunity to remove any uncertainty caused by an element in the text.

As regards mathematics examination text, many black South Africans students will find the European text less accessible not just because of linguistic reasons, but also because of cultural factors reflected in the language. A few examples illustrate this issue.

- An important void in most African languages spoken in South Africa is the word, *variable*. The absence of the *word* makes it difficult to communicate the *concept*. Variables are fundamental to mathematical thinking.

- African people in South Africa use language very descriptively, giving much detail. For example, a numerical value like 2 367 can be freely translated as: thousands which are two, hundreds which are three, tens which are six and units which are seven. When reading English text, an African needs more descriptive, explicit information than a native reader of the language. The reason is not just because of uncertainty due to a second language, but also because that is a cultural characteristic of his language. Mathematics text is very bare and very precise, giving very little unnecessary information.

- Memory is regarded very highly amongst most Africans. When interacting with mathematics text, the demand of the text is more on deductive reasoning than on memory. Moreover, the rhetorical organization of the text is according to that of an English writer and reveals a linear structure which most probably differs from that of most African thought patterns (cf. Kaplan, 1980).

Reading in a second language is therefore a bilingual as well as a bicultural experience. If the cultural nature of second language text is very different to what a reader is accustomed to in his own culture, processing constraints are inevitable, also in the case of mathematics text.

*3.3.2 Readability and cultural schemata*

The culture of a people is not only reflected in the characteristics of their language, but also by the *content* communicated by the language. Comprehending a text is always an interactive process - on the one hand, between the text and certain linguistic elements, and, on the other hand, the background or schematic knowledge of the reader. An important part of this background knowledge is the cultural experience of the reader which can also be referred to as the reader's living world. The schemata of cultural experiences are more deeply rooted than the schemata of any other experience. They have an historical background and have constantly been shaped since childhood. Stenhouse (1967:16) describes culture as *a complex of shared understandings*. This complex of shared understandings equips the reader with schemata or background knowledge that is unique to his culture.

The most extensively studied schemata are those called "story schemata." The latter is heavily dependent on cultural influences. When accessing realistic mathematics education, every verbal problem can often be seen as a small "story" with a specific cultural background.

Kintsch and Greene (1978) argue that when someone reads a long text, the difficulty is not only determined by local effects at the level of sentences or paragraphs, but also by the overall organisation of the text. Culture-specific schemata have conventional story structures known to the reader of that culture. The reported research was done with English first language students reading stories with either an English or an Alaskan-Indian background. Results revealed that stories constructed according to familiar culture-specific schemata were easier to comprehend than those constructed according to unfamiliar structures. Gains were also reported on recall and restructuring.

This research links up with that done by Carrell (1981) and Johnson (1981). Both studies report comprehension gains of stories by English second language (ESL) readers due to the cultural origin of the story. Johnson worked with Iranian university students and showed that the level of syntactic and semantic complexity of a story had a lesser effect on reading comprehension than did the cultural origin. Carrell's research was done with Japanese and Chinese students and reports not only a higher level of comprehension, but also recall gains.

Not only the cultural structure of a story but also the content and context play an important role in comprehension. ESL students are more proficient in reading when they do not have to acquire a high level of content while reading. While students should not be expected to read material almost devoid of content, they should also not be expected to develop completely new schemata on the basis of their reading (Steffensen, 1987). For mathematics reading this is an important issue. The capacity of STM is limited. Students need as much of their cognitive energy as possible for the rest of the problem solving process. The *comprehension* of the problem is only the first step.

In their research on cross-cultural influences on reading comprehension, Steffensen and her colleagues (1979) used American and Indian adults to read letters about an Indian and American wedding. Subjects read the native passage more rapidly, recalled a larger amount of information from the native passage, produced more culturally appropriate elaborations and produced more culturally based distortions of the foreign passage. The results show the profound influence of cultural schemata on comprehension and memory. But the results also illustrate another important issue. Not only the context is important. Both the Americans and the Indians were acquainted with the context of a wedding. What is even more important is the *cultural behaviour within the context*. Only if this behaviour is compatible with the reader's cultural experience can comprehension of the context be optimized.

Behaviour of variables within a certain context is important when writing "stories" used in mathematics questions. Strange behaviour can cause processing constraints. An example is a question telling the "story" of two pumps, X and Y, that are used to pump more than 60 000 litres of water to a reservoir every day. The following situation is described: *The number of hours that X operates per day may not be more than twice that of Y.* Apart from the fact that the language is used in an uncommon way, the behaviour of the pumps seems strange. Most students are not accustomed to this restriction on the working hours of pumps, especially if more than 60 000 litres of water must be pumped. The unfamiliar behaviour of the pumps could slow down comprehension as well as mathematization.

All of the above-mentioned research shows cultural effects on four levels:

1. comprehension
2. recall
3. structural issues
4. reading time

This has important implications for mathematics reading. If mathematics is tested in the context of a story, the text is more accessible if the cultural background of the story is known to the reader. A familiar cultural context gives the reader much more inferencing power. He is more able to discern between important and less important information. Not only does the readability of the text increase, but the culture-specific schemata aid mathematization. Familiar structures and extra time support the recognition of relationships which is necessary for the translation of common language to mathematical symbolism - a type of restructuring. This support is especially needed for ESL readers of mathematics (Austin & Howson 1979; Berry 1985). Mathematical symbolism and concept formation are interdependent. This means that if selected correctly the context can give the student a context-connected way of tackling the problem (Van den Heuvel-Panhuizen, 1993). In other words, the familiar context not only makes the text more readable but familiar contexts also activate mathematical thinking.

In mathematics the use of familiar contexts also helps students to grasp the intention of the problems more quickly, without an extensive oral explanation. This is extremely important for second language readers. Shorter text can reduce linguistic problems. Familiar contexts can be communicated even better with the help of pictures. In the Netherlands for instance, all Mathematics A examination tasks at pre-tertiary level are accompanied by visual aids.

An analysis of senior secondary examination papers shows that the contexts of questions are very often culturally remote from students' lives. African students in particular have problems in this area. The examiners are often European, so the contexts used are heavily influenced by Western European culture. The following linear programming question written in the context of a meteorological team measuring weather patterns on Gough Island is not only culturally remote from most students' lives, but is sure to cause many linguistic problems as well.

> *Every year a meteorological team is sent to Gough Island in the Atlantic Ocean to measure the weather patterns, giving predictions for the weather for Southern Africa.*
>
> *At least 8 weather experts are needed to ensure that all the tests are carried out and at least 3 additional members are required to ensure the harmonious running of the operation (recording, cooking, cleaning, etc).*
> *Obviously the greater the number of experts taken along would make their task easier, yet they are more highly paid, each receiving R4 500 p m while the service people are each paid R3 000 p. m. The total monthly budget is R63 000. Added to this is the fact that the transport ship can only accommodate 16 people. Let x denote the number of weather experts and y the number of back-up workers in the team.*
>
> *From the above information,*
>
> *List any inequalities which are relevant.*
>
> *Sketch these inequalities on the graph paper provided, showing clearly the feasible region.*
>
> *The optimal number of members of the team is determined by a selection panel so as to maximise P in the formula, $P = 5x + 4y$. Calculate the maximum number of experts*

*who can be taken on the expedition to achieve a maximum value of P. What is this maximum value of P?*

(JMB, 1990).

A few reasons will be given why the above text will probably have processing constraints:

1. To most students, the context is culturally unknown. It is not part of their real-life experience. Moreover, many have never heard of Gough Island and what do *meteorologists* do? How does one *record and measure* a weather pattern? Culturally speaking this described context is a violation of the given-new contract (cf. Haviland & Clark, 1977) or the co-operative principle of Grice (1967).

2. Linguistically, there is a heavy burden on STM. The text is long with a high proposition density. Words and phrases like *meteorological; harmonious running of the operation; which are relevant*, need quite a high level of language proficiency. Then there is the problem of three distinctly different words being used to denote the same people: *additional members, service people and back-up workers.*

3. Perhaps the greatest readability problem is caused by the time factor. The above question had to be completed in about 14 minutes. This included reading time.

The last cultural issue to be addressed is the possible link between cultural behaviour and mathematics as a cultural product and how this could possibly affect the accessibility of mathematics text. It has been pointed out that literature extensively reports a very definite influence of cultural behaviour on cross-cultural communication (Wierzbicka, 1991).

### *3.3.3 Cultural dimensions and mathematics*

One of the goals of mathematics education is to introduce students to the culture of mathematics as a discipline. However, mathematics is also part of the culture of all people. It is not only mathematicians that do mathematics. In fact, it is likely that people from different cultural backgrounds have different mathematical cultures. (Here, mathematical culture is viewed as a mathematical component of culture). Bishop pleads that in our attempt to introduce students to the culture of mathematics the culture of the learner has to be accommodated (Bishop, 1988).

Historically, mathematics, as an internationalized discipline, developed as part of Indo-European culture (Austin & Howson, 1979:176). It is therefore particularly important to recognize that the culture of mathematics has strong Indo-European roots and it is highly possible that mathematics as such reflects a cultural bias. The fact that mathematics as a discipline does not have an African basis does not automatically imply that those belonging to an African-related culture find the necessary skills for reading and solving mathematical problems more difficult. In this regard another weaker form of the Sapir-Whorf hypothesis could state that one is able to aquire a foreign culture *and* language besides that of one's own. There are, however, important questions to ask about the *reading behaviour* of black students reading mathematics text. Does this differ from those learners whose culture is more compatible with a European related culture? It is important that examiners of mathematics should be aware of the reading behaviour of their reading audience.

To answer questions like the above, a closer look at the phenomenon of culture is needed. A framework can then be developed to compare various cultural features of the white and black student in South Africa to various cultural features of mathematics. This framework should help to identify possible readability problems related to cross-cultural factors. There is a surmise that cultural behaviour could influence mathematical reading behaviour.

3.3.3.1 The phenomenon of culture

According to White (in Bishop, 1988:16) the functions of culture are to relate man to his environment and to his fellow-men. He divides the components of culture into four categories:

- *ideological,* relating to beliefs
- *sociological*, referring to customs
- *sentimental*, meaning attitudes and feelings
- *technological*, referring to the manufacture and use of tools and implements.

The categories are, of course, interrelated. White argues that the technological category is the most important because *all the others depend on it*. Washburn (1974) supports this argument when he says that *man became man by the use and development of tools*. Naturally this technological phenomenon of culture is not just limited to the development of machinery and tools like spades and shovels. Writers like Bruner (1964) rightly argue that man has developed by *linking himself with new implementation systems*. He calls the new systems, *amplifiers*, systems that increase the capacity of man to relate to his fellow-men and his environment. Bruner distinguishes between three kinds of implementation systems:

- amplifiers of human motor capacities, i.e. systems to increase the capacities of human motion
- amplifiers of sensory capacities, i.e. systems to increase the capacities of perceiving
- amplifiers of human ratiocinative capacities, i.e. systems to increase the capacities of logical and methodical reasoning

The last amplifier system is crucial to human development as it relates to symbols. Man is different to animals in his desire and ability to create symbols and symbol systems. The greatest of these symbol systems is speech, the forerunner of many other systems. For the purpose of this study the symbolic system of written language and the system of mathematical symbols is of prime importance.

Mathematics can therefore be considered to be

- an *amplifier*, developed to increase the thought processes of man
- an *instrument* to increase man's capacity of problem solving
- a *cultural* phenomenon.

One of the cornerstones of cultural development is therefore *the way or manner* in which a people react to specific problems. Washburn (1974) argues that a people's culture is a direct result of their tool-using way of life. Different cultures react differently to the challenge of different problems. They have a different *tool-using way of life*. Mathematics was mainly the result of the manner in which the European countries reacted to the challenge of certain problems. It was the product of their tool-using way of life. Bishop (1988) argues that if one were able to cultivate the strategies and styles relevant to the employment of mathematics, then that range of technology is open to one's use.

This concept of cultural development links up with what Hofstede (1991) calls, *culture two.* To him culture two has a much broader meaning than is usually allotted to the term culture. It not only refers to the refined fruits of civilization like arts and literature (sometimes called *culture one* by Hofstede), but also includes the ordinary things in life like how people eat, how they greet each other or respect their elders. Essentially, it reflects the way people *behave* or *react* in certain situations. Using the computer as metaphor, a more complete definition of what Hofstede refers to as culture could read as follows:

> *Culture is the collective programming of the mind which distinguishes the members of one category of people from another. National culture is that component of our mental programming which we share with more of our compatriots as opposed to most other world citizens. Besides our national component, our cultural programs contain components associated with our profession, regional background, sex, age group, and the organizations to which we belong. National culture leads to patterns of thinking, feeling and acting.*

The roots of one's culture are found within the environment in which one grew up and are learned, not inherited. There are no scientific standards that can be used to consider one cultural group superior or inferior to another. Rather, differences in culture suggest a position of relativism.

Hofstede (1991:10) distinguishes between different components or levels of culture. For the purpose of this study, the national level is important. Hofstede considers this level of culture to have four dimensions according to data he obtained from an extensive survey done by IBM, covering their personnel in 50 different countries. These dimensions of culture correspond to four basic problem areas, believed to be common to all societies. It confirms the conviction of middle twentieth century anthropologists that all societies, modern or traditional, face the same basic problems. They were convinced that the differences between societies lay in the *different* answers they provided to the *same* problems (Inkeles & Levinson, 1969:447).

3.3.3.2 Cultural dimensions

According to Hofstede (1991:13) a cultural dimension is an aspect of culture that can be measured relative to other cultures. National cultures are seen to differ mainly along the following four dimensions:

- Ways of dealing with inequality - *the power distance dimension (PDI).*
- The degree of integration of individuals within groups - *the individualism/collectivism dimension (IDV).*
- Differences in the social roles of women versus men - *the masculinity/femininity dimension (MAS).*
- The degree of tolerance for the unknown - *the uncertainty avoidance* dimension (UAI).

The four dimensions link to four behavioural areas where cultures have different *tool-using ways of life*. Cultures can therefore be positioned relative to one another, varying from a relatively high to a relatively low score in each cultural dimension. This concept of "relative positioning of cultures" will be used to develop a framework for identifying the reading needs of the two largest cultural groups in South Africa, *relative to one another*. Before doing this, the dimensions will be discussed in more detail. The decision whether a low or high score (also called index) in each dimension is more compatible with certain characteristics of mathematical culture is a personal view of the researcher and is open to discussion and more research.

***The power distance dimension (PDI)***

Power distance can be defined as *the extent to which the unequal distribution of power is accepted by the less powerful members of communities or institutions within a country.*

In a culture with a high power distance index (PDI), children are obedient to their parents, respectful to their elders and are seldom allowed to question the behaviour of those in authoritative positions. Independent behaviour by the child is not encouraged. It seems that a low PDI would link up better with the nature of mathematical culture. Mathematics needs students who are inquisitive, innovative and feel free to ask questions if they do not understand or disagree. Mathematical thinking is developed by asking questions. Students who read a mathematics question with the aim of solving it need to be able to find their own intellectual paths, must be used to critical evaluation of situations and must not feel inhibited to try out their own ideas.

***The individual/collective dimension (IDV)***

*Individualism* refers to *communities in which the ties between individuals are loose: there is a high premium on individual responsibility, for oneself and one's immediate family. Collectivism,* as its opposite, refers to *societies in which people are closely tied to each other, forming strong, cohesive groups which, throughout their lifetime, continue to protect one another in exchange for unquestioning loyalty.*

In an individualistic society, the purpose of learning is directed rather at *how to learn* than at *how to do.* Speaking one's own mind is considered to be a virtue. There is enough scope for individual approaches to problems. Confrontations and open discussions are welcomed and considered good exercise to stimulate creative thinking. In a collectivist society, confrontations would be considered to be rude and undesirable. Children from this culture seldom speak up in class. They are reluctant to give their own opinion, unless they are prompted by the group. Confrontations and conflicts are avoided or at least formulated not to hurt anybody. Children are careful not to lose "face" by asking silly questions and in this way bringing disgrace to themselves and the group. In a collectivist society the focus at school is rather on *how to do*, than on *how to learn.*

When considering the demands of mathematics in general and the reading of examination questions in particular, a culture with a high IDV would equip its children more efficiently with the skills that are necessary to tackle unknown and challenging problems. In an examination the *how to do* approach mostly does not work because the questions are different from those done in class. One is confronted with a new situation and is forced to tackle the problem on one's own. A careful, individual analysis of the text is necessary. There is a strong appeal on independent reading and thinking. When doing mathematics, one should also not be scared of making mistakes or asking critical questions. This is all part of the learning process. By eliminating incorrect solutions, one gets closer to the correct one.

***The masculinity/femininity dimension (MAS)***

This refers to the degree of *assertiveness* (masculinity) as opposed to *modesty* (femininity). The "masculine" pole is associated strongly with the importance attached to earnings, recognition, advancement and challenge. The other end of the continuum or "feminine" pole, treasures virtues like co-operation, a desirable living area and security of employment. Classroom behaviour is highly competitive in masculine cultures. Failure at school is considered disastrous. In feminine cultures failure at school is not viewed so seriously and open competition is not so eagerly pursued.

When reading a piece of mathematics text for examination purposes, the need for a challenging experience as well as the satisfaction of personal accomplishment are features that would suit the situation well. Self-confidence and an aggressive approach are needed when tackling a verbal problem in examination conditions. On the masculine/feminine continuum, it seems that reading mathematics would favour more assertive behaviour - thus a medium to high MAS.

***The uncertainty avoidance dimension (UAI)***

This can be defined as *the extent to which people feel threatened by unknown or uncertain circumstances.* At school level a culture with a high uncertainty avoidance index (UAI) expects a well structured curriculum, with a teacher knowing the exact answers. A culture with a low UAI does not feel uncertain if the teacher does not know everything. Word assignments need not be so well structured. Questions with more than one answer are easily accepted.

Referring to mathematics in general and especially the solving of verbal problems, a low to medium UAI seems more compatible with what is needed from the student. Most of the verbal problems necessitate students finding their own way. The demand for translation into mathematical symbols causes a high level of uncertainty. If the translation is wrong, the rest of the question usually cannot be done. Even if the functions are known, the computational directions are not explicitly given. Reflection and interpretation of information are expected to guide the reading process. To feel threatened by uncertainty can cause much anxiety. This is detrimental to mathematics. However, mathematical models relieve ambiguity. Strong uncertainty avoidance cultures would probably support formal mathematical solutions.

The above dimensions by Hofstede can be linked to the work of Hall & Hall (1989) who also made an extensive study of cross-cultural communication - communication across cultures. To them, culture is essentially communication. The essence of effective cross-cultural communication has more to do with triggering the correct responses than sending the correct messages. This approach is illuminating if one considers reading in mathematics. Reading the text must trigger the correct response.

Two related dimensions of culture, as suggested by Hall & Hall, will be used to identify cultural needs when reading mathematics text. One is *high* versus *low* context. The other is *mono* - and *polychronic time.* By context, Hall & Hall mean the information that surrounds the event. A high context communication is one in which the information is already in the person. Very little information or background knowledge is explicitly needed during communication. Low context is just the opposite. Cultures who are closely involved in personal relationships are high context. They have extensive information networks among their people.

When reading a verbal problem in mathematics a reader should be in the habit of searching for more information than only that which is given in the superficial lay-out of the text. Activities like the *interpretation* of non-verbal functions or the *identification* of variables and their relationship to each other are essential for expanding the given information of raw input data. The initial problem space must necessarily be augmented (cf. Newell & Simon, 1972). In general, high context people should therefore find the readability of mathematics text more difficult than their low context counterparts. (This does not imply that high context people find mathematics more difficult than low context people. The teaching experience is but one other factor that could influence a person's mathematical performance). High context people are used to communicating in another way - they usually have most of the information within themselves. But mathematics language is not a natural part of most students' information network. They usually need more explicit and extensive information. According to the work by

Hall & Hall (1989) students from low context cultures are more accustomed to *looking* and *asking for* more information.

A culture with a monochronic time structure pays attention to one thing at a time, concentrating on the work it is doing. Polychronic people are involved with many things at the same time. For the reading of mathematics examination text, it seems that a monochronic time behaviour would be more profitable. One has to concentrate on the one problem being done, without being distracted by external interferences. Although the logical development of the solutions operates to and fro between possibilities, concentration is needed for one problem at a time. The examination also has to be finished within a fixed time, which is more monochronic in nature. To people from a polychronic time culture, monochronic time is a difficult dimension to cope with.

Hall & Hall give the following list that helps to convey a pattern among cultures. The generalizations do not apply equally to all cultures.

**Table 3.2 Dimensions of culture as suggested by Hall & Hall (1989)**

| MONOCHRONIC PEOPLE | POLYCHRONIC PEOPLE |
|---|---|
| Do one thing at a time | Do many things at once |
| Concentrate on the job | Are highly distractable and subject to interruptions |
| Take time commitments seriously | Consider time commitments to be achieved if possible |
| Are low context and need information | Are high context and already have information |
| Adhere religiously to plans | Change plans often and easily |
| Emphasize promptness | Base promptness on the relationship |
| Are accustomed to short term relationships | Have a tendency to build long term relationships |
| Show respect for private property - seldom borrow or lend | Borrow and lend often and easily |

Although the parameters identified by Hall & Hall seem to be very useful as windows to cultural differences, the suggestion that some cultures are "monochronic" and others "polychronic" has to be interpreted with great caution.

Usunier (1993:442) refers to two aspects of African culture that link up with the two related dimensions mentioned above by Hall & Hall. One refers to the use of language and the other to time. Usunier confirms the traditional Africans' love of talking. Although they delight in speech, this does not make them a low context people. They use language as an instrument for the pleasure of speaking, not explicitly for the purpose of gathering important or useful information. Traditional African text is very descriptive and verbally rich which links up with their cultural attitude towards time. One seldom sees a traditional black South African doing his chores in a rush. Their concept of time is simply not the same as Europeans' (Usunier, 1993). Furthermore, their time structure reveals a polychronic nature, which often makes highly verbal negotiations seem ill-directed to people of a monochronic society.

With the different perspectives on culture, as researched by Hofstede, Hall & Hall and Usunier, it is possible to create a framework to find a relationship between the two main cultural groups of South Africa. It must be kept in mind that the positioning is relative and not absolute. For the completion of Table 3.3, the scores presented in Hofstede (1991) have been used as a guideline. Reference to black students is focused more on traditional than on urban blacks.

The framework is not intended to analyze the two cultures critically, but to develop an awareness of cultural differences. If a mismatch between cultural dimensions causes problems in oral communication, as the research of Hofstede, Hall & Hall and Usunier suggests, this is sure to be true for written discourse as well.

**Table 3.3 Relative positioning of two main cultural groups in South Africa compared with the cultural style preferred for reading mathematics text**

| Cultural dimensions | South African student of European descent | South African student of African descent | Cultural style preferred for reading mathematics text |
|---|---|---|---|
| Power distance (PDI) | Low PDI | High PDI | Low PDI |
| Uncertainty avoidance (UAI) | Medium to high UAI | Low to medium UAI | Medium UAI |
| Masculinity/ Femininity (MAS) | Medium to high MAS | Low to medium MAS | Medium to high MAS |
| Individualism/ Collectivism (IDV) | High IDV | Low IDV | High IDV |
| Monochronic/ Polychronic time | Monochronic time | Polychronic time | Monochromic time |

Ultimately people are more than the sum of their cultural dimensions. Intelligence is an example of but one other feature that makes an important contribution to the way someone tackles a problem. The above framework does, however, suggest that a Western or Euro-centric culture is more compatible with the type of cultural style needed for reading a piece of mathematics text during examination conditions. The cultural dimensions used in the framework are an indication of cultural behaviour. One might therefore conclude that cultural behaviour influences reading behaviour. Reading behaviour, on the other hand, makes text more, or less, accessible. It seems that a European related culture provides a reader with more effective "tools" to make the text more accessible. This could be due to the fact that mathematics culture has very definite European roots. More research is needed to thoroughly understand this cultural phenomenon. It could help examiners of mathematics to be more aware of the reading behaviour of their readers.

Cultural characteristics are not the only "tools" needed when attempting a mathematics question during examination conditions. There are many facets that eventually determine who a person is and what he achieves. Culture can, however, not be ignored. Every reader is automatically programmed by his culture. When teaching, one has to begin where one's students are, not where one would like them to be. Reading mathematics needs a certain tool-using way of life. This has to be taught.

## 3.4 The implications of linguistic and cultural influences on readability

Not all linguistic and cultural factors causing readability problems to second language readers can be addressed by an examiner. For example, achievement in mathematics necessitates a certain level of language proficiency in both languages. An examiner must be able to assume a certain level of cognitive academic language proficiency (CALP) in the bilingual. However, the writer of examination text should consider his audience carefully and be sensitive to linguistic and cultural needs of second language readers.

All the factors causing processing constraints among first language readers need to be considered when writing for second language students. The readability problems they cause will be even worse for second language readers mainly due to the underdeveloped lexical network. Greater care should therefore be taken when writing for these students. A less serious readability factor in a first language can cause more harm in a second language.

### *3.4.1 Implications of linguistic factors*

- The access to the second language lexicon is more difficult, so misleading cognates and synonyms in the same passage should be avoided. On the other hand, the use of true cognates should be encouraged. This helps to gain access to the second language lexicon more efficiently.

- The unstable language network of a second language reader can be overloaded more quickly than that of a first language reader. The shorter the discourse the smaller the possibility of overloading, especially in time-limited conditions like examinations.

- The call for common language is greater. Common language is more frequently used, therefore this part of the lexical network is stronger.

- Core concepts are not so easily recognized, therefore greater care must be taken to place important information in superior positions of the text. The hierarchical organization is very important because second language readers are more dependent on the overall organization of the text.

- The use of pictorials or other visual aids will do much to support the less developed abstract part of the lexicon. Sketches will also aid the visualization process so that less verbal text is needed. This implies that the visual aids be given as early as possible in the text structure. If students are expected to draw a graph and this could help to comprehend the situation better, the question asking for the graph should be done as quickly as possible.

- Examiners must be aware that text written for English second language students must necessarily differ from text written for English first language students. The reading needs of the two groups are distinctly different. First and second language groups should therefore have separate examination papers with respect to language use although the mathematical demand should remain the same.

- At the present moment Afrikaans and English students receive a bilingual examination paper. This should also be true for African students. The latter group could have a paper in English and the relevant African language of the area.

### *3.4.2 Implications of cultural factors*

- Examiners should be careful in the choice of contexts. One group of students should not be privileged above another. Contexts should be culture-free unless the culture-strange context has become part of the reader's known world during the teaching/learning situation. Abstract contexts should be avoided.

- Care must be taken that not only the context, but also the *cultural behaviour* within the context is compatible with the everyday experiences of the students.

- If possible, pictorials should be used to aid the activation of the necessary cultural schemata. This is especially important if the context is less known. Visual aids seem necessary for linguistic, cultural and mathematization reasons.

### *3.4.3 An implication of both linguistic and cultural factors*

- Examination papers should be written by examiners who share the examinees' language and culture. Only then can subtle linguistic and cultural differences that could cause readability problems be prevented.

## 3.5 Possible research questions

Two important issues evolve from the discussion of psycholinguistic theory. One is related to the effect of syntactical simplification of mathematics text on achievement. The other refers to the effect of improved readability on readers with different levels of language proficiency, the E1, E2 and E3 reader.

### *3.5.1 Questions regarding syntactic simplification*

- Will readers of mathematics text prefer a simplification of syntactical features, like changing passives into actives?

- If they do prefer a syntactic simplification - what kind of simplification?

- Will first and second language readers ask for the same kind of syntactic simplification?

- Will this syntactic simplification influence achievement levels?

### *3.5.2 An issue regarding improved readability*

- If readability were to be improved, what language group would profit most?

The reason for asking this question is two-fold. On the one hand a second language reader experiences readability problems more severely than a first language reader. This could mean that such a reader could profit more by an increase in readability. On the other hand, the weaker language proficiency could cause a ceiling effect. If the discourse is too long and not on the level of single sentences, the available reading energy, released by the increased readability, may not be sufficient to cope with the overloaded network. This could imply that the longer the discourse the more a first language reader will profit by improved readability.

After all is said and done, it is not quite clear how students themselves experience the readability of examination text. The research results used in this study up to now were taken from empirical studies done on reading comprehension of expository or narrative prose in ordinary English. Where research results did involve mathematics, it was mainly limited to primary and junior secondary students. It is not known if older students experience the same readability problems when attempting to comprehend verbal problems in mathematics text. It is also not known if these factors could affect achievement.

The best way to find out seems to be to go to the students themselves. Therefore the next chapter reports on empirical research done by using protocol analysis. Verbal reports were used as data to identify readability problems experienced by different language and cultural groups as they read and solved verbal problems in mathematics.

# CHAPTER FOUR

# THE PROTOCOL STUDY

## Rationale

*The empirical research addressed two aspects. First, think-aloud protocols were used to identify comprehension difficulties and formulate a hypothesis about readability problems students may experience with typical examination questions. The analysis of these protocols is reported in this chapter. The second aspect of the empirical study provided material for testing the hypothesis and is described in the next chapter.*

*Three different language groups did the protocol study i.e. students of European descent with English as a first language (the E1 group), students of European descent with English as a second language (the E2 group) and students of African descent with English as a second language (the E3 group). Students were all in their final year of secondary education and 17/18 years of age. The protocols were analyzed to identify possible readability problems experienced by the students as they read and solved nine mathematical verbal problems. Information gained in this way was used in conjunction with results from previous readability research to generate a hypothesis regarding readability and achievement.*

## Chapter contents

## 4.1 The aim of the experiment

In this study, the aim of think-aloud protocols was to find out from the students themselves what type of readability problems they experience when attempting to solve verbal problems in mathematics and whether these difficulties have any effect on their problem solving activities. In the previous two chapters a number of factors influencing the readability of common language was identified. The question now pursued is whether these factors also influence the readability of common language used in mathematics examination text. Examination questions in mathematics normally consist of relatively short text. It could be that the familiar real-life context or the information communicated by the given mathematical model provides enough support to overcome any comprehension difficulties caused by ordinary language. Think-aloud protocols provide an opportunity to find out more about this issue.

Knowledge of students' internal mental processes is extremely important because without it, writers of mathematics examination text have to resort to intuition and guesses about students' thoughts. Normally, what goes on in the minds of students is hidden from outside observers. As the student proceeds through the different phases of problem solving, many a struggle or search for meaning passes by unnoticed. This struggle for meaning is at the core of reading comprehension and other integrated problem solving strategies (cf. 1.5). Listening to think-alouds is a means of detecting unknown, hidden data. It is a method of identifying those linguistic variables that prevent a clear understanding of examination questions. By using talk-alouds, one can detect whether these factors inhibit mathematical thinking and eventually influence students' performance.

## 4.2 Verbal reports as data

A think-aloud protocol is a type of verbal report in which students speak out their thoughts and behaviour *without* any interference from an outside observer. The potential disadvantage of intruding and so disturbing the whole problem solving process is therefore minimized.

Thinking aloud, as method of direct observation, was developed by Newell & Simon (1972) to study human problem solving strategies. To them all problem solving takes place within a problem space (cf. 3.2.3). The nature of the initial problem space is continually changing as the knowledge states regarding the problem change. Reading can be considered a form of problem solving, since there is a chain of changing knowledge states within a hypothetical problem space. Think-aloud protocols have therefore been applied successfully in a number of research projects related to reading (Waern, 1979; Block, 1986; Samuels, 1988).

Some criticism could be raised against the use of verbal reports as reliable data. They are most informative when students are aware of difficulties and how they go about solving them. However, *automatic* processes are not easily verbalized and therefore not readily available for study. Fortunately, it is exactly this limitation that makes talk-alouds so suitable for identifying readability problems. Because students are more conscious of comprehension difficulties, they are more able to talk about them. Apart from being able to talk about the specific problem, readers usually revert to comprehension *strategies* as soon as they experience readability problems. Most of these strategies are also easily verbalized (Block, 1986). Comprehension strategies include activities like rereading, reformulation, a slower reading rate or translation into the mother tongue. Verbal reports should therefore contain important information on this matter.

Another concern is whether verbalization will not distort cognitive processes. Ericsson and Simon (1980) answer critics by pointing out that verbalization only affects problem solving activities if subjects are asked to attend to information they normally do not attend to. They firmly contend that verbal reports, elicited and interpreted with care, are a valuable and thoroughly reliable source of information about cognitive processes.

## 4.3 Constructing the experiment: methodological considerations

### *4.3.1 The pre-test*

A pre-test was designed to explore the experimental terrain first, therefore the talk-aloud investigation started with some preliminary trials. Six volunteers took part in the trial run. Each of them solved four to five verbal problems while talking aloud into a tape recorder. This experience proved to be invaluable for the actual experiment in two ways:

1. The analysis of these protocols provided evidence to support the continuation of the study. The various readability problems identified by the analysis of examination papers reported in 1.8.1 - 1.8.6 were confirmed. The analysis also pointed to some readability factors generated by the literature study of Chapters 2 and 3, like the importance of familiar contexts and the call for more common language.

2. Practical requirements like the following became clear:

- A good quality tape recorder
- The necessity for students to check whether they were being taped efficiently
- Enough time for the experiment. The talk-alouds took much longer than expected. Some E1 readers needed as many as two hours to complete only five questions whereas some of the E3 readers took three hours or more.

*4.3.2 Sampling of questions*

For the actual investigation, nine questions were chosen. The number of questions had to be a multiple of three to suit the method of testing and the statistical design reported in Chapter 5. Time-wise, not many more than nine questions could be coped with comfortably in the schools.

All the questions were original questions from previous std 10 HG examinations. They represented the following parts of the syllabus: linear programming, calculus, the quadratic function and simultaneous equations. Verbal problems to test the ability of applying mathematical knowledge in real-life contexts are often based on these sections. The selection of questions was based on the following considerations:

1. All questions test mathematical knowledge in the context of a supposedly real-life situation. They are often referred to as verbal or word problems because the text consists of much more verbal language than the usual abstract mathematics problems.

2. Some of the questions were known to have caused comprehension problems because students had complained or asked for clarification after previous examinations. It was not known how serious the problems were, but there was hard evidence of *some* comprehension problems experienced in authentic conditions.

3. All the questions represented a good mix of verbal and non-verbal language. This was important for the second half of the experiment where the same students were asked to adapt the questions by changing the language of the question to a more verbal version (cf. 4.3.5).

Many other questions could have been selected for the same three reasons. In this sense, the questions were chosen at random. The nine questions, as formatted for the talk-alouds, have been included in Appendix A. All the questions were given in the original form. Some sub-sections of certain questions were omitted for the simple reason that the available time did not allow for the whole question to be asked. The nine questions are representative of about 35% of an authentic std 10 HG mathematics examination. The rest of the examination paper consists of non-verbal problems which are not set in the context of a real-life situation.

*4.3.3 Sampling of subjects*

The following considerations guided the sampling of subjects:

1. Subjects had to be in their final school year (std. 10) and had to present mathematics on the higher grade to be able to answer the questions referred to in 4.3.2. Three schools in the Cape Peninsula region and two private schools in the Transvaal were asked to take part in the experiment. These schools had been recommended as being of a high standard. Three different groups of English readers were represented i.e. white students with English as a first language (E1), white students with English as a second language (E2) and black students with English as a second language (E3). Although the E2 and E3 groups both represented second language groups, the mother tongue of the E2 group is Afrikaans which means that to them English is a related language. This is not

the case with the black students of the E3 group. The mother tongue of this group has no European roots and is therefore unrelated to English (cf. 3.1).

2. Subjects had to be bright students of mathematics. This criterion became clear during the pre-test reported in 4.3.1. When solving a mathematics problem, mathematical knowledge interacts with language. If the mathematics proficiency of a student is too low, it is virtually impossible to distinguish between a mathematics problem and a language comprehension problem. For instance, if an average to weak student continues to reread a certain phrase, it is virtually impossible to decide whether he has problems related to readability or whether he is struggling with the mathematics. To eliminate this problem as far as possible, it was decided to confine the talk-aloud experiment to bright students of mathematics. Most students had school achievement levels above 80%. Mainly because of this criterion the students doing the talk-aloud experiment were not representative of the whole student population at senior secondary school level.

3. Six students per language group (18 students in all) were considered to be an adequate sample to provide enough verbal data for the analysis. However, more than 18 students were asked to participate in the experiment. This was to make sure there would eventually be six *useful* tapes per language group. Another important fact revealed by the pre-test was that not all students come over effectively on tape. Some of them become so involved in the mathematics that they sometimes forget to talk aloud. Others start talking so softly that, even if they do speak up, their voices are not audible enough to follow their arguments. The talk-alouds were eventually done by nine to ten students per language group. From these, the six most useful ones per language group were chosen.

### *4.3.4 Developing a framework to analyse the protocols*

Initially a framework to analyse the protocols of the pre-test was developed by using the possible readability problems identified by the analysis of examination papers discussed in 1.8.

Comprehension difficulties due to poor formulation (1.8.2) and senseless answers (cf. 1.8.3) were grouped under the category *obscure information* so that the initial framework had three categories i.e. *obscure information*, *non-verbal factors* and *visualization problems*. As the analysis of the protocols progressed, in the pre-test as well as in the analysis reported in 4.4, similar readability factors as those generated by the literature study of Chapters 2 and 3 were recognized. Not only did the literature study confirm the readability problems reported in Chapter 1, but it helped to identify additional comprehension difficulties in the talk-alouds of the students, difficulties that were not recognized during the analysis of examination papers reported in 1.8. For example, it soon became clear that students experienced the order of information input and a high level of information density as readability factors influencing comprehension (cf. Haviland & Clark, 1974; 1977; Kintsch et al., 1973; 1978). These factors play a role in the overall structure of text and therefore a category referred to as *structural problems* was added to the framework. The initial framework was ultimately extended to include all readability problems experienced by the students doing the talk-aloud experiment.

The categories of the framework represent those readability factors that prevented clear understanding of a question. These factors lead to uncertainty or confusion within the reader which again retards the processing of information. This made the text less accessible and therefore necessitated a higher level of information processing than if the questions had been presented more simply. In more serious cases the comprehension problems resulted in a complete communication breakdown.

The framework eventually comprised the following five categories:

- Difficult vocabulary
- Structural problems
- Obscure information
- Visualization difficulties
- Misleading non-verbal factors

One could consider the framework as a continuum with *difficult vocabulary* at the more linguistic pole at one end and *non-verbal factors* at the more mathematical poll at the other end. A detailed description of each category, as well as the way the readability factors relate to reading and readability theory, is given in 4.6.

Unfortunately the time factor could not be included in the framework because it was decided not to limit the time of the talk-aloud tasks. This was a pity, especially since time had been identified as a possible major problem (cf.1.8.1 and 4.3.1). However, for the sake of the protocol study it was more important that students finish all the questions so that enough verbal data could be collected on *all* nine questions. During the pre-test it became evident that a time limit on the think-alouds would have left some questions untouched. However, the questions were formatted in such a way that students had to indicate the time they had spent on each question.

The framework is not an exhaustive list. The empirical research was confined to nine word problems that were selected according to certain criteria (cf. 4.3.2). Other questions may have revealed some other comprehension difficulties. There is also a possibility that readability problems could have passed unnoticed.

### *4.3.5 Method of conducting the talk-aloud experiment*

The execution of the empirical research was heavily dependent on the support and co-operation of the principals of schools and heads of mathematics departments. They had to be willing to give their time and were co-responsible for identifying suitable students as well as arranging for venues and times for conducting the various talk-alouds. Therefore, before the actual talk-aloud experiment, the schools concerned were visited to discuss the project with them. After the permission and support of the schools had been guaranteed, meetings were arranged at each school to explain the project to the students who were willing to do the talk-alouds as well as the adaptations of the questions. This meeting was also necessary because on the day of the actual talk-aloud experiment there was very little time to sort out any uncertainty or develop a positive attitude amongst the students. It was important that the students realize just how crucial their contribution to the rest of the experiment would be so that they would take the experiment seriously. However, no problems were experienced to motivate the staff or students. In fact they were all interested as they recognized the research problem to be relevant to some of their own experiences.

The original idea was to give each of the students all nine questions. However, because of the fact that the protocol experiment proved to be very time-consuming (cf. 4.3.1), it was decided that six students per language group would still be used, but that the nine questions would be divided into two sections. Three students per language group would do Questions 1 to 4 while the other three students would do the remaining five questions. This meant three protocols would be available per question, per language group for all nine questions.

On the day of the talk-aloud experiment each student was supplied with a tape recorder, a calculator, a watch and the necessary question paper. Apart from time indications, the formatted questions also allowed space where students could write down anything they found difficult to understand (cf. Appendix A). All answers were done on the question sheets. The instructions to the students were very simple: "*First check the tape recorder to see if you are being taped efficiently. Then solve the questions to the best of your ability, but remember to talk aloud about what you are doing. Start talking aloud from the moment you start reading and don't stop talking aloud until you have finished all the questions.*" In this way data was available on all the problem solving episodes reported in 1.5. Each student then moved to an empty, quiet classroom that had been arranged by the head of department.

After they had finished, the students handed back the tapes, but were asked to take their question papers home to make the necessary adaptations which would be used in the composite test (cf. 5.1).

### 4.4 The analysis of the protocols

The analysis of protocols was done in order to identify readability problems as experienced by the *intended* audience when attempting *real* questions. As has been mentioned before, this partly answers the call to test the effect of readability in real writing (cf. Selzer, 1983). To test the effect of readability when students do mathematics is not so simple. Mathematics and reading have one important element in common and that is *meaning*. When listening to the think-alouds one has to be very careful to distinguish between a readability and a mathematics problem. The fact that the students were all very bright in mathematics did help a great deal because that meant that the mathematics part of the questions did not pose any serious problems. If one had just marked the questions without listening to the protocols one would have been inclined to think there were no significant comprehension problems either. The analysis of the protocols revealed the contrary.

To get an idea of what difficulties could be attributed to readability, the following three main sources of information were used:

- the talk-alouds
- the verbal adaptations
- a personal interview with specific students if deemed necessary.

By combining these different sources of information, it was possible to gain a better idea of the extent of the readability problem. This method of combining reader protocols with other sources of information is strongly recommended by Samuels (1988). The personal interviews were informal and were held with students from all three language groups, depending on the type of information that was needed. Sometimes a group of students was gathered for an interview and at other times a one-to-one conversation was more practical.

Because the aim was to identify *readability* problems, the main focus was on the reading and understanding episodes mentioned in the research by Artzt & Armour-Thomas (1992; cf. 1.6). A typical analysis proceeded as follows:

1. To start the analysis one carefully listened to the student as he read the question. Simultaneously the corresponding information was followed on the student's question paper. Any disturbance in the normal reading rhythm was taken as an indication that there was an interference in the ease of comprehension. Disturbances in the reading rhythm were signalled by a variety of interruptions, like some or other comprehension strategy, a stumbling over words, heavy breathing, a change in the tone of voice or

even remarks of discontent. As soon as a disturbance was detected in the reading process, the tape was stopped, replayed and listened to once again. Readability problems were analyzed and classified according to the framework developed for this purpose (cf. 4.3.4).

2. If there was doubt about a readability problem, listening to the rest of the student's problem solving strategies helped to gain clarity. For example, in Question 1 a student kept on rereading *the running costs* of the pumps before starting any mathematization. At first it sounded as if he was having problems with mathematization per se. However, when listening to the way in which he was trying to solve the problem, it became clear that he was having problems in trying to translate the information of the *running costs* into an *inequality*. A closer look at the question made one realize that the question had ambiguous information because *running costs* was not part of the constraints, but belonged to the objective function (cf. Question 1, Appendix A). Therefore the experienced difficulty was obviously due to obscure information and was classified as such.

3. Sometimes it was necessary to look at the adaptations to make sure whether the student had actually experienced a problem as a *readability* problem. For example, if a student stumbled over a word like *simultaneously* but it could not be ascertained whether the word had caused processing problems, the adaptations were consulted to see if the word had been changed to something more comprehensible. If this was the case, the word (or phrase) was counted as a readability factor. In this way the adaptations often helped to clarify the situation.

4. Often the adaptations could not help to identify the kind of problem either, because not all students were able to adapt the questions to a more readable form. It was therefore sometimes difficult to decide what kind of readability factor was responsible for the comprehension difficulty. For example, in Question 9, it was not clear what readability factor was causing pupils to revert to comprehension strategies like rereading or reformulation. It could have been a visualization problem, a structural problem or even both. In cases like these it often helped to ask what could be done to relieve the

comprehension difficulty. If the use of a sketch could ease comprehension, the readability factor was coded as a *visualization difficulty* or if a more familiar word could help, the factor was noted as *difficult vocabulary*. The following table gives an idea of the types of questions that were asked to help decide what readability factor was applicable.

**Table 4.1** **Questions asked to decide what readability factors were applicable**

| Questions | Readability factor |
|---|---|
| Will more familiar words help? | Difficult vocabulary |
| Will it help to change the composition of the sentence or discourse? | Structural problems |
| Will it help to remove the ambiguity? | Obscure information |
| Will a sketch help?<br>Will more concrete information help? | Visualization problems |
| Will the use of other letter symbols help?<br>Will it help to disentangle the verbal/non-verbal information? | Non-verbal factors |

5. Sometimes it was necessary to call students back for an interview because the adaptations signalled a readability problem that had not been identified while listening to the think-alouds. Many structural problems fall into this category. An example is to be found in Question 1 where the information for the first inequality is scattered across three sentences. Students adapted this by keeping like information together in the same sentence. When asking students why they had adapted the verbal information in this way they argued that all information referring to *litres* of water belonged together and should therefore be put together. They said they always found it difficult to understand the described situation as a whole if information that actually belonged together was scattered around in different sentences. This was especially true if time was limited.

During the personal interviews much additional data on comprehension were generated. This made one realize that not all comprehension problems can be traced in talk-alouds and also, not all comprehension problems have an effect on achievement. It seems that bright students are often able to overcome comprehension difficulties more or less automatically. They know what to look for. This seems to be the main reason why the students in this study did not verbalize *all* their comprehension problems.

For the study as a whole as well as the formulation of the hypothesis, it was important to have an indication whether comprehension difficulties could have a possible effect on achievement. Therefore, during the analysis of the protocols, all answers of students were marked according to a previously set memorandum. Because the maximum scores were not the same for all questions, achievement levels were converted to percentages. Different questions and language groups could therefore be compared more easily. During this part of the study, test scores did not play such a prominent role as the focus was more on the identification of readability problems. However, the talk-alouds made it possible to follow all the problem solving strategies of the students and therefore one could form an idea of whether achievement levels of students were influenced by a specific readability problem. Achievement levels are reported in 4.7. The results of the analysis are now discussed in more detail.

### 4.5 Results and discussion of the analysis

The distribution of readability problems as identified in the talk-alouds is indicated in Tables 4.2, 4.3 and 4.4. The tables contain the categories of the framework discussed in 4.3.4. Categories vary from a more linguistic level to more mathematical one. There are three talk-alouds per language group for each question.

The tables reflect the number of times a type of readability problem *was verbalized* by the total number of students. It should be noted that the frequency at which problems were verbalized is not necessarily an indication of the difficulty of a readability problem. It is clear that one major readability problem could cause more harm to the ultimate comprehension of a question than a few minor readability problems. For example, in Question 1 the readability problem caused by a certain non-verbal factor was reported only twice by the E1 and E2 groups, but caused more serious problems than the various readability factors mentioned under obscure information in Question 8. Moreover, some readability problems, like visualization, are by their very nature experienced less often than others. The quantitative results reflected in the tables were used to detect some similarities and differences between the three language groups.

**Table 4.2 Number of readability problems verbalized per category per question**
**Language group: E1 (English first language, white students)**

| | Question number | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Category** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **T.** | **R.O.** |
| Difficult vocabulary | - | - | - | - | 3 | 1 | - | - | 1 | **5** | **5** |
| Structural problems | 5 | - | - | 3 | 1 | - | 3 | - | 1 | **13** | **2** |
| Obscure information | 3 | 2 | 3 | 2 | 4 | 5 | 3 | 8 | 3 | **33** | **1** |
| Visualization problems | - | 6 | - | - | - | 2 | - | - | 3 | **11** | **4** |
| Non-verbal factors | 2 | 2 | 6 | - | 3 | - | 3 | - | - | **16** | **3** |
| **Total per question** | **10** | **10** | **9** | **5** | **11** | **8** | **9** | **8** | **8** | | |

**T:** Total number of readability problems verbalized per category.

**R.O.:** Rank order. (1 = verbalized most frequently. 5 = verbalized least.)

**Table 4.3 Number of readability problems verbalized per category per question**
**Language group: E2 (English second language, white students)**

| | Question Number | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Category** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **T.** | **R.O.** |
| Difficult vocabulary | 3 | 4 | 2 | - | 5 | 2 | 2 | 1 | 3 | **20** | **2** |
| Structural problems | 2 | 2 | - | - | 2 | - | 2 | - | - | **7** | **4** |
| Obscure information | 5 | 1 | 4 | - | 3 | 6 | 2 | 6 | 3 | **30** | **1** |
| Visualization problems | - | 2 | - | - | - | 1 | - | - | 3 | **6** | **5** |
| Non-verbal factors | 2 | 1 | 4 | 1 | 2 | 3 | 2 | - | 4 | **19** | **3** |
| **Total per question** | **12** | **9** | **10** | **1** | **12** | **12** | **6** | **7** | **13** | | |

**T:** Total number of readability problems verbalized per category.

**R.O.:** Rank order. (1 = verbalized most frequently. 5 = verbalized least.)

Before reporting on the E3 group, the results of the E1 and E2 students will be compared. Other than the E3 group, these two language groups share related mother tongues, English and Afrikaans.

Comparison of Tables 4.2 and 4.3 show similarities and differences. Both groups show a high frequency count for obscure information. Non-verbal interference is a definite problem for each group, while the higher count for unfamiliar vocabulary, as experienced by the E2 readers, is understandable. Visualization and structural problems were reported more frequently by the E1 group. It is difficult to say why this is so. It could be because the E1 students found it easier to verbalize their difficulties, thus giving them a higher count in this category. However, one should be careful not to infer too much from this. Nine other questions could have given a different frequency pattern. A different examiner with a different writing style or other contexts are but two factors that could influence the number of readability problems. What is important is not so much a high or low frequency count, but the fact that the readability problems actually occurred. Once the readability problems are identified, one can attempt to remove them.

When comparing the results of the E1 and E2 groups, the following facts must be taken into consideration:

- One should be careful not to conclude too much, quantitatively. The groups were relatively small. There were only three students per question, per language group.

- These two groups performed equally well on the HSRC's English and Mathematics proficiency tests. Moreover, English is related to Afrikaans, the native language of the E2 group. Some of these E2 students could be classified as balanced bilinguals (cf. Table 3.1). As regards language and culture, the two groups are therefore more similar than when compared to the E3 group.

This last fact is probably the reason why, when listening to the two groups, no significant differences between them as regards comprehension difficulties were found. However, the differences were more noticeable when listening to their *reading behaviour:*

- E1 students read more fluently and at a quicker rate.

- E1 students were able to talk more fluently and more extensively about comprehension difficulties or the way they tackled a problem.

- E2 students were emotionally more upset by comprehension difficulties than their E1 peers - they were inclined to start breathing more heavily and a change in their tone of voice betrayed a certain degree of anxiety. E1 readers were able to control their emotions better, probably because they could control the comprehension problem more effectively. As pointed out in Chapter 3, first language readers are more able to discern just how important the comprehension difficulty is for the global understanding of the verbal problem. Often the difficulty can be ignored.

The talk-alouds of the E3 group were done at a much later stage than those of the other two groups. This was mainly due to political reasons. Teaching at schools was disrupted for a large part of 1993, the year of the empirical study, making it difficult to find suitable students for the experiment. The following table shows the distribution of readability problems as verbalized by E3 readers.

**Table 4.4 Number of readability problems verbalized per category per question**
**Language group: E3 (English second language, black students)**

| | Question number | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Category** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **T** | **R.O.** |
| Difficult vocabulary | 1 | 4 | 1 | - | 4 | 2 | - | 1 | 2 | **15** | **3** |
| Structural problems | 2 | - | 1 | 2 | 2 | 1 | 2 | - | 1 | **11** | **4/5** |
| Obscure information | 2 | 7 | - | - | 6 | 4 | - | 7 | 1 | **32** | **1** |
| Visualization problems | - | - | 3 | 2 | - | 1 | - | - | 5 | **11** | **4/5** |
| Non-verbal factors | 1 | 3 | 4 | - | 1 | - | 2 | - | 5 | **16** | **2** |
| **Total per question** | **6** | **14** | **9** | **4** | **13** | **8** | **4** | **7** | **14** | | |

**T:** Total number of readability problems verbalized per category.

**R.O:** Rank order. (1 = verbalized most frequently; 5 = verbalized least).

The E3 students matched the E1 and E2 readers in the written HSRC tests on mathematics proficiency. On the nine-point stanine scale their score on the English proficiency test was seven, compared to eight and nine of the E2 and E1 readers respectively.

Before comparing the results of all three language groups, the data generated by the talk-alouds of the E3 students will be discussed in more detail.

According to Table 4.4 readability problems were experienced in all five categories. Difficulties with *obscure information* were reported most frequently and *difficult vocabulary* seemed to have caused more or less the same number of obstacles as *non-verbal factors*. Questions 7 and 9 represent extreme cases. It is not clear why. It could be that students could relate more to Question 9 - persons walking along roads - so it was easier to talk aloud about this familiar context in a second language. Question 7 had an unfamiliar context for most E3 students so it could be that they avoided any comment on the issue. Although Question 3 was similar to Question 7, there was more verbalization on Question 3 due to the redundant information in this question.

Apart from readability problems, the *reading behaviour* of the E3 group followed the same pattern as the E2 group, only the verbalization difficulties were experienced more intensely. The following three aspects were clearly audible on the tapes:

1. *The reading rhythm of the E3 students was less fluent and the reading rate was slower than the E2s'.*

As already mentioned, the protocol experiment took markedly longer with the E3 group than with the other two groups. The fact that the E3 group was asked to verbalize in English could have been responsible for this. Some of the E3 students could be classified as subordinate bilinguals (cf. Table 3.1).

2. *E3 readers were inclined not to verbalize as easily or as much as the other second language group.*

This can be understood if one takes into consideration that they were asked to verbalize in English. English is not related to their mother tongue and many E3 readers do not have as much opportunity to speak English as the E2 group. Some E3 readers from the rural areas said they did not speak more that 15 minutes of English per day. Unfortunately, therefore, the E3 tapes had much less verbal data. What made matters even worse was the fact that two pupils had worked through all the word problems before switching on their tape recorders. It seems as if the students were scared of losing face (cf. 3.3.3.2). Another student switched off the tape recorder as soon as he had serious problems. This meant that even fewer comprehension problems were verbalized on these three tapes than on the other tapes from E3 readers.

3. *Comprehension difficulties seemed to cause even more anxiety to E3 readers.*

They sounded more upset and more discouraged, especially if they noticed that time was passing without their being able to make any progress. For example, one E3 reader took 47 minutes to do Question 5 but could not get further than Question 5.3. (If students were to take more than about 16 minutes for a similar question in authentic examination conditions, they would start running out of time.) The mathematization and graphs were done correctly but the rather difficult language of the otherwise relatively easy Question 5.3 caused a complete comprehension breakdown. She kept on saying, *"I don't understand the question... Oh, I don't know...Oh, I'm taking too much time...."* (More or less the same thing happened with the short and easy Question 7). After another 13 minutes she despondently said, *"I'll come back to this one later if I have time..."*

Irrespective of differences in reading behaviour, the talk-alouds revealed definite *reading needs* specific to E3 students. All the implications of linguistic and cultural factors discussed in 3.4.1 and 3.4.2 were confirmed by the talk-alouds. In particular, the following aspects were re-affirmed:

1. *There is a definite need for more time to read and process information.*

2. *There is a definite need for more plain language.*

Although comprehension problems were experienced by all groups, a complete communication breakdown, as illustrated above for Question 5.3, occurred more often with E3 readers. Although this question caused some E2 readers comprehension difficulties as well, they were able to overcome the problem and complete the question. This reminds one of what was reported in Chapter 3 that second language readers find it more difficult to recognize core concepts and abstract words especially if the mother tongue is not related to the language of the text (cf. 3.2.1). They are therefore less able to distinguish between important and less important information in a second language.

3. *There is a need for more explicit and more extensive information.*

Because of a weaker language proficiency, E3 students are not able to make as many inferences as their first language peers (cf. 3.2.1). It is quite natural that the less one knows of a language or context the more explicit and extensive the information must be. The adaptations of the E3 pupils had one conspicuous recommendation in common: more explicit information. Consider the following two versions of information for Question 6. The first is the original version and the second, the adaptation as suggested by an E3 pupil.

***Original version***

*Two cyclists A and B travel from the same place and leave simultaneously. The distance S, in kilometres, that both of them are from the starting point after t hours, is respectively given by the formulae,*

$$S = t^3 - 12t^2 + 36t \text{ and}$$
$$S = 24t - 4t^2$$

*Both are back at the starting point simultaneously.*

***Adapted version***

*Two cyclists, John and Peter, start from the same starting point at the same time. They travel along two different straight roads.*

*Distance of John from the starting point to his destination*

$$= (t^3 - 12t^2 + 36t) \text{ kilometres}$$

*Distance of Peter from the starting point to his destination*

$$= (24t - 4t^2) \text{ kilometres.}$$

*The time (t) is in hours.*

*Both Peter and John are back at the starting point at the same time.*

Although the above adaptation addresses a variety of readability problems, the need for more explicit information is obvious.

4. *The possibility of cultural interference due to* ***unknown contexts*** *is greater.*

Because many E3 readers come from a pre-industrial society, the concept of profit is not well developed. Questions 3, 5 and 7 all required students to maximise the profit. All E3 students had problems with these questions which could possibly be traced back to cultural interference. For example, to some students the word *cost* triggers the message of selling price. The phrase *cost price*, well-known in Western culture, is less familiar to many a secondary-school African student. Another example is the difference between speed and average speed (Question 6). E3 students in rural areas do not often experience situations where average speed is applicable or needed. This is contrary to the culture of E1 or E2 students. To many of these students average speed has been a well-known concept since early childhood.

Cultural interference is a readability problem that could cause some controversy. One could rightly argue that an important objective of mathematics education is to extend the known world of students and in this way prepare them for adult life. This would mean introducing *unknown* cultural contexts and expanding the everyday world of the learner. On the other hand, if this extension has not been realized in teaching practice, it is unfair to assume these contexts as familiar in examination conditions. Depending on the nature of the question, this could affect the validity of test scores (cf. 2.1). Real-life contexts should be part of a student's available schemata if they are to support comprehension and problem solving during examinations. This matter relates to the following issue.

5. *The possibility of* ***unfamiliar cultural behaviour*** *is greater.*

Although the context may be known to students, there is a possibility that the behaviour within the context is not. Referring to Question 3, one could say that even though the context of producing and selling calculators may be familiar to students, the behaviour of the selling price is not - especially in a pre-industrial culture. Many students are not aware of the fact that the selling price of an article could be a function of the number of articles sold. To them the selling price is fixed - whether you sell one or one hundred articles. This study has indicated that to aid readability, students should be able to relate not only the context, but also the behaviour within the context, to their real-life experiences. (cf. Steffensen et al., 1979; 3.3.2)

6. *E3 students are more dependent on the structure of text to guide comprehension.*

Although structural problems were verbalized more frequently by the E1 group (cf. Tables 4.2 and 4.4 ), structural problems caused the E3 group more serious problems. The need for well structured information could have a linguistic as well as a cultural basis. Good structure can support a relatively weak language proficiency and aid the construction of a contingent knowledge structure. Furthermore, there are traditional African people who prefer to tackle a task consisting of a number of smaller tasks by doing one task at a time. They are often confused with a number of simultaneous commands. For example, when purchasing a number of goods, an African traditionally prefers to buy and pay for one item at a time. This perhaps explains why the entangled information of Question 5 was more difficult for E3 students to cope with than for the E1 and E2 groups. One student said, "*The information is all mixed up - this tends to be confusing.*" One should remember that students may not ask for clarification of text during examination conditions. In real life, one could have an on-the-spot experience of a situation or one could ask for more information. This makes oral communication distinctly different from written discourse.

The fact that much less verbal data was available on the E3 tapes makes it rather difficult to compare the frequency table of the E3 readers (Table 4.4) with that of the other two groups (Tables 4.2 and 4.3). However, in spite of this restriction one does notice more or less the same pattern as that of E1 and E2 readers. The following Table is a summary of Tables 4.2, 4.3 and 4.4. There were six students per language group.

**Table 4.5 Number of readability problems verbalized per category per language group**

| Category | E1 | | E2 | | E3 | |
|---|---|---|---|---|---|---|
| | Language group E1 | | Language group E2 | | Language group E3 | |
| | **T** | **R.O.** | **T** | **R.O.** | **T** | **R.O.** |
| Difficult vocabulary | 5 | **5** | 20 | **2** | 15 | **3** |
| Structural problems | 13 | **3** | 7 | **4** | 11 | **4/5** |
| Obscure information | 33 | **1** | 30 | **1** | 32 | **1** |
| Visualization difficulties | 11 | **4** | 6 | **5** | 11 | **4/5** |
| Non-verbal factors | 16 | **2** | 19 | **3** | 16 | **2** |

**T**: Total number of readability problems verbalized per category.
**R.O.:** Rank order (1 = verbalized most frequently; 5 = verbalized least).

The rank orders of all three language groups are more or less the same except for the more extreme linguistic category, *difficult vocabulary*. It is clear from the table that second language readers, both the E2 and E3 groups, reported more problems in this category than their first language counterparts. This tendency is understandable if one considers the weaker second language lexicon (cf. 3.2.1).

The interpretation of the tables should, however, be done with care. In this respect it is important to point to at least two findings that seem to be obvious at first sight, but on closer scrutinity do not correspond with the evidence of the talk-alouds.

1. It seems as if structural problems are not such a problem for second language readers because the rank order for both the E2 and E3 groups is fairly low in this category. Listening to the talk-alouds one realizes this is not true. Although *structural problems* were not verbalized as often as for instance *obscure information*, structural problems in Questions 1 and 5 caused some of the second language readers serious difficulties. In some cases a complete communication breakdown occurred - something that never happened with first language readers. An analysis of the adapted verbal versions (cf. 5.1.2 and 5.1.3) shows that overall structure, in sentences as well as text, was one of the important issues addressed by second language students.

2. According to the rank orders in the table, *visualization difficulties* seem to be less of a problem for all language groups. This is not necessarily true. Visualization problems, because of their very nature, were not reported as often as for example *non-verbal factors* and yet only one visualization problem could be enough to upset comprehension seriously. For example, for Question 2 most students could not visualize the context. As for Questions 1 and 5, these caused more uncertainty to second language readers.

The *visualization difficulties* in the tables can therefore not be interpreted as being of lesser importance without listening to the protocols themselves. This is true of readability problems in all categories. One has to keep in mind that both second language groups did not verbalize as much as the first language group. For example, if there had been more verbal data of the E3 group, the frequency counts would probably have been higher, especially that of difficult vocabulary.

The importance of Tables 4.2 - 4.5 lies in the fact that they provide enough evidence of readability problems that interfere with comprehension. When listening to the protocols one realizes that these types of problems are common to all language groups and that most problems cause more difficulties to second language readers. For readers who veer towards being subordinate rather than balanced bilinguals (cf. Table 3.1), most problems are more serious.

## 4.6 Discussion of readability problems per category

The discussion is done by means of the framework suggested in 4.3.4. All five categories of readability problems were identified in the first question, therefore the comprehension difficulties of this question are discussed in more detail and serve as an example for readability problems in general. However, other questions will be referred to as well although it is obvious that not all types of readability problems have the same level of difficulty in the different questions. Furthermore, it is not practical to discuss *all* the comprehension difficulties experienced by *all* the students in detail. Enough information is nevertheless given to illustrate readability problems sufficiently and indicate the possible influence on achievement. When noting the comprehension difficulties, one should keep in mind that all the students were high achievers of mathematics. Weaker students could have experienced the readability of the questions differently.

### *4.6.1 Difficult vocabulary*

This refers to unfamiliar words or phrases. Difficult vocabulary is less able to activate the necessary schemata in the minds of readers (cf. 2.3.1). Far less information can be deduced than if the vocabulary is familiar. Question 1 has the word *reservoir* that was not familiar to some of the second language readers, both E2 and E3. One student stumbled over the word while reading the text and omitted the word in his adaptation. A *phrase* that interfered with comprehension ease in Question 1 was, *running costs.*

Question 5 seems to have a few other words and phrases that caused problems even to first language readers: *utilized, profit margins, optimal search line, daily capacity.* Question 5.6 contains a number of these words. Most students could not answer this question correctly, mainly because of comprehension difficulties. One student reread the question a few times and then said, *"Mayo, Mayo! I really don't know what's going on here... I'll do my best... I don't understand the question."*

### *4.6.2 Structural problems*

This category refers to problems related to the overall organization of text, whether in sentences or general discourse.

Factors related to the overall structure of text have been reported to be one of the most important reasons why readability formulae are not able to guide writers effectively in the production of more readable text (cf. 2.2 ). Readability formulae focus on readability of sentences and fail to address the comprehensibility of connected discourse. In mathematics, the logic reflected in the structure of text not only influences comprehensibility, but also affects deductive reasoning that is necessary for the ongoing process of problem solving. Mathematics educators agree that the method of solution often arises from the structure of the verbal problem (Van den Heuvel-Panhuizen, 1993).

Structural problems were caused by:

#### 4.6.2.1 Order of input

This problem refers to the order of input not being the same as that needed for the construction of a contingent knowledge structure or CKS (cf. Bobrow & Brown, 1975; 2.4.3.4).

The order needed for constructing a CKS is not to be equated with the order needed for the solution strategy. However, often the wrong order for constructing a CKS coincides with the wrong order needed for the construction of a mathematical sentence. For example, in Question 1 there is a restriction on the number of litres of water to be pumped to a reservoir, but the information referring to litres of water has been scattered across different sentences. The same applies to the information regarding the costs. This not only makes it more difficult to understand the situation, but the mathematization is also more difficult than if like information had been grouped together. As reported previously, students argue that like information should be put together (cf. 4.5). However, if the order of information input in this question were changed to aid comprehension, the mathematical demand would be reduced.

Question 5 is an example of the same comprehension problem. The "story" is about two departments that both experience constraints regarding the number of hours available for producing two products. The restriction is within each department, therefore it would be easier to understand the situation if the data belonging to each department were given separately. However, here the information about the two departments is entangled and scattered across different sentences, a mixture of a one-to-many and a many-to-one distribution of information (*one* verbal sentence has information belonging to *more* than one mathematical sentence and *more* than one verbal sentence has information belonging to *one* mathematical sentence). This type of discourse structure necessitates a high level of information processing, not only for ordinary comprehension, but also for mathematical meaning. Most students were able to overcome this difficulty, but only after they had mathematized the first inequality wrongly as

$$2x + 4y \leq 60 \qquad \text{instead of} \qquad 2x + 3y \leq 60$$

On the question paper each student had an opportunity to indicate what he had found difficult to understand. One E2 pupil wrote a remark at the end of Question 5 while talking aloud at the same time. He said: "*All last sentences have something important about the equations. Difficult to extract these aspects from the sentences*". On the tape he explained this further by saying, "*What one finds difficult here to understand is... as I made the error myself,... is that if you look at it and you see A requires 2 hours per unit in $D_1$ and 4 hours per unit in $D_2$...* ***Usually*** (accentuation by student) *you would think it will be 2 hours per unit in $D_1$ and then* ***B*** *will take say 3 hours in $D_1$ also*". He was trying to say that one would think that the information in the first sentence would all belong to $D_1$. He continued: "*But this time they give it... you ... eh ... so that you require the information that A takes two hours in* ***$D_1$*** *for the first but not for the second time... ja..eh, in the first equation and then the four hours you require in the second equation. So it is different equations and you have to work the* ***whole*** *thing...eh.. the* ***whole*** *last three sentences, you have to work* ***together*** *to get just two equations. You have to use each of those last three sentences which makes it rather difficult.*" The error to which he is referring is the wrong inequality, $2x + 4y \leq 60$.

4.6.2.2 Hierarchical progression

Text composed according to this principle is arranged according to main ideas, subordinate ideas and their interrelationships (cf. 2.3.3). The composition of the text in Question 3 has in a sense violated this principle by inverting the hierarchy or order of importance. The irrelevant information regarding 8000 calculators is given in the prominent first sentence, whereas the crucial information that the selling price refers to only *one* calculator is reserved for the inferior last position of the text. As has been mentioned, some pupils did not even verbalize the important last two words. Whereas oral communication can accentuate more important information by a heavily stressed word or phrase, written communication must rely strongly on the hierarchical structure of text. Research reported in Chapter 2 has confirmed this by pointing out that good readers process text hierarchically, therefore important points in a text should be placed in superior positions according to a hierarchical order (cf. Huckin, 1987).

4.6.2.3 High proposition density

Propositions are units of information used as building blocks by both readers and writers. Selzer (1983) refers to proposition density as possibly one of the more important determiners of readability. As mentioned before, research has shown that reading time increases as the number of propositions increase, even if the number of words and phrases remains the same (cf. 2.4.3.4).

During the analysis of protocols, a high proposition density was detected in sentences as well as in paragraphs, thus influencing syntax as well as the overall composition of discourse. Reading of verbal problems in mathematics necessitates a receptive and critical reading style making proposition density an important readability factor (cf. 2.4.1.2). All nine verbal problems contain condensed information which is often highly congested. For example, in Question 1 different facets of the problem are all communicated in just one paragraph. This makes the proposition density of the paragraph relatively high. Students are used to the practice of ordinary reading, where different paragraphs distinguish between ideas and in this way relieve the proposition density to some extent.

Once again E2 and E3 readers found it more difficult to cope with this readability factor than their first language counterparts. A high proposition density usually increased the reading time of both second language groups. In the adaptations they addressed this problem more frequently, thus confirming the problem they had experienced. Readers seem to need a certain amount of reading space to comprehend text more readily. If text is too congested with information, it seems to smother comprehension. Time-limited tests would naturally increase this problem because the rate of proposition-input would be even higher.

4.6.2.4 Syntax

Syntax seems to be more of a problem to second than to first language readers. Although the E2 and E3 readers performed well on the HSRC language proficiency tests, their talk-aloud reading behaviour was distinctly different to that of their first language counterparts. This was especially true of E3 readers to whom English is a non-related language. They read much slower and were more inclined to follow a serial reading process than the partial parallel pattern described in Chapter 2 (cf. 2.4.1.1; Fig. 2.5). Instead of a top-down strategy, they tended to follow a bottom-up strategy, giving much more attention to superficial detail. This often resulted in their getting entangled by the syntax at the expense of the global message. To mathematize a verbal problem, comprehension of local detail is necessary, but the global understanding of the problem is just as important for the translation of relationships between variables.

An example of syntactic problems was illustrated by Question 4. The hypothetical construction of two relatively long sentences made the information difficult to process. Both first and second language readers had problems. These sentences were reread a few times. Two E1 readers reread these sentences even before reading the posed questions of the metatext. Remarks by E3 readers indicate that not all the information implied by the subjunctive mood was understood correctly. But E1 readers also seemed to have problems in understanding the two hypothetical conditions. They used x as the number of pupils for each of the *different* hypothetical situations. (The initial sentence refers to x as the number of pupils in the class). This could have been a mathematics error, but it was strange that two out of three bright E1 students made the same mistake. Apart from this error, their mathematics was executed perfectly, but x works out to be a negative fraction. Both realized this could not be correct. After much rereading and cross-checking of calculations they gave up. The discourse structure of this question is an example of hypothetical conditions being communicated by hypothetically constructed sentences. This readability problem confirms the earlier finding from literature (cf. Au, 1983; Dale & Cuevas, 1985).

The syntactical issues concerning active and passive voice, as discussed in Chapters 2 and 3, were not detected while listening to the protocols. However, the adaptations revealed some interesting data that will be discussed in the following chapter when the adaptations are reported in more detail.

### *4.6.3 Obscure information*

Information of this kind is not easily or clearly understood and causes uncertainty within the reader. Different forms of obscurity were identified.

#### 4.6.3.1 Meaning not clear

Information of this type could either be ambiguous or the meaning could be obscure for another reason. One of the important reasons for not understanding the information clearly seemed to be the use of unnatural verbal language which sometimes reminded of a direct verbalization of the mathematical model. Referring to Bachman (1992), this means the real-life experience is being described in a language which is not *authentic* - not the same as normally used in the situation. Part of the information in Question 1 serves as an example: *The sum of the number of hours that the pumps operate per day may not exceed 20.* Almost all the students reread this sentence a number of times. Although the language may be common to a specialist in the field of mathematics, pupils at school level do not speak this way. The information also contained a degree of ambiguity because some students were not sure whether *the sum of the number of hours* referred to each pump separately or to the two pumps together. When reconsidering this last uncertainty, one realized that if a pump were to work for several different periods during a day, one could also refer to *the sum of the number of hours* of one pump only.

A question that appeared to be very obscure was Question 8. Some of the readability problems of this question have already been discussed in 1.8.2. One of the major problems of this question was the behaviour of the magazines as communicated by the function. Although there may be something like *the rate at which the circulation of a magazine is changing three years from now*, this behaviour is strange to pupils. It is not part of their real-life experience. After rereading and a few sighs, most pupils gained full marks for the question, but they used the mathematical cue, *rate of change*, found the first derivative and plugged in the number, three. There was little understanding of what the answer meant in real life. If the rationale behind realistic mathematics education is to bring the students' daily life into the classroom, this word problem partially failed.

When trying to do Question 4, two E3 pupils commented,"*Why do some pupils receive less books if there are books left over*?" The *some pupils* most probably refers to the five pupils who joined the class and who, according to the way the readers understood the problem, received fewer books than the others. It seems that for some E3 readers, the information of this question was too implicit to be clear.

4.6.3.2 Confusing information

The information is confusing because of factors such as inconsistent use of words or phrases, redundant information or even missing information. It has already been mentioned how confusing some students found Question 1.1. They were asked to represent the given information as a system of inequalities but not all the given information belonged to the constraints of the problem (cf. 4.3.5). This question confused one student so badly that his achievement level for the whole of Question 1 was only 25% although he attempted 62% of the question. He kept on saying, *"How can I find the inequality of the cost, if I don't know how much money is available?"* He was so confused that he started reading other information incorrectly. For example, he read 60 000 as 6000, three times. Therefore readability problems in one part of the text sometimes caused comprehension difficulties in other parts where there was no reason for such a problem. As Smith (1982) has indicated, there can be different reasons, quite independent of the reader's reading ability, that make reading more difficult - *by* *being given difficult material, by being required to pay a lot of attention to every word, or by being put in a condition of anxiety.* All of these reasons increase the demand for visual information and have the paradoxical consequence of making it harder to see the text. The struggle for meaning seems to have a ripple effect.

Other examples of confusing information are to be found in Questions 3 and 5. In Question 5, the departments are said to have available 60 and 80 hours per *week* but the students are asked to determine what percentage of *daily* capacity is utilized in each department. In Question 3 there is an example of missing information. The unit of the selling price is not given. When students tried to interpret the function so as to experience the situation better, they became confused. How should the function be interpreted? A selling price of 400 rands or 400 cents is not true to real life.

Question 3 also has redundant information in the very first sentence. Research has shown that the initial mention of an idea serves to guide the reader's abstract processes - readers are inclined to accept the information in the first sentence as the most important information of the paragraph (cf. Kieras, 1980; Faigley & Witte, 1983). In mathematics, pupils are used to *doing* something with important information. What makes matters even more complicated is the fact that students in South Africa are virtually never taught how to cope with redundant information. It was therefore not surprising to listen to just about all students as they used the redundant information of 8000 calculators in their attempt to answer Question 3. Research by Arter & Clinton (1974) found irrelevant information to be a serious problem in time-restricted tests. Although Question 3 was relatively easy, students experienced the question as difficult because of a variety of comprehension problems.

#### 4.6.3.3 Cultural interference

This could happen when the described context is not part of the student's everyday or cultural experience. The answer to Question 6.2 is -8 km/h. Pupils not doing Physical Science do not have "negative speed" as part of their frame of reference. One student said, *"This can't be right because he won't be riding backwards".* In Question 3, the concept of *profit* is not part of the cultural experience of many students from rural areas. In these parts much trading is done by the exchanging of goods. Part of the empirical research was done in a rural area and personal conversations with the students confirmed that to many, profit was not a common concept.

Cultural experiences form part of the necessary schemata or background knowledge that help readers to gain access to text more easily. When these schemata are missing, no activation of semantic contexts can take place (cf. 2.3). Instead of acting as the initial data for the given-new contract (cf. 2.4.1.2), the unfamiliar experiences become an interference that impedes comprehension, as illustrated above. In examination conditions this could have important repercussions.

4.6.3.4 Contradictory or senseless information

This occurs when data generated by the interpretation of a mathematical model does not fit the given verbal information. For example, in Question 2 the verbal information refers to *distance*, but the given mathematical function describes a *position*-time function. Some students understood the *distance* to mean *displacement* due to the negative answer to Question 2.1. During a conversation, one student remarked, "*I wasn't sure. The examiner referred to distance when he actually meant displacement. Does that mean that when he refers to speed in 2.3, he actually means velocity?*"

The functions used in Questions 3, 6 and 7 do not fit the described situations in every way. As soon as students started interpreting the functions, comprehension difficulties arose. The unreal behaviour of the function representing the selling price in Question 3 has already been discussed in 4.5.2. In Question 6, interpretation of the function generates information of two cyclists who begin cycling at 36km/h and 24km/h respectively and immediately start slowing down. This of course is not what happens in real life. One cannot *start* at that speed and what sense is there in immediately slowing down? What is more, they apparently turn back when their speed is zero - for a moment. This also cannot be true, but it is the only way it seems possible to answer Question 6.4. Cyclists do not behave like thrown-up stones. If one throws up a stone vertically into the air, the stone "*turns around*" when the speed is momentarily zero, but a cyclist on the ground turns back to his/her starting point in a variety of other ways and for various other reasons. Students experienced many problems when tackling Question 6.4. A common question was, "*How do I know when the cyclists turn back?*" The behaviour of the cyclist, assumed by the examiner, was not the way students experience their cycling. Some students understood that the cyclists had begun to return to the starting point the moment they were back at the starting point, that means when s = 0. In the original question, a graph was asked for at the end of the 35 mark question. This could perhaps have helped pupils to solve the problem the way the examiner had in mind, provided they did not try to interpret the function too intelligently. A graph would not have saved the use of an inappropriate function. In spite of these comprehension problems some students still managed to score high marks. They did what they thought they should do or looked for mathematical cues like *speed* which usually signals the use of the first derivative.

When one tries to interpret the functions of Question 7, the comprehension difficulties are similar to those of Question 3. What makes this type of information difficult to process is the fact that the cost price function refers to the total cost of *all* x hi-fis, while the selling price function refers to the price of *one* hi-fi in terms of *all* x hi-fis sold. Most problems were caused by this last function. It is not true to real life. This is one of the reasons students did not pick up this "vital" piece of detail among much other information. Those who did try to interpret the function were confronted with the following: The function suggests that in order to sell zero hi-fis, the selling price must be 1000 rands but in order to sell one (and no more), the selling price must drop by one rand. So every time one drops the price with one rand one is able to sell *exactly one* more hi-fi. "*It doesn't work that way*", one student remarked. Another student could solve neither Questions 3 nor 7 correctly, in spite of the fact that his achievement level is usually well above 90%. During the talk-aloud, he considered the profit function he had constructed and said, "*This seems to be wrong. I can't figure out why*". In the personal conversation he said the information on the selling price had confused him. This question is an example of a question where the context is known to many pupils but the *behaviour within the context* is not. Information of this kind is counter-productive. Instead of the real-life situation encouraging mathematical thinking, the *unreal behaviour* inhibits the process.

*4.6.4 Visualization problems*

Information that is either too abstract or too condensed often makes it difficult to form an image of the communicated context. Readers find sentences that are easy to visualize easier to understand than low-imagery sentences (Holyoak, 1974; cf. Paivio, 1979).

The limited information on the pumps in Question 1 makes it difficult to visualize the situation and understand why one pump is able to pump so many more litres than the other one, or why one pump is so much more expensive. This was one of the readability problems mentioned by the students during the personal conversations. The unfamiliar behaviour of the pumps was not elucidated by the text and students found it difficult to imagine or experience the constraints of the problem. As has been mentioned before, real-life experiences aid pupils' comprehension as well as their methods of solving verbal problems. This is one of the pillars of realistic mathematics education (cf. 1.1). If, however, the behaviour within the context is difficult to visualize, the information remains abstract, no matter how known the context may be. It is clear that if real-life contexts are difficult to visualize, they cannot trigger comprehension or mathematical thinking the way they could. To be able to experience a story within a certain context, familiar *behaviour* seems to be just as important as the familiarity of the context itself. If the behaviour is unfamiliar, more explicit elucidation is necessary.

In Question 5 the rather abstract information does not allow students to form an idea of the type of product or the kind of departments to which the examiner is referring. In one talk-aloud a student remarked, "*Too many figures and numbers; it gets very confusing.*" During the personal interview another student said, "*I tried to form an idea of what was going on by taking product A as ties and product B as shoes, but that caused more problems when moving over to the departments. Ties would need too much time, especially when compared to the shoes.*"

Apart from the problems caused by the inappropriate function, Question 6 also gave rise to visualization problems. In one of the talk-alouds a subject reread the text three times before he started to answer the questions. He kept on stressing the words, "*Both are back...*" Then continued, "*What do they mean?...Sa must be the distance of cyclist A*" (The notation Sa was unknown to him. An analysis of his textbook revealed that this notation was not used anywhere in the book. By trying to give meaning to the notation, he slipped up on other exact information. The function value is *not* the distance covered by the cyclist, but is the distance from the starting point). One of the problems seemed to be that to him the place mentioned at the beginning of the question and the starting point mentioned later on were not the same place. He also had problems understanding Question 6.4. He reread the question and then said, *"After how many hours ... each return to the starting point... so ... dammit... uhm... when do they return to the starting point?"* Then he reread the relevant question three more times. He could not visualize the situation. He had trouble knowing what the track looked like. If the track is curved, what point is the turning-around point? Then he continued, *"Shit!.. Ugh!.. I'm not sure...I think ... maybe...No... They return to the starting point?... How do I know when they turn around?...It will be at half the distance covered"*. He started to divide the "distance formula" by 2. This did not lead to anything because ***t*** still remained unknown. Moreover, the same misconception regarding the position-time function picked up in the beginning of the question prevailed.

An analysis of another talk-aloud on Question 6 proves that a student can perform quite well mathematically in spite of visualization problems. The comprehension difficulty can be cancelled by subject proficiency. Question 6 has enough mathematical cues that signal what calculation is needed in most cases. An E1 student achieved 90% for Question 6 without really understanding the situation. A few of his remarks are as follows, *"I'm not quite sure what's going on in this question...I don't know what's going on here"*. After reading Question 6.4 he once again says, *What's going on here?...They must be going in circles.* He nevertheless finds the times when the speed is zero, but the fact that he gets two answers for the first cyclist confuses him more. He says, *"I don't know what's cracking here"*. He could not interpret his answers. After doing the question he wrote: *"I wasn't sure what my answers meant...that is how to interpret my answer"*.

A remark from a student trying to do Question 2.4 shows that she could not form an image of the movement to which the question was referring. *"How can a particle move in a straight line and pass the same fixed point a few times?"* she asked. She was unable to answer the question.

### *4.6.5 Non-verbal factors*

This category refers to the use of a letter symbol in such a way that it interferes with the processing of information. This can happen in various ways:

#### 4.6.5.1 Inappropriate use of the letter symbol

In Question 1 for example, X and Y are used as the names of pumps i.e. the letter symbols refer to objects. Usually x and y are used to indicate variable quantities. In this particular question it caused many problems because X and Y were used in the same sentence in which the numbers of litres of water were given. In a sense, this suggested that the number of litres should be taken as x and y and resulted in misleading information. In Question 1 the variables are not explicitly given and the aim of the linear programming problem is only mentioned in the last part of the metatext. Most students began to mathematize as they started reading and used x and y for litres of water as suggested. This misleading non-verbal factor caused serious comprehension problems. Most of the students started off with the wrong inequality, $x + y \geq 60\,000$. After quite a bit of struggle and loss of time, some managed to link the variable to the number of hours. One E3 student experienced the suggestion of x and y for litres of water so strongly that he said he had problems because x and y referred to hours but also to litres of water. If the time had been limited, most students would not have been able to solve this problem. During the personal conversations, students expressed their discontent with this unfamiliar, misleading use of the letter symbol. The textbook used by most students who did the protocol experiment does not use the letter symbol in this way (De Jager et al. 1985).

#### 4.6.5.2 Ambiguous use of the letter symbol

In Question 2 the letter symbol 0 is used to indicate a point as well as the number, zero. Students mentioned this as an unnecessary impediment. Some of them understood the point 0 to be the point, zero. To them that meant the starting point and where the speed is subsequently zero.

A similar problem was experienced with Question 3. The last phrase uses P for selling price, but the immediate question after this information is about profit. First language students who read quickly were inclined to link P to profit and not to selling price. So according to them, the profit function was given and they did not need the cost price. One girl said, "*I got confused with the cost - I wasn't really sure if I needed it.*" She did the sum by using the "profit" function, P = 400 - 0,02n correctly in all respects, except of course it was not the profit function. Needless to say she refrained from using P in her adaptation of the problem. More than one student was confused in this way.

4.6.5.3 Entangled verbal / non-verbal information

This could interfere with reading rhythm or comprehension ease, especially if students are non-native readers of mathematics (cf. Pimm, 1987:1). Most students fall into this category. An example of a verbal problem being communicated in this type of language is Question 3. One of the sentences reads, *The cost of n calculators is C = 100n - 200 and n calculators can be sold at a price P = 400 - 0,02n per calculator*. The phrase, *n calculators can be sold at a price P = 400 - 0,02n per calculator*, was very confusing and did not succeed in communicating the writer's intended message. The two phrases, *n calculators* and *per calculator* in one sentence did not seem to match. Few pupils who know the concept of profit have the experience of a selling price changing with every *one* article being sold. Although the information, *per calculator* was important, the entangled verbal/non-verbal information as well as the inferior position of this data had the effect of hiding the crucial detail. Some students did not even read the words, *per calculator*. Most of the other students who did read these words did not register that this was an important cue. As mentioned in Chapter 2, research by Huckin (1983; 1987) suggests that important information be put in prominent positions to aid successful communication.

Entangled verbal/non-verbal language was a readability problem specifically mentioned by the students during the personal interviews (cf.4.4). Most of them experienced this type of language as unnatural or non-real. Referring to questions like Question 3, one of them said, "*If you want to give us real-life problems, why don't you give them to us in real-life language?*" According to them the "non-real" language is more difficult to read and takes longer to understand, which implies that it interferes with reading and comprehension ease. This readability problem links up with what Bachman (1992) calls the need for *authentic* language during testing. Referring to language tests, he points out that for tests to be authentic, they should put the same requirements on test takers as does the language in non-test situations. Drawing a parallel to mathematics testing, one could say that the language of the described context should be the same authentic language as that used in non-test situations.

When studying the adaptations it was clear that the students preferred the real-life situation to be described in verbal language as far as possible without the interference of non-verbal symbols or mathematical models. They preferred a language that could be linked to the language of their everyday world. During the personal conversations, they emphasized the fact that *real-life contexts should be communicated in real-life language*. They were convinced that if they could receive the message in more natural language, it would be easier to read *and* understand. Furthermore they said it would be easier to relate the verbal information to the non-verbal function if the two types of information were more distinctly separated.

When reading the original questions once more, one realized that many of the verbal problems were already semi-mathematized. This was especially true of Questions 1, 5 and 9. The examiner was communicating real-life information in a type of short-hand language which was quite natural to him, but difficult for the students to process. Although one would think that this semi-mathematized form would help students to move more easily to the fully mathematized form, this aid seems to interfere with comprehension. Some students stumbled over the non-verbal data and even skipped this information when reading some real-life "story-sums" for the first time.

The issue of non-verbal information can prove difficult for writers of mathematics text. Students who are specialist readers of mathematics could prefer data presented in semi-mathematized form. These readers normally have no problems extracting concrete information from a mixture of verbal and non-verbal text. They could even find too much verbal information irritating. At school level however, most students are not specialist readers yet - they do not read or speak mathematics like a "native" (cf. Pimm, 1987:1). Only one student doing the talk-aloud showed preference for semi-mathematized text. This student was an exceptionally high achiever in mathematics. To him it was quite natural to speak mathematically about real-life situations. Most of his peers did not share this experience.

The above discussion of the five main types of readability problems is by no means exhaustive. It is nevertheless sufficient to indicate that the analysis of protocols generated important evidence regarding comprehension difficulties. This information as well as the personal interviews with the students confirm the need for more readable text.

## 4.7 Readability and achievement - a few general remarks

Every question solved by the students during the protocol experiment was marked. When considering the different achievement levels, a few interesting issues emerged. The following table represents the average performance per question per language group. The last two columns contain the average achievement for all nine questions and the students' average school performance in mathematics. There were six students per language group although only three students per language group did a talk-aloud on each question.

**Table 4.6 Percentage achievement per question per language group**

| Questions: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A.9 | A.S |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Language group: E1 | *41* | *43* | *22* | *14* | *67* | *70* | *44* | *67* | *30* | *44* | *76* |
| Language group: E2 | *57* | *76* | *44* | *89* | *81* | *65* | *44* | *73* | *52* | *68,5* | *83* |
| Language group: E3 | *48* | *42* | *44* | *8* | *68* | *63* | *55* | *60* | *52* | *44* | *68* |

**A.9:** Average percentage for the nine questions
**A.S:** Average school performance in mathematics of students doing the talk-alouds

Test scores referring to the nine questions (A.9) are markedly lower than the average school performance of the students (A.S). This confirms the well-known fact that most students find verbal problems in mathematics exceptionally difficult. This difficulty is often increased by unnecessary readability problems as illustrated by the recorded talk-alouds. The reason for the relatively high performance of the E2 group compared to the other two language groups is not quite clear. When listening to the talk-alouds it seemed as if the reading rate of the E2 group was more conducive to good comprehension than that of the other two groups. E1 students read too quickly to pick up important detail whereas E3 students read too slowly and seemed to lose the global message of the verbal problem.

When comparing the achievement levels of the different questions, one should keep in mind that only Questions 1, 4, 5 and 9 necessitated mathematization. This mathematical skill is usually considered to be one of the main reasons for low scores on verbal problems. However, the table reveals that except for Question 4, these questions were not experienced as the most difficult. For Questions 2, 3, 6, 7 and 8 the necessary functions were virtually all given. In spite of this, the average performance was weak compared to the students' school achievement. Although one should be careful not to conclude too much, one could say that factors other than mathematization seem to cause important problems as well. It could not have been a lack of motivation or even the absence of a positive attitude. There was evidence enough to confirm the opposite. Lack of mathematics proficiency was also not a major factor with these relatively bright students.

The analysis of protocols showed up two other factors. One of course was readability. In a sense this research confirms the research results by Cummins et al. (1988) that much of the difficulty students experience with word problems can be attributed to difficulties in comprehending abstract and ambiguous information. The other was the fact that few pupils attempted to generate additional information by interpreting the non-verbal functions. The few who did try interpretation were confronted with even more comprehension problems due to inappropriate functions (cf. 4.6.3.4). The lack of interpretation could also be due to a time problem. Students are so used to working against time in examination conditions that they do not allow themselves the time to develop this necessary skill. The evidence of the protocols shows that students can achieve quite well by executing calculations signalled by mathematical cues without really understanding the verbal problem. The analysis of protocols suggests that time-restricted tests as well as other readability problems tend to encourage rote learning.

Although there could be many reasons why students find verbal problems difficult to solve, the protocols have generated enough evidence to suggest that readability has an important influence on achievement.

### 4.8 Important issues leading to the main hypothesis

A few important issues emerging from the protocol experiment need to be reconsidered before formulating the hypothesis.

1. The analysis of protocols has revealed definite readability problems that would not have been detected otherwise. When the nine verbal problems were initially selected, one had no idea that so many different types of readability problems were latently present in the texts. The identified problems emphasized the unequal partnership between the interlocutors, the examiner and examinee.

2. Although much verbalization was done, it was clear that not all readability problems were verbalized. Not all second language readers found it easy to talk about the problems they were experiencing and during the personal interviews readability problems were raised that were not mentioned in the talk-alouds. However, enough data is available to form an idea of the range of the readability problem.

3. The results of readability research reported in Chapters 2 and 3 were to a large extent confirmed by the talk-alouds (cf. Table 5.1). This answers the important research question whether readability problems experienced in ordinary common language also affect the readability of mathematics text.

4. The results of the talk-alouds not only compared well with readability factors reported in the literature study, but generated additional factors that are specific to mathematics text. Many of these factors refer to non-verbal features (cf. 4.6.5 and 6.3).

5. A reduction of certain readability problems will sometimes result in the reduction of the mathematical demand (cf. 4.6.2.1). In cases like these one would carefully have to reconsider the aim of the question. On the one hand one should be careful to maintain the mathematical standard. On the other hand, a readability problem should not be allowed to jeopardize the primary aim of a specific examination question.

6. Not all readability problems influence comprehension. The talk-alouds provide enough evidence that readers at times knew enough about a subject to comfortably by-pass possible readability problems. However, Steffensen (1987) argues that an improvement in readability always influences reading ease and reading time.

7. Not all the readability problems that did influence the comprehension of verbal problems influenced achievement. Readers were often able to overcome the comprehension difficulties they had verbalized and successfully proceed to solve the problem. Had it not been for the evidence of the talk-alouds, the high achievement level of some students could make one believe that there were no comprehension difficulties at all.

8. The fact that the talk-alouds were not time-limited levelled down the influence of readability problems on achievement. Students had enough time to sort out many of their difficulties. If however the think-alouds had been time-limited, the situation might have been different. For example, two E2 students gained 81% and 90% respectively for Question 1. Had the time been limited to the 16 minutes normally allocated for a question like this, they would have lost virtually all their marks. The research by Arter & Clinton (1974) as well as that by Lepik (1990) points out that certain readability factors play no significant role in the solution of word problems, *unless* time is limited.

9. The question remains whether improved readability will have any *significant* effect on achievement. It is obvious that if improvement in readability does not always improve comprehension, it also will not always improve achievement. On the other hand, there is evidence from the talk-alouds that students could possibly have achieved better had they understood the verbal problem better.

10. Readability is not the only variable that could have an influence on achievement. Another influencing factor is the difficulty level of the question which again depends to a large degree on the subject proficiency of the student. Ultimately achievement is influenced by the interaction of a variety of factors, readability being only one.

However, the data generated by the talk-alouds provides enough evidence to believe that readability is one of the important keys to students' performance on verbal problems. In this respect the protocol study has confirmed the important role of reading and understanding in the whole problem solving process (cf. 1.5).

Years of teaching experience have shown that in most cases students are able to execute the non-verbal operations underlying verbal problems. However, as soon as they are expected to apply their mathematical knowledge in a real-life situation their performance decreases dramatically. After considering the data of the protocols and the influence readability could possibly have on achievement, the following main hypothesis has been generated:

*Improved readability of the common language used in verbal problems in mathematics examinations will improve achievement.*

This hypothesis will be tested in the next chapter by using two versions of the same nine questions answered by students in the talk-aloud experiment. The first version is the original version and in the second version the common language of the questions has been changed to a more readable form mainly by using the results of the protocol study. These results compare well with those generated by the literature study.

# CHAPTER FIVE
# ADAPTING THE QUESTIONS AND TESTING THE HYPOTHESIS

## Rationale

*To test the hypothesis that improved readability will improve test scores, the readability of the nine questions used in the protocol study had to be improved. The same students who answered the nine questions in the protocol experiment were asked to adapt the questions to a more readable form.*

*The adaptations suggested by the students as well as the guidelines from readability research were used to improve the readability of the original nine questions. Test papers were compiled by using the nine questions in three different versions: original, adapted and non-verbal. Test scores of the original and adapted verbal versions were used to test the hypothesis whereas test scores of the non-verbal versions were used to determine in what way the verbal form of the question prevented students from demonstrating other mathematical abilities. The original and adapted verbal versions of each question were mathematically the same, but differed in readability, the adapted version representing the version with improved readability.*

*Test scores of 108 selected students were used to test the hypothesis. Three different language groups were represented in this sample, i.e. students with English, Afrikaans or an African language as their mother tongue. To the Afrikaans and African groups, English is a second language. Tests were not time-limited to ensure that all questions were attempted by all students. The statistical significance of the differences between the test scores for the original and adapted versions of each question was assessed by using the non-parametric Mann-Whitney U-test. The hypothesis was tested for the combined language group as well as for each language group separately and was confirmed in a number of important cases.*

*Results emerging from the statistical analysis point to the importance of the study from a quantitative point of view. Simultaneously, the results guide interpretations regarding the* *practical significance of improved readability during assessment and teaching practice. Results are therefore reported both quantitatively and qualitatively.*

**Chapter contents**

## 5.1 The adaptations

After the students had completed the talk-aloud experiment, they were asked to adapt the questions for better comprehension. None of their answers were to be changed, only the formulation of the original questions. The aim of the adaptations had been discussed with the students before, but written instructions were nevertheless given to each of them to remind them of the discussion (see instructions for adaptations, Appendix B). Initially there was a surmise that the entangled verbal/non-verbal information would be the main obstacle to comprehension ease. That is why the instruction sheet refers to the adapted verbal version as *an extreme verbal form*. When giving students the instructions, they were once again reminded of the aim of the adapted verbal version: to make the original question more comprehensible. They were only to adapt the questions to a more verbal form if they were convinced it would be more student-friendly. Care had to be taken not to change the mathematical demand.

All nine original questions were to be changed into two versions, an adapted, more readable verbal form and a non-verbal version. Adaptations were to be completed within two days to make sure students would not forget the comprehension difficulties they had experienced during the talk-alouds. The adapted, more verbal forms of the questions were analysed with a view to establishing what they considered as more comprehensible text. Although students also presented a non-verbal version, this version was not analyzed for readability purposes as the hypothesis was directed at the two verbal versions.

The analysis of the adapted verbal versions was done directly after the analysis of the protocols. Doing these two operations more or less simultaneously facilitated direct comparison between students' adaptations and their comprehension problems. There was a constant cross-reference between suggested adaptations and identified comprehension problems (cf. 4.4). Eventually, the readability of the questions was improved by using the suggested adaptations of the students as well as the guidelines generated by readability research (cf. 2.5 and 3.4 ). Test papers for the composite test contained all nine questions in three different versions.

### *5.1.1 The three versions: original, adapted and non-verbal*

The ***original version*** of each verbal problem was the version that was used in a previous, formal examination and was the same as that used in the protocol experiment.

The ***adapted version*** was the same verbal problem as the original version, but with improved readability. As mentioned, it was originally surmised that a main cause of readability problems in "story sums" (as students call them) was the entangled verbal/non-verbal information. This type of language is not the language in which students usually tell their "stories". When considering the adaptations of the students, the following became clear:

- They changed much more than the entangled verbal/non-verbal information.

- Their adaptations addressed the main readability problems they had experienced during the talk-alouds. This became very apparent when comparing their adaptations with their comprehension problems and vice versa. For example, in Question 1, the talk-alouds indicated that readability problems were caused by three main factors: the inappropriate use of the letter symbols, X and Y; unstructured information, and ambiguity. All three these factors were addressed in the students' adaptations. This trend was true for all nine questions.

- Most of the students' recommendations compared well with the guidelines generated by the readability research reported in Chapters 2 and 3. This is a salient characteristic of the adaptations. When comparing the guidelines reported in 2.5 and 3.4 with the students' suggestions discussed in 5.1.2, it is remarkable how many aspects correspond and that no contradiction was reported. However, not all readability factors reported in literature were identified for the obvious reason that the nine questions did not contain all possible readability problems. Furthermore, students were asked not to change the mathematical demand of questions. This was necessary to be able to compare test scores of the original and adapted versions. Certain readability guidelines reported in literature, like the use of the metatext to elucidate the text, could possibly have changed the mathematical demand of the question.

The following table is an indication of how the readability guidelines reported in literature as well as those identified by the preliminary analysis (cf. 1.8) were confirmed by the students' talk-alouds or adaptations. By confirmation of readability guidelines is meant that there was a general tendency amongst the students to point out the same readability factors reported in literature or identified in the preliminary analysis.

**Table 5.1 Confirmation of readability guidelines by students' talk-alouds or adaptations**

| Guidelines | Guidelines reported in literature or the preliminary analysis | Guidelines confirmed by talk-alouds or adaptations |
|---|---|---|
| *re:Vocabulary*<br>1 | Use high frequency words | Yes |
| 2 | Use the same words consistently | Yes |
| 3 | Be careful not to pre-suppose familiarity of words | Yes |
| 4 | Use true cognates. Avoid misleading ones | No |
| *re:Syntax*<br>1 | Avoid unnecessary parentheses | Yes |
| 2 | Avoid hypothetical situations and constructions | Yes |
| 3 | Use interrogatives instead of sentence-completion sentences in multiple-choice questions | No |
| 4 | Curb proposition density per sentence | Yes |
| *re:Discourse*<br>1 | Use metatext to elucidate text | No |
| 2 | Be careful not to presuppose background knowledge of contexts | Yes |
| 3 | Contexts as well as behaviour within contexts must be familiar | Yes |
| 4 | Format the text according to the anticipated reading style | Yes |
| 5 | Use topic sentences | Yes |
| 6 | Put important information in important positions | Yes |
| 7 | Repeat important information | Yes |
| 8 | Curb proposition density in paragraphs, across sentences | Yes |
| 9 | Obey the principles of the given-new contract | Yes |
| 10 | Use a separate sentence for every question. Place information before question | Yes |
| 11 | Structure of information must support the construction of a CKS | Yes |
| 12 | Allow enough reading time | Yes |
| 13 | Use pictorials to elucidate the text | Yes |
| 14 | Avoid obscure information | Yes |
| *Non-verbal factors*<br>1 | Use appropriate functions | Yes |
| 2 | If given functions have a restricted domain, mention the restrictions | Yes |
| 3 | Avoid the ambiguous use of letter symbols | Yes |

Other factors that became clear when considering the students' adaptations were:

- E1 students gave the *most* and *best* adaptations. This came as a surprise. One would have thought that second language students who had experienced many comprehension problems during the talk-aloud experiment would address more problems more successfully. This was not the case. Some E2 and E3 readers did not make any adaptations. A second language student wrote the following which helps one to understand the situation: "*...I think if I had to change the questions I would just confuse a student or myself. I attempted to change the questions, but after I read them they didn't make any sense or rather they made the question very complicated...*" Other second language readers did adapt, but the language of their adaptations was often oversimplified, awkward or clumsy. It seems that E1 readers, because of their superior control of English, were able to improve the readability without using awkward or oversimplified language. To them the more informal, real-life English came naturally. It was the language in which they experienced the real-life problems described in the given questions.

The fact that second language readers did not give as good and as many adaptations as their first language counterparts reminds one of Clarke's short-circuit hypothesis (1980) which implies that the transfer of first language reading strategies to a second language is dependent on second language proficiency (cf. 3.2.2). Studies by Cummins & Swain (1986; cf. 3.1) have shown that although second language learners may have basic interpersonal communication skills (BICS) they often lack the necessary cognitive academic language proficiency (CALP).

- Although the attempts of the second language readers were often less successful, the analysis of their adaptations revealed very much the same suggestions as those of first language readers. A marked difference was the definite call of second language readers for simpler language and more explicit and extensive information. In this regard their adaptations were exceptionally important.

- Some students did not realize they were having comprehension problems during the talk-aloud experiment, so they did not address the specific readability problem in their adaptations. For example, an E2 student who had a complete breakdown in Question 1 because of the non-verbal factor, X and Y, did not address this problem in his adaptation. To him X and Y remained objects, the pumps in this case. He could not escape this. A sentence in his adaptation reads, *Running costs are 90 and 60 cents per hour for X and Y respectively.* The adaptations proved that the inappropriate use of letter symbols could reinforce existing misconceptions (cf. Hart, 1981).

- Sometimes the adaptation was a simplified version of the student's miscomprehension, proving that he still did not understand the question correctly. The position-function in Question 6 was frequently misunderstood as referring to the distance travelled. Often this misconception was explicitly described in the adaptations. For example, one adaptation reads: *...The distance (S) that both of them have travelled from the starting point after a certain number of hours (t) is given by*

$$S \text{ of cyclist } A \text{ is: } = t^3 - 12t^2 + 36t$$
$$S \text{ of cyclist } B \text{ is: } = 24t - 4t^2$$

- Students found some questions virtually impossible to adapt to a more readable form because they found it difficult to understand what the "story" was all about. Examples of this dilemma are Questions 2, 5, 6 and 8. Referring to Question 8.1, one E1 student wrote: "*I don't understand the question or given information so it is difficult to change.*" She did not adapt Question 8.3 either, but instead wrote: "*The whole question is confusing, therefore it is difficult to change.*" During the personal interviews, some suggested that in Questions 2 and 6 a graph be introduced at a much earlier stage to make the information clearer. Although this could have helped a great deal, the graph would not just have changed the readability but also the mathematical composition of the question. To be able to compare the achievement levels of the adapted and original versions, the mathematical demand of the problems had to be kept the same as far as possible. This pre-requisite often made it almost impossible to change the readability of a question effectively.

- Even though the performance of the E3 group on the HSRC's English and Mathematics proficiency tests was relatively high, they were the least able to adapt the questions for improved readability. It seems that another skill, other than that tested by the HSRC test, is needed when adapting for improved readability. An E3 reader wrote, "*I don't really know what to change, but the information is all mixed up - this tends to be confusing.*" When changing a piece of discourse into a more readable form, one must not only *understand* the original message correctly, but also be able to *manipulate* the language like a mother tongue user.

The ***non-verbal version*** contained as little verbal language as possible (see Appendix C for the nine non-verbal versions). The non-verbal version was not used for testing the hypothesis, but was introduced into the experiment and included in the composite test to ascertain what the achievement level of the students would be if the mathematics were *not* tested in a real-life context. Of course, this version did not have the same mathematical demand as the original and adapted verbal versions. For example, the non-verbal versions in Questions 1, 4, 5 and 9 did not necessitate the translation of information to equations or inequalities because they were already included. Although mathematization is an important and often difficult skill for students to master, in actual examinations, most marks for verbal problems are allocated to the execution of calculations. So the aim of including the non-verbal version in the composite test was to ascertain to what degree the verbal part of a problem and the mathematization requirements prevented students from demonstrating other mathematical abilities.

An analysis of the ***adapted*** verbal versions revealed various types of suggestions.

### *5.1.2 General discussion of students' adaptations*

Students' suggestions can be divided into four categories: lexical, syntactical, discourse and non-verbal features. Single examples will be used to demonstrate the different features, but the recommendations will be exemplified in more detail when the nine adapted questions are discussed one by one in 5.1.3.

#### ***Lexical features***

In this category students suggested that:

1. The same word should be used consistently to refer to a concept, either to prevent uncertainty or to emphasize important information.

In Question 5 the word *inequalities* is used in 5.1 and 5.2 whereas the original version first refers to *constraints* and then to *inequalities*.

2. More natural or authentic vocabulary should be used under test conditions, that is, language that is more in line with the type of language students use in non-test situations (cf. Bachman, 1992; 4.6.3.1). This refers to words or phrases.

The words *simultaneously* and *respectively* have been omitted in Question 6. Very often *per* week or *per* calculator was replaced by, *one* week or *one* calculator.

Not all suggestions were always acceptable. For example, students should be taught the importance of cue words like *per*. However, this specific suggestion points to a more general appeal to communicate real-life situations in real-life language.

3. Words or phrases that could cause confusion or ambiguity should be avoided (cf. Akmajian et al. 1984).

In Question 5 there was confusion because the text referred to *each week* while Question 5.6 referred to *daily*. Students changed *daily* to *weekly* so that text and meta-text were the same.

***Syntactical features***

These suggestions addressed the structure of sentences. There were definite recommendations to:

1. Reduce the proposition density of sentences (cf. Kintsch et al. 1973; 1975; 2.4.3.4)

In the original version of Question 3, the information of cost and selling price is given in one sentence. Students used at least two sentences in their adaptations.

2. The mathematical model should not be verbalized when communicating the corresponding real-life behaviour.

In Question 1, students kept on re-reading, "*The sum of the number of hours that the pumps operate per day must not exceed 20*", which is a typical verbalization of the mathematical model. The following is an example of a typical adaptation: *Together these pumps may not run more than 20 hours a day.* This adaptation clearly represents a more natural sentence structure.

3. Use more "action-specific" language. This suggestion was more than just a recommendation to use the active voice. It reminded one of the "scenario principle" described by Flower and her colleagues (1983). Although the sentences were not structured around a person performing certain actions (cf. 3.2.3), the information communicated more of a real-life scenario.

For example, instead of writing *Two pumps, X and Y are used to supply a minimum of 60 000 litres of water per day to a reservoir*, a student wrote, *Two pumps supply water to a reservoir. These pumps must pump at least 60 000 litres of water to the reservoir together every day.* This suggestion relates to the use of common or authentic language. Often the adaptation to more action-specific language was simultaneously a switch to more common or authentic language. In the same sense a change to more common language was often a change to more action-specific language use.

The above example also serves to illustrate how many different kinds of readability factors were often addressed simultaneously. The last suggestion includes:

- an introductory topic sentence
- more active language
- more authentic language
- a reduction of proposition density per sentence
- a removal of the non-verbal interference, X and Y

- more explicit communication of information

This example points to the interactive nature of readability factors. Often, when addressing one factor, other factors are addressed as well.

4. Structure sentences in the same order as necessary for the construction of a contingent knowledge structure or CKS (cf. Bobrow & Brown; 1975; 3.2.3)

This type of adaptation relates to the structure of discourse and will therefore be discussed in more detail in the next paragraph.

*Discourse features*

The overall structure of discourse was an important area addressed by students and confirms the results reported in research (cf. Selzer et al., 1983; 2.4.3.4). Like the structure of sentences, the structure of discourse relates to the manner in which the discourse contributes to the construction of a CKS or adheres to the given-new contract (cf. 2.4.3.4).

To change unstructured discourse in such a way that the information becomes more comprehensible often results in a structure that is also easier to mathematize. Question 5 serves as an example. In this question a writer could structure part of the information in two distinct ways. He could either explain what happens to *one product* in *different departments* or he could explain what happens to *different products* in *one department*. In this question the restriction is within the departments so the latter explanation would aid comprehension. This is probably one of the reasons why students suggested this type of adaptation. However, as has been indicated in 4.6.2.1, this information structure also aids mathematization. One could ask whether this type of adaptation is justifiable, especially if taking into account the importance of teaching students to translate unstructured information into mathematical language. To answer this question, one would have to consider the aim of the question. If the main objective is to assess whether a student can translate a practical problem into a mathematical model then the adaptation can not be justified. If, however, the student is also expected to demonstrate his ability to use the linear programming algorithm, an adaptation as suggested by the students is permitted, especially if the task is time-limited. Difficult structure should not prevent a student from solving the rest of the problem.

The question referring to whether a suggested adaptation should be accepted or not also applies to an issue like redundant information. Is it acceptable to remove redundant information to make the text more comprehensible? One could argue that students should be able to extract essential components from a muddle of quantitative facts. If, however, questions like these are asked, similar problems should have been practised in a classroom situation. If not, redundant information in examination conditions can be considered to be unfair and should be removed.

Therefore, to decide whether certain adaptations were justifiable or not the following aspects were taken into consideration:

- the aim of the question
- the time allowed for the execution of the solution
- teaching practice as inferred from relevant textbooks as well as personal experience.

For example, if the primary aim of a question was not to disentangle unstructured information, the adaptation that addressed this problem was accepted as justifiable. (The structure of the question and the marks allocated for the different parts of the metatext gave a reasonably good indication of what goals the examiner had in mind when he set the question).

The decision to accept or reject a recommended adaptation was also guided by the issue of reliability and validity. One should be careful that factors other than those one wants to measure do not affect the reliability of the measuring instrument, here the specific mathematics question (cf. Bachman, 1990:160). Referring to teaching practice, one could say that the fairness of an examination question can be judged by considering the degree of congruency between teaching and examination practice (cf. 6.4.1).

Regarding ***discourse*** features, students' adaptations included the following suggestions:

1. Reduce the proposition density per paragraph.

This recommendation is an extension of the one referring to sentences. Density of information can vary per sentence, but also per paragraph i.e. across sentences. It is clear that although the proposition density of individual sentences may be low, the total information density per paragraph, contributed by single sentences, may be high.

In the personal interviews some students spoke of *reading space*. Separating the information into paragraphs helps to make the text more accessible. As soon as one begins to use paragraphs, the information becomes more structured or cohesive. For example, in the adapted version of Question 1, the three paragraphs not only give the necessary reading space, but communicate the situation in three distinct packages of information: the first refers to the pumps, the second to the litres of water and the third to the number of hours.

With this suggestion students follow the same pattern usually followed in ordinary expository prose. Written communication uses paragraphs to suggest a degree of cohesion between units of information - it groups together what belongs together (Kaplan, 1980:411).

2. Use a topic sentence as an initial sentence for introducing the context (cf. Faigley & Witte, 1983; 2.4.3.4). In this way readers are better able to identify the main idea of the passage.

For example, it was suggested that Question 5 start with a short sentence to introduce the context. *A factory produces two products, A and B.* In this case the use of a topic sentence automatically reduces the proposition density of the first sentence.

3. Remove redundant information.

Only Question 3 had this type of information. South African textbooks contain few, if any exercises that require students to cope with redundant information, so it was clear why students felt information of this kind should be removed.

4. Relieve visualization or imagery problems.

One way of doing this is by using a sketch or graph. A sketch was suggested for Question 9.

The imagery problems caused by inappropriate functions could not be addressed because the students did not know what alternative function to use.

5. Use more explicit and extensive information.

This recommendation also addressed visualization or imagery problems. It was especially E3 readers who advocated this suggestion. Due to language difficulties or questions that were culturally biased, they needed more explicit and/or extensive information to properly understand the examiner's intention. This is quite natural. The less one knows of a situation the more exact and the more extensive the information must be. Schemata of less familiar situations are not as well developed and cannot automatically induce the necessary inferences (cf. 2.3.1). The activation of these schemata is even more difficult if it has to be done in a second language (cf. 3.2.1). To illustrate the need for more explicit and extensive information, an E3 reader's adapted version of Question 9 is presented in full. Parts of the explicit information are highlighted.

> *2 straight roads intersect at right angles at* ***a point*** *O. P is a point on* ***one of the roads*** *so that* ***the distance of*** *OP = 10 km. 2 people -* ***one at P and the other at O****, start to walk at the same time,* ***the person*** *P walking* ***towards*** *O on one road while* ***the person*** *at O walks on a* ***different*** *road away from* ***point*** *O.* ***The person*** *at P walks* ***with a speed*** *of 3 km/h while the person at O walks* ***with a speed*** *of 4 km/h. After* ***a time*** *of t hours, they reach positions A and B on* ***these*** *two* ***different*** *roads.*

- *Find an expression for* ***the distance*** *AB*
- *After what length of time are the 2 persons closest*
- *What is the shortest distance of AB*

The structure of the above paragraph reminds one of the cultural thought pattern described by Kaplan as circulatory, only here the information is not indirect but very direct (cf. 3.3.1). More extensive information is communicated by the repetition of explicit information. For example, consider the explicit repetition of important aspects like, *a point O. P is a point... one at P and the other at O...the person at O... the person at P.*

*Non-verbal features*

These readability factors are specific to mathematics text and were only identified by the analysis of the students' adaptations. The literature study of Chapters 2 and 3 had few suggestions in this regard because it was focused more on the readability of ordinary English. Students suggested the following:

1. Entangled verbal and non-verbal information should be separated.

This means that the narrative part of the text should be communicated, as far as possible, in verbal language only. Not only should the mathematical model be separated from the narrative part of the text, but the information should not be communicated in a semi-mathematized form as in Question 7.

In a sense this also addresses a syntactical and discourse feature, because the removal of non-verbal information relieves the proposition density per sentence and/or per paragraph.

In Question 3 a sentence like

*The cost of producing n calculators is* $C = 100n + 200$ *and ...*

is replaced by,

*The total cost to make these calculators is given by,*

$T(x) = 100x - 200$

*T(x) is the total cost in rands.*

*x is the number of calculators.*

2. Non-verbal factors that give rise to confusing or misleading information should be avoided.

The letter symbol, P, in Question 3 confused a number of students and was associated with profit. They all omitted this symbol in their adapted versions.

3. Identify the variables if this is not the primary aim of the question.

In Question 1 students correctly remarked that once the variables were identified the verbal problem was more comprehensible (cf. discussion of Question 1 in 5.1.3).

4. Use function notation wherever applicable.

Students felt quite strongly on this issue. Questions 2, 3, 6 and 7 were adapted accordingly.

More readability problems relating to non-verbal features were generated by the analysis of protocols, but they were not directly addressed by the students' adaptations. An important problem was the use of inappropriate mathematical models (cf. 4.6.3.4). These additional non-verbal issues are included in the readability checklist presented in 6.3.

The manner in which the above suggestions were incorporated, in order to adapt the nine original verbal problems to a more readable form, is discussed in the following section.

### *5.1.3 Detailed discussion of the nine adapted questions*

To be able to discuss the adaptations more meaningfully, the original and adapted verbal versions will be presented alongside each other. By comparing the two versions, one is better able to understand the readability problems as well as the corresponding adaptations. Since the hypothesis was eventually tested by using the test scores of the original and adapted versions only, the non-verbal version is not presented here. It has been included in Appendix C.

The questions consist of *text* (the information) and *metatext* (the questions set on the information). Every verbal problem will be considered on its own, but not all adaptations will be discussed in detail.

Unfortunately, the talk-alouds and adaptations of the E3 readers could not be considered when preparing the final adaptations. By the time the composite test had to be completed, very few protocols (or the corresponding adaptations) from E3 students were available because of political unrest in schools. It was also not certain *when* this data would be ready, so the original questions were adapted using mainly the suggestions of E1 and E2 readers.

Finally, the adaptations were guided by three main sources of information:

1. Readability research reported in Chapters 2 and 3.
2. The results from the protocol analysis (cf. Chapter 4).
3. The students' adaptations (cf. 5.1.2).

The sources were by no means contradictory. On the contrary, results generated by one source were confirmed by the next (cf. Table 5.1).

When discussing the adaptations, no judgement is passed regarding other qualities of the verbal problems. More readable does not necessarily mean better questions. Questions may be criticized on other grounds, but that is beyond the range of this study. The adaptations address only readability.

The adaptations are by no means perfect. When considering the adaptations at a later stage, one realizes that readability could be improved even more. This emphasizes the fact that good writing is an iterative process and not the product of a few single attempts (Ulijn & Strother, 1994). After the data from the E3 students became available, one realized anew that in certain areas the needs of these readers are often different to those of E1 and E2 readers. In particular, the need for more explicit and extensive information became obvious. However, the following adaptations are the ones that were used to test the hypothesis.

| **Original version: Question 1** | **Adapted version: Question 1** |
|---|---|
| Two pumps X and Y are used to supply a minimum of 60 000 litres of water per day to a reservoir. X pumps 5 000 litres per hour and the running costs are 90 cents per hour. Y pumps 2 000 litres per hour and the running costs are 60 cents per hour. The sum of the number of hours that the pumps operate per day must not exceed 20. The number of hours that X operates per day may not be more than twice that of Y.<br><br>1.1 Represent the above information as a system of inequalities.<br><br>1.2 Use graph paper to represent these inequalities graphically and clearly indicate the feasible region.<br><br>1.3 Use your graph and determine the number of hours that each pump will have to operate per day in order to minimize the total running costs. | Two pumps pump water to a dam. One is an electric pump. The other is a diesel one.<br><br>These pumps must pump at least 60 000 litres of water together every day. The electric pump can pump 5 000 litres in one hour. The diesel pump can pump 2 000 litres of water in one hour.<br><br>Together these pumps may not pump for more than 20 hours a day. Also, the electric pump may not pump for more than twice as many hours as the diesel pump.<br><br>Let the electric pump work for x hours a day.<br>Let the diesel pump work for y hours a day.<br><br>1.1 Write down inequalities to show the restrictions in the above information.<br><br>1.2.1 Represent these inequalities on graph paper.<br>1.2.2 Clearly indicate the feasible region with the letter F.<br><br>1.3 The running costs for the electric pump are 90 cents an hour. For the diesel pump the running costs are 60 cents an hour.<br><br>Use your graph to calculate how many hours each pump has to pump to minimize costs. |

In this question the main readability problem was the inappropriate use of the letter symbols X and Y. When comparing the two versions, one recognizes adaptations related to lexical, syntactical and discourse features as well as non-verbal factors.

***Lexical features***

- More common language is used. For example, *at least, every day* and *one hour* are substituted for words like *a minimum of, per day* and *per hour*. The paragraph starting with *Together these pumps...* replaces two sentences of rather unnatural language.

*Syntactical features*

- The proposition density of the original first sentence has been reduced.

- Apart from using more authentic language, the sentences of the paragraph starting with *Together these pumps...* use more action-specific language. The same applies to the very last sentence of the metatext.

- Sentences are structured to communicate like information and in this way aid the construction of a CKS. For example, all the sentences in the second paragraph communicate information about the restrictions on the litres of water. Although restructuring for better comprehension also makes mathematization easier in this case, it was decided to retain the adaptation because mathematization was not the final aim of Question 1.

- The information belonging to the objective function has been separated from other information and reserved for the last question of the metatext. This last adaptation addressed the ambiguity caused by the original Question 1.1 where *the above information* was not all part of the *system of inequalities*.

- Question 1.2 originally included two questions. These have been separated.

*Discourse features*

- The first sentence is also the topic sentence.

- The proposition density per paragraph has been reduced. The original text had all the information in one paragraph. The adapted version communicates this information in three paragraphs.

- The discourse is structured so that the different paragraphs report on different topics which once again aids the construction of a CKS.

- Visualization or imagery problems are relieved by explicitly mentioning the types of pumps. Diesel pumps normally cost less and the fact that one is electric could explain the difference between the 5000 and 2000 litres per hour. The somewhat strange behaviour described by the last two sentences of the original text could not be addressed.

***Non-verbal features***

- The letter symbols, X and Y, have been omitted.

- The variables have explicitly been identified. This adaptation also changes the mathematical demand of the problem. When students were told that this part of their adaptation would change the mathematical demand, they were adamant not to withdraw this suggestion. They said, *"That is true. By identifying the variables, you help the reader, comprehension-wise* ***and*** *mathematics-wise. But the amount you help him mathematics-wise is small compared to the big help you give, comprehension-wise. In any case, how many marks does the examiner award for the identification of the variables?"* The answer to the question was of course, *"None"*. But before incorporating this suggestion into the final adaptations, the advice of an applied mathematician was sought. His answer was that the students were right since the identification of variables is one of the difficult tasks of an applied mathematician and added, *"Once the variables have been identified you also* ***understand*** *the problem much better"*. Since the identification of variables was not the main objective of Question 1, the students' suggestion was accepted.

The next question is an example from kinetics theory.

| Original version: Question 2 | Adapted version: Question 2 |
|---|---|
| A particle moves along a straight line so that. t seconds after observations have commenced, its distance, s metres from a fixed point 0, is given by<br><br>$s = \frac{1}{2}t^3 - {}^{11}/_2 t^2 + 19t - 20\ (t > 0)$<br><br>CALCULATE:<br><br>2.1 how far the particle will be from the fixed point 0 the moment when observations commence<br><br>2.2 the values of t for which the particle is momentarily at rest<br><br>2.3 the speed of the particle at t = 2<br><br>2.4 when and how many times the particle will pass the fixed point 0<br><br>2.5 the distance covered by the particle during the fourth second. | An object moves up and down a straight line. On the line is a fixed point. The **distance** of the object **from the fixed point** is given by<br><br>$s(t) = \frac{1}{2}t^3 - {}^{11}/_2 t^2 + 19t - 20;\ (t \geq 0)$<br><br>s(t) is the distance in metres.<br>t is the time in seconds.<br><br>2.1 What is the distance of the object from the fixed point in the beginning?<br><br>2.2 After how many seconds will the speed of the object be zero?<br><br>2.3 What is the speed of the object after two seconds?<br><br>2.4.1 After how many seconds does the object pass the fixed point?<br>2.4.2 How many times does this happen?<br><br>2.5 What distance is covered by the object from the third to the fourth second? |

Although the given function in Question 2 is a position-time function (and not a distance-time function) it was decided to keep it as such as this is how some textbooks (cf. Laridon et al., 1987:245) and previous examination questions often refer to this type of function. The same kind of function is contained in Question 6. No example, either in a textbook or in a previous examination paper, could be found where the function was correctly referred to as a position-time function. It was argued that this was an example of a teaching problem. If the adapted version were to refer to a position-time function and the concepts of *distance, displacement* and *position* had not been taught thoroughly before, students would become even more confused.

In this question adaptations have been suggested on all three linguistic levels.

*Lexical features*

- On the whole, the vocabulary tends to be more authentic. For example, the word *object* replaces *particle* and Question 2.1 refers to *in the beginning* instead of *when observations commence.* Also the words *momentarily at rest* have been substituted by, *speed be zero.* Question 2.5 avoids the phrase *during the fourth second* and asks the question more explicitly. This last adaptation also removes the ambiguity experienced by some students. A textbook commonly used by many students (De Jager et al. 1985), never explains or uses a phrase like, *during the fourth second* and some students understood the question as asking the distance covered *after* four seconds. Luckily the object had not turned around in the given time interval. If it had, the question would have been more difficult to answer since the graph had not yet been drawn.

*Syntactical features*

- The proposition density per sentence has been reduced. In the original version all the information of the text is given in only one sentence. Three sentences have been used in the adapted version. In the metatext, Question 2.4 is communicated in two sentences instead of one, a sentence for every sub-question.

*Discourse features*

- The first sentence is a topic sentence giving more explicit information about the movement of the object and in this way relieving the visualization problem.

- The words *distance from a fixed point* have been formatted in bold letters in an attempt to clarify the message that the distance being referred to is *not* the distance travelled, but the distance from the point.

- In the adapted version a full sentence has been used for every question in the metatext. The original version uses one verb to serve all the questions. This is quite demanding on the short term memory of the reader and he will probably have to go back to the beginning of the metatext each time he wants to know exactly what is asked.

*Non-verbal features*

- A noticeable difference between the adapted and original versions is the separation of verbal and non-verbal information. All the non-verbal letter symbols have been removed in the first paragraph. The semi-mathematized information has been changed to common language. This increases reading ease and students experience the real-life situation in real-life language.

- The function notation s(t) replaces s. Students felt quite strongly about this adaptation. They argued that this notation was necessary to communicate the function concept that was operating within the given situation.

- The unnecessary interference of the non-verbal symbol, 0, has been omitted.

In the next question a story is told about the manufacturing and selling of calculators.

| **Original version: Question 3** | **Adapted version: Question 3** |
|---|---|
| A manufacturer has the capacity to produce 8 000 calculators of a certain type per week. The cost of producing n calculators is $C = 100n + 200$, and n calculators can be sold per week at a price $P = 400 - 0.02n$ per calculator.<br><br>Determine the value of n which maximises the profit. | A manufacturer makes calculators. The **total** cost to make these calculators is given by<br><br>$T(x) = 100x + 200$<br><br>T(x) is the total cost in rands.<br>x is the number of calculators made and sold.<br><br>The selling price of **one** calculator is given by<br><br>$S(x) = 400 - 0.02x$<br><br>S(x) is the selling price of one calculator in rands.<br><br>Determine how many calculators must be sold for a maximum profit.<br><br>**Hint:**<br>Total profit = Total selling price - Total cost price. |

Two major readability problems in this question were the redundant information in the very first sentence and the entangled verbal/ non-verbal information. When comparing the two versions, the following adaptations are noticeable:

*Lexical features*

- Overall the language is simpler. For example, the word *produced* is substituted by *makes* and the question of the metatext is posed in more common language and is more action-specific.

*Syntactical features*

- The proposition density of the second sentence has been reduced.

*Discourse features*

- The redundant information has been removed. Students had had no previous practice in dealing with this kind of information.

- The proposition density per paragraph has been reduced. In the original question the information is communicated by only one paragraph. In the adapted version the information on the cost and selling price are given in separate paragraphs.

- The discourse is structured in such a way that the facts are conveyed more explicitly. Bold letters have been used to emphasize the important fact that the first function refers to the cost of all the calculators while the second function is the selling price of only one calculator.

- The hint was inserted to help students who were not well acquainted with the concept of *profit*. Although the students of the talk-aloud experiment had no problems in this respect, there was evidence that some E3 students from rural areas needed help. In a sense this question was culturally biased because the concept of profit was not familiar to some African students. As the hint did not interfere with the primary aim of the question, the suggested adaptation was accepted.

*Non-verbal features*

- The non-verbal and verbal information have been disentangled
- Function notation has been used

Question 4 depicts a classroom situation.

| **Original version: Question 4** | **Adapted version: Question 4** |
|---|---|
| The number of students in a class is x and there are exactly enough books to supply each student with y books. If there had been 5 more students each would have received 2 books less and there would have been 6 books left over. If however there had been 3 students less each would have received 1 book more and there would have been 11 books left over.<br><br>How many students are there in the class and how many books are available? | There are exactly enough books in a classroom to give each student the same number of books.<br><br>If 5 more students join the class. each student will receive 2 fewer books. 6 books will be extra.<br><br>If 3 fewer students join the class each student will receive 1 book more. There will be 11 books extra.<br><br>Let x be the number of students in class.<br>Let y be the number of books for each student.<br><br>4.1 Find the number of students in the class.<br>4.2 Find the **total** number of books available. |

The adaptations were as follows:

*Lexical features*

- Words like *give* instead of *supply* simplify the language.

*Syntactical features*

- The hypothetical constructions of the original version have been avoided.
- The proposition density of all the original sentences has been reduced.
- The two questions included in the one sentence of the metatext have been separated into two different sentences.
- The structure of the metatext sentences is more action-specific.

*Discourse features*

- A topic sentence introduces the context.

- The proposition density per paragraph has been reduced. The three different situations are described in three different paragraphs instead of one paragraph as in the original version.

*Non-verbal features*

- The non-verbal letter symbols are separated from the verbal information.

Question 5 is once again a question on linear programming.

| **Original version: Question 5** | **Adapted version: Question 5** |
|---|---|
| In a factory x units of a product A and y units of a product B are processed through two departments. $D_1$ and $D_2$. A requires 2 hours per unit in $D_1$ and 4 hours per unit in $D_2$. B requires 3 hours per unit in $D_1$ and 2 hours per unit in $D_2$. $D_1$ and $D_2$ have 60 and 80 hours respectively available each week.<br><br>5.1 Use the above information to write down the set of constraints in terms of x and y.<br><br>5.2 Represent these inequalities on the graph sheet provided and shade the feasible region.<br><br>5.3 If the profit margins are R30 and R40 for A and B respectively, write down the objective function.<br><br>5.4 Draw the optimal search line and hence write down the recommended product-mix that will maximise profit.<br><br>5.5 Determine the maximum profit.<br><br>5.6 What percentage of daily capacity will be utilized in each department? | In a factory two products are made, product A and product B. The products are made in two departments.<br><br>One unit of product A needs two hours and one unit of product B needs three hours in the first department. This department may only be used for 60 hours a week.<br><br>The second department may only be used for 80 hours a week. One unit of product A needs four hours and one unit of product B needs two hours in this department.<br><br>Let x be the number of units made of product A.<br>Let y be the number of units made of product B.<br><br>5.1 Write down the inequalities in terms of x and y to show the above restrictions.<br><br>5.2 Draw these inequalities on the graph paper. Shade the feasible region.<br><br>5.3 The profit on product A is 30 rands per unit and on product B, 40 rands per unit. Write down the objective function.<br><br>5.4 Use your graph to calculate the maximum profit.<br><br>5.5 Calculate what percentage of the available time in each department is used when the profit is at a maximum. |

The main readability problems of this question were related to unstructured discourse and difficult vocabulary. Adaptations include the following:

### *Lexical features*

- The vocabulary is more common to everyday life and in a sense also more action-specific. For example, in the first paragraph the verb *made* replaces the verb *processed* and Question 5.2 changes from *Represent these inequalities on the graph sheet...* to *Draw these inequalities on the graph paper*. The word *inequalities* is used consistently in 5.1 and 5.2 whereas in the original version *inequalities* and *constraints* are interchanged.

The rather difficult vocabulary of the last question has also been simplified.

### *Syntactical features*

- The proposition density of a few sentences has been reduced. For example, the information of the first sentence of the original version is communicated by four sentences in the adapted version.

### *Discourse features*

- The reduction of proposition density per sentence changes the first sentence of the adapted version to a type of topic sentence.

- The proposition density per paragraph has been reduced. The text of the original version has only one paragraph, whereas the adapted version has four paragraphs.

- The discourse has been restructured so that the information about each department is grouped together. Information communicated in this way tends to give a clearer focus. Instead of fusing, the information is communicated in discrete packages. This also leads to the more effective construction of a CKS.

*Non-verbal features*

- The identification of the variables has been reserved for later.
- D1 and D2 have been omitted so that narrative part of the text can be communicated in as much verbal language as possible.

Question 6, like Question 2, is an example from kinetics theory.

| **Original version: Question 6** | **Adapted version: Question 6** |
| --- | --- |
| Two cyclists A and B travel from the same place and leave simultaneously. The distance. s in kilometres, that both of them are from the starting point after t hours, is respectively given by the formulae.<br><br>$S_A = t^3 - 12t^2 + 36t$ and<br><br>$S_B = 24t - 4t^2$<br><br>Both are back at the starting point simultaneously.<br><br>6.1 Find the average speed of A over the first 2 hours.<br><br>6.2 Find the speed of B 4 hours after he has left. Interpret your answer.<br><br>6.3 After how many hours has each begun to return to the starting point? Interpret your answer. | Two cyclists, Frances and Lizzy, start from the same starting point at the same time. They travel along two different straight roads.<br><br>The **distance** Frances is **from the starting point** is given by<br><br>$F(t) = t^3 - 12t^2 + 36t$<br><br>F(t) is in kilometres.<br>t is in hours.<br><br>The **distance** Lizzy is **from the starting point** is given by<br><br>$L(t) = 24t - 4t^2$<br><br>L(t) is in kilometres.<br>t is in hours.<br><br>Frances and Lizzy are back at the starting point at the same time.<br><br>6.1 Find the **average** speed of Frances after the first two hours.<br><br>6.2.1 Find the speed of Lizzy after four hours.<br>6.2.2 What conclusion can you reach from your answer?<br><br>6.3.1 After how many hours does Frances begin to ride back to the starting point?<br>6.3.2 After how many hours does Lizzy begin to ride back to the starting point?<br>6.3.3 What conclusion can you reach from your answer? |

The functions of Question 6 are both *position*-time functions, but because of the same reasons mentioned when discussing Question 2, the word *distance* was retained. *Displacement* would also have been a wrong word choice because displacement is defined within a certain time interval only. Besides, only students taking physical science are really acquainted with the concept of displacement.

Apart from the inappropriate functions, one of the main readability problems of this question was visualization. Adaptations tried to address this and other problems in various ways. On the whole the language use is more explicit.

***Lexical features***

- *Lizzy* and *Frances* allow students to identify with the problem.
- The word *average* has been highlighted in 6.1. This was an attempt to distinguish between the two types of questions in 6.1 and 6.2.

***Syntactical features***

- The proposition density in 6.3 has been reduced. Three answers are required so the three questions have been set using three different sentences.

***Discourse features***

- The visualization problem has been addressed by providing explicit information. For example, the information in the text concerning the two cyclists has been given in two separate paragraphs.

***Non-verbal features***

- In the narrative part of the text all non-verbal information has either been replaced by verbal information or removed.

- Function notation has been introduced.

The next verbal problem resembles the one in Question 3, only this time hi-fis, instead of calculators, are being sold.

| Original version: Question 7 | Adapted version: Question 7 |
|---|---|
| A manufacturer of hi-fi sets determines that in order to sell x units of a new hi-fi, its price per unit must be p = (1000-x) Rand. He also determines that the total cost, c, of producing x units is given by c = (3000 + 20x) Rand.<br><br>7.1 Deduce an expression for the total profit in terms of x.<br>7.2 How many units must be produced and sold in order to maximise the profit? | A manufacturer makes hi-fi sets. The **total** cost to make these hi-fi sets is given by<br><br>C(x) = 3000 + 20x<br><br>C(x) is the total cost in rands<br>x is the number of hi-fi sets.<br><br>The selling price of **one** of these hi-fi sets is given by<br><br>S(x) = 1000 - x<br>S(x) is the selling price for **one** hi-fi set in rands.<br><br>7.1 Find an expression for the total profit in terms of x.<br><br>7.2 How many hi-fi sets must be sold in order to get a maximum profit?<br><br>Hint: Total profit = Total selling price - Total cost price. |

As in Question 3, the main problem in this question is the entangled verbal/non-verbal information as well as the artificial functions used for the cost and selling price. Unlike Question 3 however, Question 7 does not contain redundant information.

The following adaptations have been suggested:

*Lexical features*

- More commonplace vocabulary is used. The verb *make* replaces the word *producing* and *one set* is used instead of *per unit*.

- The words *one* and *total* have been highlighted to emphasize the important difference between the two functions.

*Syntactical features*

- The proposition density of the original first two sentences has been reduced.

- The sentences of the adapted version are more action-specific.

*Discourse features*

- The short introductory sentence serves as a topic sentence.

- The proposition density per paragraph has been reduced.

- The information has been structured to aid the construction of a CKS. Information regarding the cost price is given *before* the information of the selling price.

- The hint was inserted for the same reason mentioned in the discussion of Question 3.

*Non-verbal features*

- The misleading non-verbal letter symbol p has been replaced by function notation S(x).

- The entangled verbal/non-verbal information has been separated.

In the next verbal problem, a function communicates information about the circulation of a magazine.

| **Original version: Question 8** | **Adapted version: Question 8** |
|---|---|
| It is estimated that t years from now the circulation of a magazine will be given by the function<br><br>$m(t) = 50t^2 - 200t + 3000$<br><br>8.1 At the time of first publication, what was the circulation of the magazine?<br><br>8.2 Derive an expression for the rate at which the circulation will be changing t years from now.<br><br>8.3 At what rate will the circulation be changing 3 years from now?<br><br>8.4 Determine by how much the circulation will be changing during the 2nd year. Will it be increasing or decreasing? | A new magazine is published. It is determined that after each year the number of magazines sold will be given by<br><br>$m(t) = 50t^2 - 200t + 3000$<br><br>m(t) is the number of magazines sold in one year.<br>t is the number of years the magazine has been published.<br><br>8.1 How many magazines are sold on the first day they begin selling the magazines?<br><br>8.2 Find an expression in terms of t to show how the rate at which the magazines are being sold changes.<br><br>8.3 In three years' time, at what rate will the number of magazines being sold change?<br><br>8.4.1 What is the difference between the number of magazines sold at the end of the second and the end of the first year?<br><br>8.4.2 Will the number of magazines being sold be decreasing or increasing? |

In this question, the behaviour within the context rather than the context itself caused problems (cf. Steffensen et al., 1979; 3.3.2). Unfortunately, this could not be addressed successfully in the adaptations because, as one student wrote, *"I don't understand the question so I don't know what to change"*.

The adaptations include the following:

***Lexical features***

- Overall the vocabulary of the adapted version is more commonplace, which here makes the information more explicit. Compare the two versions of Question 8.4.

- The word *circulation* has been avoided. Although an adult reader might not have problems with this word, some E2 and E3 readers did experience difficulties.

*Syntactical features*

- The sentences are more action-specific. For example, compare the two versions of Question 8.1. The more active sentence structure of the adapted version avoids the two nominalizations, *publication* and *circulation*.

***Discourse features***

- A topic sentence has been used.

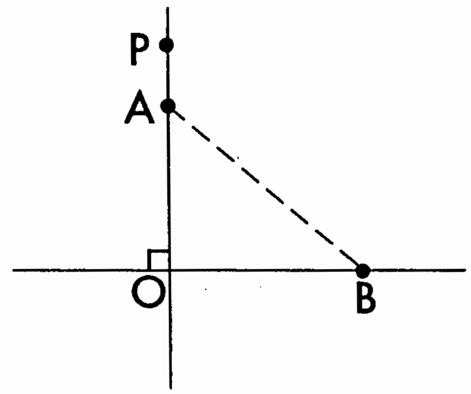
- The proposition density in Question 8.4. has been reduced. One paragraph has been replaced by two, one for each question.

*Non-verbal features*

- The non-verbal letter symbol has been removed from the narrative part of the text.

- The non-verbal formula is described more explicitly.

The last verbal problem is about two people travelling along two different roads.

| **Original version: Question 9** | **Adapted version: Question 9** |
|---|---|
| Two straight roads intersect perpendicularly at 0. P is a point on one road such that 0P = 10 km. Two persons at 0 and P respectively start to walk simultaneously, the one at P in the direction of O along the one road at 3 km/h, and the other at 0 along the other road away from 0 at 4 km/h. After t hours they reach the positions A and B on the two roads.<br><br>9.1 Find an expression for $AB^2$<br><br>9.2 After how many hours are the persons nearest to one another?<br><br>9.3 What is the shortest distance between them? | In the sketch two straight roads cross each other at 90°.<br><br>0P = 10 km<br>0 is the point of intersection.<br>P, A and B are points on the roads.<br><br>Two girls, one at P and one at 0, start walking at the same time. The one at P walks towards 0 at 3 km/h. After t hours she reaches the point A on the road.<br><br>The girl at 0 walks away from 0 along the other road at 4km/h. After t hours she reaches the point B on the road.<br><br>9.1 Find an expression for $AB^2$ in terms of t.<br><br>9.2 After how many hours are the girls closest to each other? (i.e. when does AB have its smallest possible value?)<br><br>9.3 How far is this shortest distance between them? |

The main readability problem of this question was that students found it difficult to visualize the situation. The high proposition density of the third sentence increased this difficulty, especially for E2 and E3 readers. Adaptations once again addressed lexical, syntactical and discourse features.

***Lexical features***

- Difficult vocabulary like *respectively* and *simultaneously* have been omitted.

*Syntactical features*

- The proposition density per sentence has been reduced. The information of the third sentence in the original version is communicated by three sentences in the adapted version.

- The sentences have been constructed so as to aid the construction of a CKS. Information relating to different happenings are communicated in different sentences. Information input is therefore more discrete.

*Discourse features*

- An important adaptation is the inclusion of a sketch. The situation immediately becomes more comprehensible. The sketch even aids the writer.

- The ambiguity of the information concerning the shortest distance between the two girls (Question 9.2) has been addressed by explicitly referring to what distance is meant. During the analysis of the talk-alouds it became clear that some students understood the shortest distance referred to, to be the shortest distance *road-wise,* i.e. the shortest distance of (0A + 0B).

- The proposition density per paragraph has been reduced. The one paragraph of the original text has been replaced by three paragraphs in the adapted version as well as a *described* sketch.

- Every paragraph communicates like information.

The original and adapted versions of all nine questions, as presented above, were used to gather test scores for the statistical confirmation of the hypothesis which states:

*Improved readability of the common language used in verbal problems in mathematics examination questions will improve achievement.*

## 5.2 Testing the hypothesis

### *5.2.1 The subjects*

More than 300 std 10 HG students from twelve different schools wrote the composite test containing all three versions of the nine verbal problems. All students in the std 10 HG class wrote the test as this was the easiest practical arrangement. According to previously attained information the number of students from these schools would, after selection, provide sufficient test scores for the testing of the hypothesis. E1 and E2 students came from six traditionally white schools in the Cape Peninsula while the E3 students were from six semi-private schools, mainly in the Transvaal. The selection of students whose test scores would ultimately be used in the statistical analysis was done according to the following criteria. Students had to have:

1. A scholastic achievement level between 55% and 75% for Mathematics.
2. A stanine score between 6 and 9 for the HSRC's Mathematics achievement test. (The stanine scale is a nine-point scale according to which raw scores are divided into nine intervals. It provides scores ranging from 1, very poor, to 9, very good, with a mean of 5).
3. A stanine score between 6 and 9 for the HSRC's English proficiency test.

The rationale behind the first criterion came from the protocol study. Analysis of the protocols predicted that bright students above the 75% level would probably not gain much by improved readability. These students' high subject proficiency carried them through many of the comprehension difficulties experienced in the talk-aloud experiment. It was also predicted that students achieving below the 55% level would not gain much by improved readability either, in any case not achievement-wise. Their mathematics was probably too weak. Improved readability could at most aid reading ease or comprehension. Both predictions were confirmed while marking the composite tests. The criteria from the two HSRC tests were to help with the matching of students from different schools. Using scholastic scores only might mean that students would not be matched well enough due to possible differences in scholastic standards. Eventually the results of 108 selected students, 36 from each language group, were used for the statistical analysis. The number of students from each language group had to be a multiple of three because of the questions that were posed in three different versions.

### *5.2.2 The testing method*

Students wrote the composite test as an ordinary class test. This was done in order to keep the test situation as authentic as possible. However, students knew the results would not affect their grades so the usual test anxiety that plays a definite role in all final examinations was virtually absent. Moreover, tests were not time-limited, which removed much of the normal tension as well.

Three sets of test papers were compiled. Every set included all nine questions, but the order of the three versions varied. (see Table 5.2. Reasons for the inclusion of a non-verbal version have been discussed in 5.1.1). The first set started with Question 1 in the original version, Question 2 in the adapted version, Question 3 in the non-verbal version and so on for all nine questions. The second set started with Question 1 in the adapted version and the third set had Question 1 posed in the non-verbal version. So every student answered the nine questions, three questions in each version. In this way each question was answered in all three versions but no student was asked to answer the same question more than once. The following table illustrates the composition of the three sets of test papers.

**Table 5.2 Composition of three sets of test papers**

| Questions | Set 1. | Set 2. | Set 3. |
|---|---|---|---|
| 1 | O | A | N |
| 2 | A | N | O |
| 3 | N | O | A |
| 4 | O | A | N |
| 5 | A | N | O |
| 6 | N | O | A |
| 7 | O | A | N |
| 8 | A | N | O |
| 9 | N | O | A |

**O:** Original version. **A:** Adapted version. **N:** Non-verbal version

As mentioned, all the std 10 HG students of the relevant schools wrote the tests. All test papers were marked according to three memorandums, one for each version, but only the scores of selected students were used in the statistical analysis reported in the next section. Averages of each version were converted to a number out of 10 which makes the scores more comparable.

### 5.2.3 *Test results and the statistical analysis*

The hypothesis was tested by subjecting the differences between the test scores for the original and adapted versions to the Mann-Whitney U-test, a test designed to measure the probability that a given difference occurs by chance. In this study the differences were measured relative to achievement, the dependent variable in this case. Because the questions were not comparable, mathematics- or readability-wise, the hypothesis was tested for each question separately. Test results are reported by first considering the three language groups as a combined unit. Then an account is given of the results for each individual language group.

#### 5.2.3.1 The hypothesis tested for the combined language group

When considering the students as one language group, the number of students writing the test for each of the three versions is 36, giving a total number of 108 students per question. The results are followed by a brief discussion of the scores. This has been done for each question. Although the non-verbal test scores were not used in the composite test, they are nevertheless reported. By comparing the results of this version with that of the two verbal versions, one can form an idea of the extent to which the verbal part of the problem prevented students from demonstrating other mathematical abilities. The reported p-values in Table 5.3 refer to the difference between the test scores of the original and the adapted versions.

**Table 5.3 Test scores of original, adapted and non-verbal versions for the combined language group (E1+E2+E3)**

Maximum score for each question: 10

| Question | Original version | Adapted version | Non-verbal version | Mann-Whitney test. p-values |
|---|---|---|---|---|
| **1** | 4,6 | 6,5 | 6,3 | < 0,005* |
| 2 | 4,7 | 5,5 | 7,5 | < 0,2 |
| **3** | 2,4 | 6,3 | 8,3 | < 0,001* |
| 4 | 3,4 | 3,1 | 7,4 | < 0,9 |
| **5** | 6,6 | 8,1 | 8,6 | < 0,01* |
| **6** | 5,1 | 6,2 | 9,2 | < 0,06* |
| 7 | 5,2 | 5,8 | 8,5 | < 0,6 |
| **8** | 5,7 | 7,7 | 9,6 | < 0,002* |
| **9** | 4,9 | 6,4 | 6,5 | < 0,1* |

Differences between the original and adapted versions were taken to be significant for a p-value ≤ 0,1. According to this criterion the hypothesis was confirmed in six of the nine cases. The strongest confirmation came from Questions 1, 3 and 8. As expected, improved readability did not improve the achievement levels of all questions in the same way. For this, the nine questions, as well as their corresponding readability problems, differed too much from one another.

When considering the test scores of the different versions, the fact that improved readability improved achievement in all but one case, is significant. A few other interesting findings are discussed for each question:

- Apart from the fact that the hypothesis was confirmed by **Question 1**, another surprising result of this question is the fact that there is little difference between the test scores of the adapted and non-verbal versions. The same applies to **Question 5,** a similar linear programming problem. This is a surprising result because the non-verbal versions all had the inequalities given, whereas the adapted versions of these two questions required students to translate the verbal problem into non-verbal language, normally a difficult skill for students to master. Even though the adapted version of Question 1 identified the variables, the verbal information still had to be translated into mathematical models (cf. 5.1.3, p.16). Achievement levels of the adapted and non-verbal versions of Questions 1 and 5 seem to suggest that students find mathematization less difficult if the text allows them to construct a contingent knowledge structure more effectively. One of the main readability problems addressed in these two adaptations was related to the overall structure of text.

- In **Question 2** there was no need for mathematization in any of the versions and yet there is quite a difference between the achievement levels of the verbal and non-verbal versions. The suggested improvement in readability did not manage to improve achievement significantly. This could have been due to the inherent mathematical difficulty of the question, but when considering the question once more, one realizes that the overall mathematical demand is relatively easy. This is confirmed by the achievement level of the non-verbal version which shows that students were able to do most of the computations. The real problem of this question seems to be the fact that the context remained remote, despite the adaptations. Students had no idea what the object was and to what movement the problem was referring. Because the context could not be experienced the attempt to relate mathematics to the real world failed and the intended support for mathematical thinking was counter-productive.

- In **Question 3**, the hypothesis was confirmed at a level less than 0,001, the most significant for all the questions. Redundant information, discussed in 5.2, and inappropriate functions were the two main readability problems of the original version. The achievement level of this version is the weakest of all 27 versions used in the experiment. This version did not require mathematization of the necessary basic functions and the achievement level of the non-verbal version shows that students could execute most of the computations successfully. So the differences in test scores demonstrate in no uncertain terms that the difficulties encountered with the original version of this problem were more related to comprehension than to mathematical incompetency.

- **Question 4** gave an unexpected result. The achievement level of the adapted version is lower than the original version, which means that improved readability has done the opposite of what was intended. This is the only question where this happened. In all the other questions, improved readability improved achievement, although not always significantly. When considering the adaptation of this question once again, the reason for the unexpected result could not immediately be found. So the adaptation of this question was discussed with different groups of students who did the tests. From the reasons given by the students for this strange phenomenon, the following can be concluded:

* Improved readability does not always improve comprehension and improved comprehension does not guarantee improved achievement levels.
* The mathematization of the question was experienced as exceptionally difficult. Improved readability was not enough help to remove this difficulty.
* The adaptation had introduced another readability problem for some students. Students did not know to which of the three conditions the letter symbols x and y were referring.
* For some E3 students, the main readability problem had not been removed. They experienced the books as belonging to a subject, say mathematics, so to them, $x = y$. By the first sentence they understood everyone was lucky enough to receive **one** book.
* Another comprehension problem that had not been removed effectively was the strange situation (to some readers) that while on the one hand there is a shortage of books (students receive fewer books), on the other there is a surplus (books are left over).

This problem was experienced mainly by E3 students and they apparently lost much time and confidence when trying to understand this "strange information".

* When reconsidering Question 4, it became clear that this question was actually a type of mathematical puzzle and not true to the real-life experiences of students. It would be interesting to know what the students' performance would have been if the question had been supplied with a heading to introduce the idea of a puzzle.

From the above it is clear that there were factors that superseded the principle of improved readability in the adapted version.

- **Question 6** is similar to Question 2 because it requires no mathematization and is also an example from kinetics theory. Unlike Question 2, the hypothesis was confirmed by Question 6 for the combined group. It is possible that students could identify more with a bicycle ride than with an unknown moving object. However, the important readability problems of Question 6, visualization and inappropriate functions, could not be removed successfully because of the fact that the mathematical demand of the original and adapted versions had to be maintained. For example, a graph asked for at an earlier stage could have aided comprehension, but would have altered the mathematical likeness of the two versions.

- **Question 7** is similar to Question 3, but without redundant information. The hypothesis was not confirmed by Question 7. There is very little difference between the test scores of the original and adapted versions. This can be understood if one takes into account that the inappropriate functions, causing much confusion during the talk-alouds, were not changed in the adapted version. Achievement levels of the non-verbal versions show that students were once again able to execute most of the necessary computational skills. Moreover, this question, like Questions 2, 3, 6, 7 and 8, did not require mathematization of basic functions so difficulties were once again more related to comprehension of the situation than to mathematical inabilities.

- The adapted version of **Question 8** managed to address some of the comprehension problems successfully as the hypothesis was confirmed at a level of 0,0015. One of the important readability problems of this question was the rather unfamiliar word, *circulation*, experienced as difficult, especially by E3 readers. Question 8 had the necessary non-verbal formula given in the original and adapted versions. Most of the questions required a few easy computations, so the fact that there is still such a great difference between the test scores of the adapted and non-verbal versions suggests that not all readability problems have been addressed by the adapted version. When considering the adapted version, one realizes this to be true because much of the information still remains vague: It is still not clear when the magazines are published for the first time and the *changing rate at which the magazines are being sold* is not a situation that is experienced by most students.

The last of the nine questions also confirmed the hypothesis for the combined language group. Two major adaptations, a sketch and the reduction of proposition density, characterized the more readable version of **Question 9**. Both were aimed at removing the visualization problem picked up during the analysis of the protocols. The adaptations seem to have reduced the readability problems to such a degree that the achievement level of the adapted version is virtually the same as that of the non-verbal version, although the adapted version requires mathematization and the non-verbal version has the mathematical model given.

Overall, two important results from the combined language group are:

1. Improved readability improved achievement in eight of the nine cases, differences being statistically significant at the 0,1 level in six cases. The improved achievement can be explained in terms of improved readability except in the case of Question 1. (The introduction of the variables in Question 1 not only improved comprehension, but could have aided mathematization as well).

2. Three of the six significant cases refer to Questions 1, 5 and 9. All three these questions required students to mathematize and yet improved readability was able to improve test scores to more or less the same level as that of the non-verbal versions. Traditionally one would say the mathematical demand of the non-verbal versions is considerably less than the verbal versions. This matter is discussed in more detail in the next chapter (cf. 6.4.2).

The hypothesis was also tested for each language group separately.

5.2.3.2 The hypothesis tested for the separate language groups

It was necessary to determine what effect improved readability would have on the individual E1, E2 and E3 groups, as this was one of the main purposes of the study. It was also necessary to determine which language group would gain most by improved readability, first or second language readers.

Testing the hypothesis for each language group separately seemed desirable for another reason as well. One could argue that the test scores, reported above for the combined group, might have been biased by the influence of a specific language group. The hypothesis might have been accepted because of the meaningful influence of readability problems on one language group only. For example, it could be that first language readers, due to their superior language ability, pick up ambiguities more easily and are therefore prejudiced more seriously by subtle differences between intended and written information. First language readers experience a situation more quickly and more intensely than second language readers because they need not waste time and energy trying to cope with the language.

To learn more about the above issues, test scores were analyzed for each language group per question. Once again the Mann-Whitney U-test was applied to measure the differences between the test scores of the original and adapted versions for E1, E2 and E3 readers. In total, the number of students remains 108, with 36 in each of the three language groups. However, considering each group on its own meant that there were only 12 test scores available for each of the three question versions. This had implications for the statistical analysis because the relatively small number of test scores meant that it was highly probable that the differences in scores would occur by chance. The small number of test scores therefore made the significance test less sensitive which in turn made it more difficult to confirm the hypothesis quantitatively.

**Table 5.4 Results of the statistical analysis for separate language groups (E1, E2 and E3)**

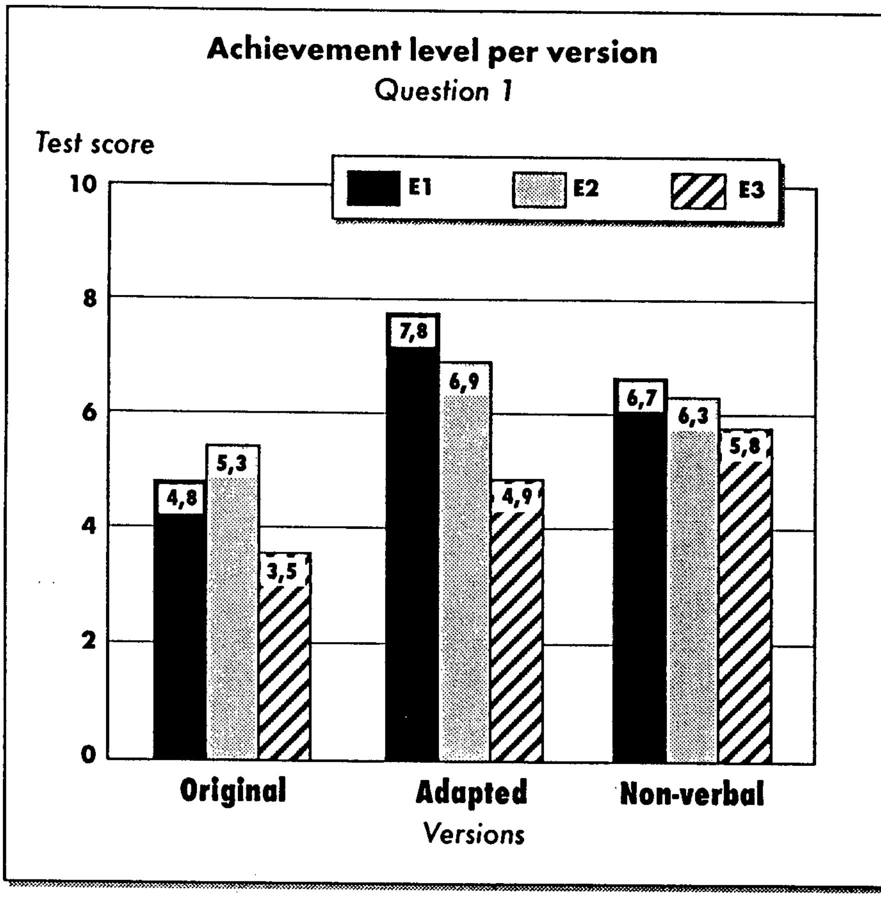
| Question no | p-values E1 readers | p-values E2 readers | p-values E3 readers |
|---|---|---|---|
| 1 | < 0,03 * | < 0,06 * | < 0,07 * |
| 2 | < 0,9 | < 0,5 | < 0,05 * |
| 3 | 0,01 * | 0,005 * | < 0,03 * |
| 4 | 1 | 0,9 | < 0,7 |
| 5 | < 0,01 * | < 0,2 | < 0,3 |
| 6 | < 0,5 | < 0,2 | < 0,4 |
| 7 | < 0,8 | < 0,6 | < 0,9 |
| 8 | 0,1 * | 0,13 | 0,02 * |
| 9 | < 0,3 | 0,01 * | < 0,3 |

As with the combined language group, differences were taken to be significant for a p-value $\leq 0.1$.

The fact that the hypothesis was not confirmed in all cases does not make the readability problems of the other questions less important. Evidence from the talk-alouds proved beyond doubt that all questions had readability problems that seriously interfered with comprehension. The above table indicates that quantitative results are not always able to point to difficulties students may be experiencing, whether they be comprehension or other mathematical ones.

To be able to report on the above results more qualitatively, the actual test scores of each question are presented *graphically* for each language group. As each question was done in three versions by three different groups, there are nine average test scores for every question. The number of students for each of the 9 possibilities was 12, giving the 108 students who participated in this part of the study.

Figure 5. 1



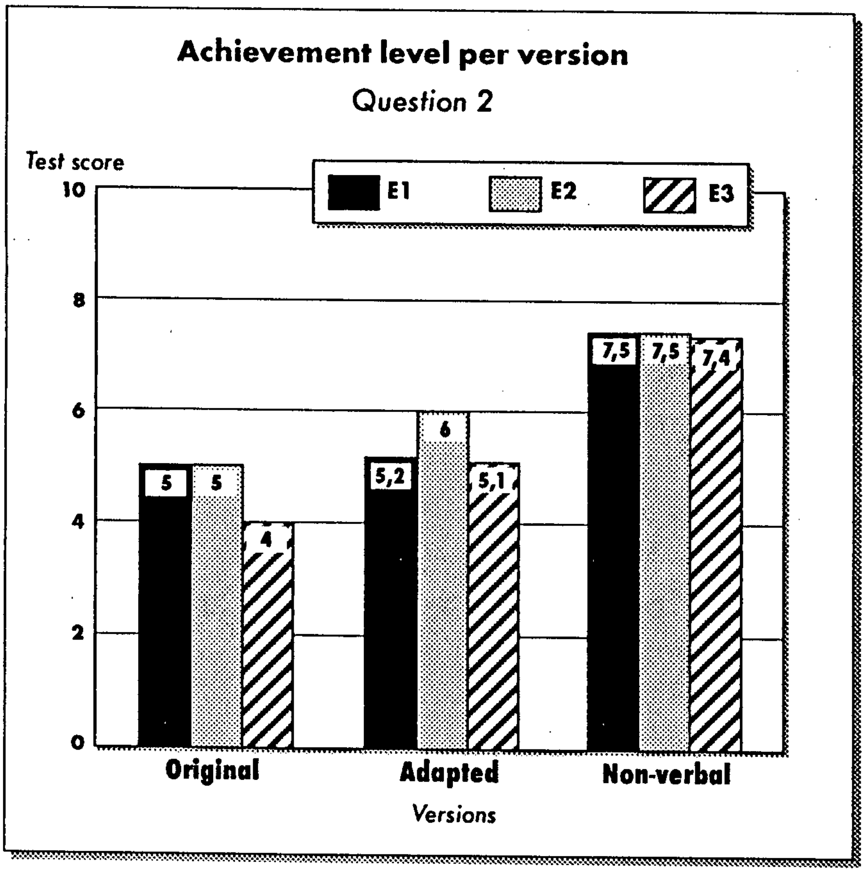

The graph reveals some interesting information regarding Question 1:

- All language groups gained considerably by improved readability. (The hypothesis was confirmed for all groups doing Question 1).

- E1 readers seem to have gained more by improved readability than the other two groups. This could be because of the relatively longer discourse of the question. The longer the discourse the more "reading energy" is needed. The extra reading energy released by the improved readability seems to benefit first language readers more.

- All three groups had a lower test score for the non-verbal version compared to the adapted version. This is a surprising result since the non-verbal version did not require mathematization whereas the adapted version did.

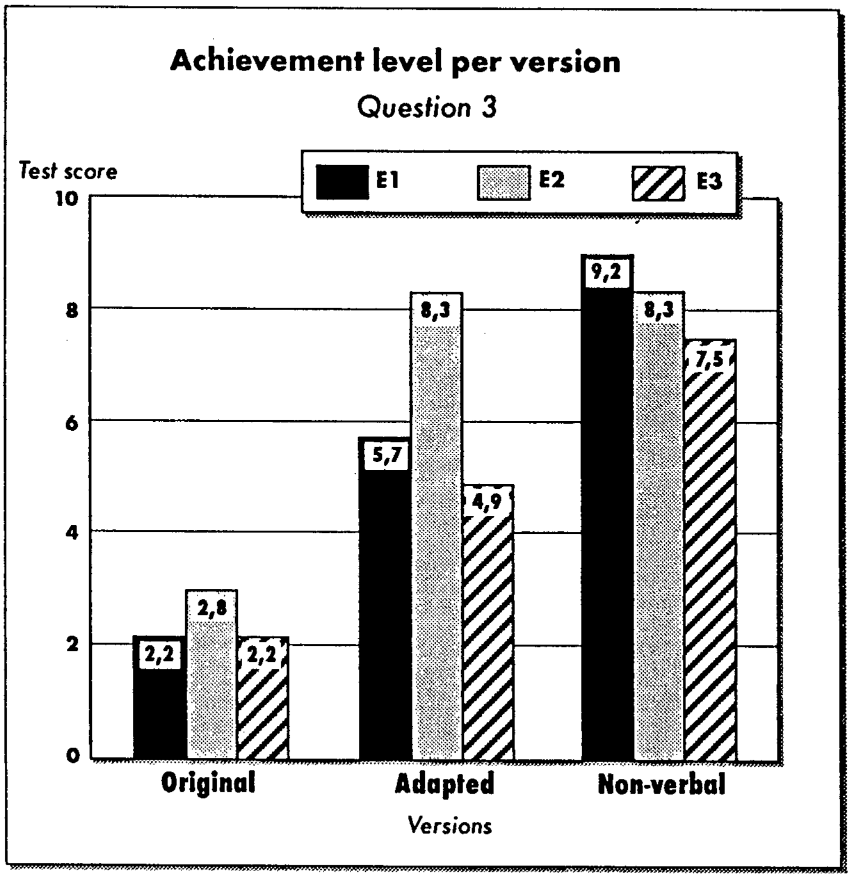
Figure 5.2




In Question 2, differences in test scores between the original and adapted versions increase from virtually nothing for E1 readers to a significant difference for the E3 group, which means that in this question the hypothesis is accepted for the E3 group only. The E3 groups' relatively weak performance on the original version is cancelled out by the improved performance on the adapted version. This is easily understood if one takes into account that the context of Question 2, an example from kinetics theory, is even more unrealistic to the E3 than to the other two groups. Fewer E3 students take Physical Science as a subject, making the rather abstract information of *a particle that moves along a straight line* even less known to African students than to white students. The more explicit description of the situation in the adapted version enabled E3 students to improve their test scores significantly.

All groups achieved equally well on the non-verbal version showing that all students were able to execute most of the necessary manipulative skills. Furthermore, it is remarkable that there are marked differences between the test scores of the non-verbal and verbal versions, in spite of the fact that Question 2 requires no mathematization. Evidence from the protocols proves that the weak performance on the original version was mainly due to imagery problems. The adapted version was unable to address this problem successfully.

Figure 5.3

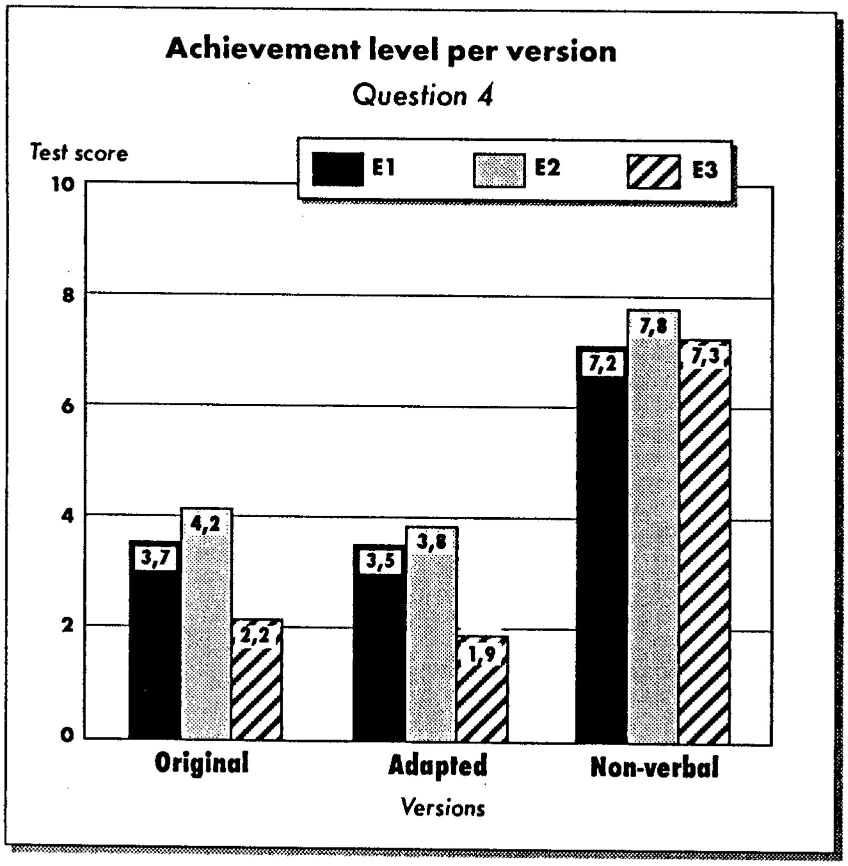



All groups performed extremely poorly on the original version and excellently on the non-verbal one. Major problems of the original version were redundant information, entangled verbal/non-verbal information and inappropriate functions. Except for the inappropriate functions, all these were removed in the adapted version. Differences between the original and adapted versions were significant for all groups so the hypothesis is confirmed in all cases.

Once again the basic formulae were given in this question, but in spite of this, *only* the E2 group performed equally well on the adapted and non-verbal versions. Evidence from the talk-aloud experiment could possibly explain this phenomenon. Two reasons can be given for the relatively weak performance of the E1 group on the adapted version. In the first place their mother tongue language proficiency allows them to experience more of the situation and concentrate less on language difficulties. They are therefore more apt to be puzzled by factors like the inappropriate functions used in this question.

Another reason for the weaker performance of the E1 and E3 groups on the adapted version, as compared to the E2 group, may possibly be found in their reading rate. Listening to the talk-alouds, it becomes clear that E1 students read very much faster than their second language peers, especially if the text is as short as in Question 3. By doing this, they could have missed important cues like *total cost* and *selling price of one calculator* which was an adaptation made explicitly to distinguish between the cost price of *all* calculators and the selling price of *one*. E2 readers, on the other hand, are aware of their language backlog and are inclined to concentrate more when reading shorter pieces of information. They also read markedly slower, which allows them to pick up important key words more effectively. During the protocol experiment, E3 students read English even more slowly than their E2 counterparts. It could be that whereas E1 readers miss out on important detail because of reading too quickly, E3 readers give undue attention to unimportant detail because of their relatively slow reading rate and other second language difficulties - like the inability to recognize key words (cf. Hamers & Blanc, 1989:95). A slower reading rate, which follows a bottom-up reading pattern, also tends to miss the global message of the text. Important *local* detail as well as *global* understanding are needed for successful application of mathematical knowledge.

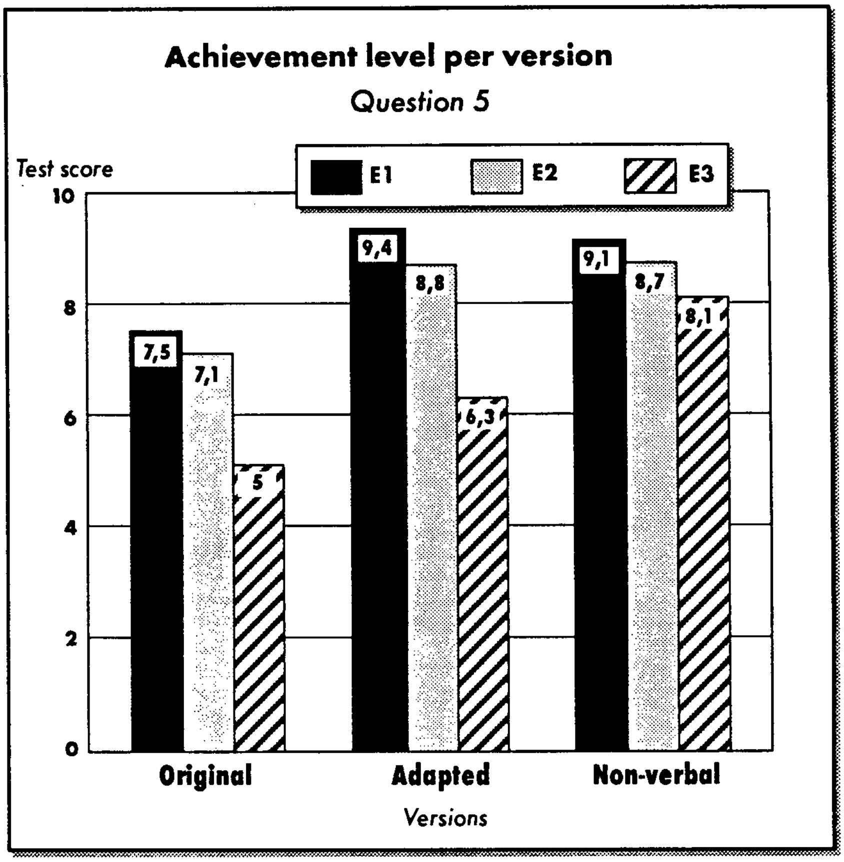
Figure 5.4




The weak performance on both verbal versions, as well as the inability of the adapted version to relieve the readability problems, have been discussed extensively in 5.2.3.1. It is clear that differences measured between the test scores of the original and adapted versions could not confirm the hypothesis for any language group.

As reported, comprehension problems were experienced more acutely by the E3 group, which could be the reason for their weaker performance on both the adapted and original versions as compared to the E1 and E2 groups. However, test scores of the non-verbal version show no meaningful difference between the groups. They were all able to solve the two simultaneous equations satisfactorily. The manipulative skills required in this question were the most difficult of all nine questions and yet all groups performed much better in the non-verbal version, especially the E3 group.

Figure 5.5

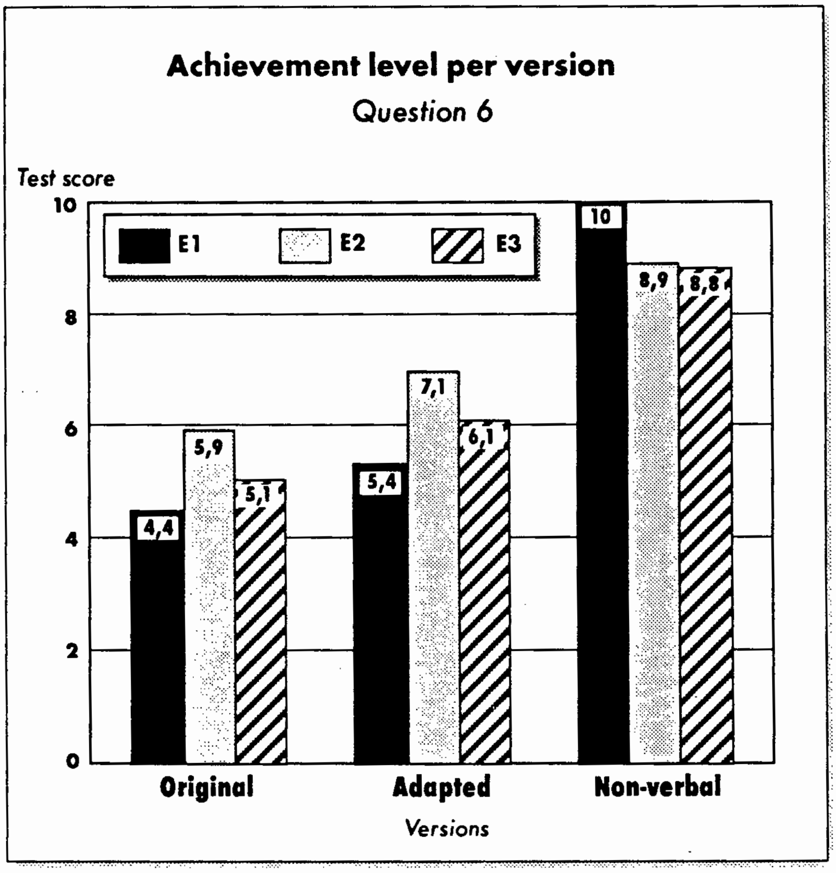



Improved readability was able to improve the achievement of all groups substantially, but only the E1 group confirmed the hypothesis. As mentioned before, the relatively small number of students, 12 in each of the nine cases, made it more difficult to obtain significant outcomes.

As in the case of Question 1, it is surprising that the test scores of the non-verbal and adapted versions are so close to each other in spite of the fact that the adapted version required mathematization whereas the non-verbal version did not. In fact, test scores of the E1 and E2 groups are even slightly higher for the adapted version. This tendency for both Questions 1 and 5 seems to suggest that, when applying mathematical knowledge, a real-life context provides students with more support than a mathematics context.

More research is needed to find the reason for the significant difference between the test scores of the adapted and non-verbal versions for the E3 group. Were mathematization or comprehension skills responsible for most of the difference? It must be remembered that all adaptations were mainly done by E1 and E2 students. Adaptations by E3 students could have addressed their comprehension problems more successfully and could have increased their test scores even more. Their achievement levels on all non-verbal versions indicate they have sufficient operational skills to achieve good scores.

Figure 5.6



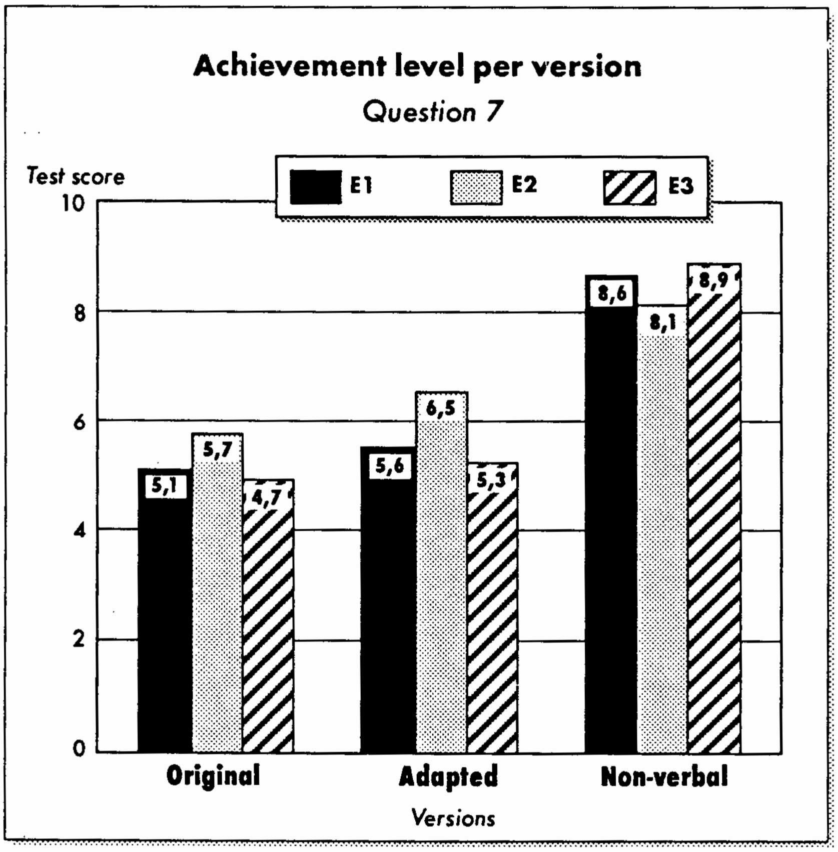

Question 6 is very similar to Question 2 and the test scores of the two questions reveal much the same pattern: All language groups have gained by improved readability with the E2 group having the highest score for the adapted version.

It is remarkable that the E2 and E3 groups performed better than the first language group on both *verbal* versions. It could be that, due to their weaker language proficiency, the second language readers are so preoccupied by language difficulties that they do not really get to the point of experiencing the situation and asking critical questions about the formulae. For example, much of their energy is used to work their way through the text by translating the information into their own language. They have less time and cognitive energy left to be troubled by the artificial functions. During the talk-alouds, the comprehension problems caused by artificial functions were mentioned mainly by first language readers.

Although the hypothesis was confirmed by this question for the combined group, it was not confirmed by any of the individual groups. Possibly the smaller number of subjects could have been the reason for this.

All language groups performed excellently on the non-verbal version. The verbal versions of the question did not require any mathematization so once again it seems that factors related to language influence achievement quite considerably.

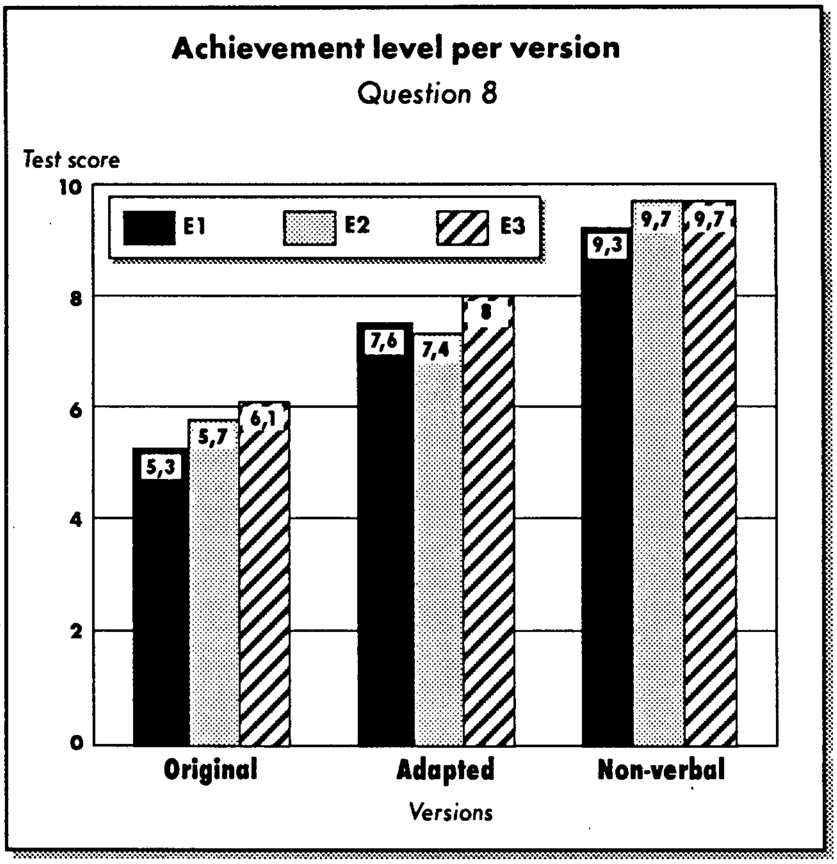
Figure 5.7



Question 7 and Question 3 are similar except for the redundant information. The fact that the hypothesis was confirmed by Question 3 for all language groups but not by Question 7 could be attributed to the redundant information in Question 3.

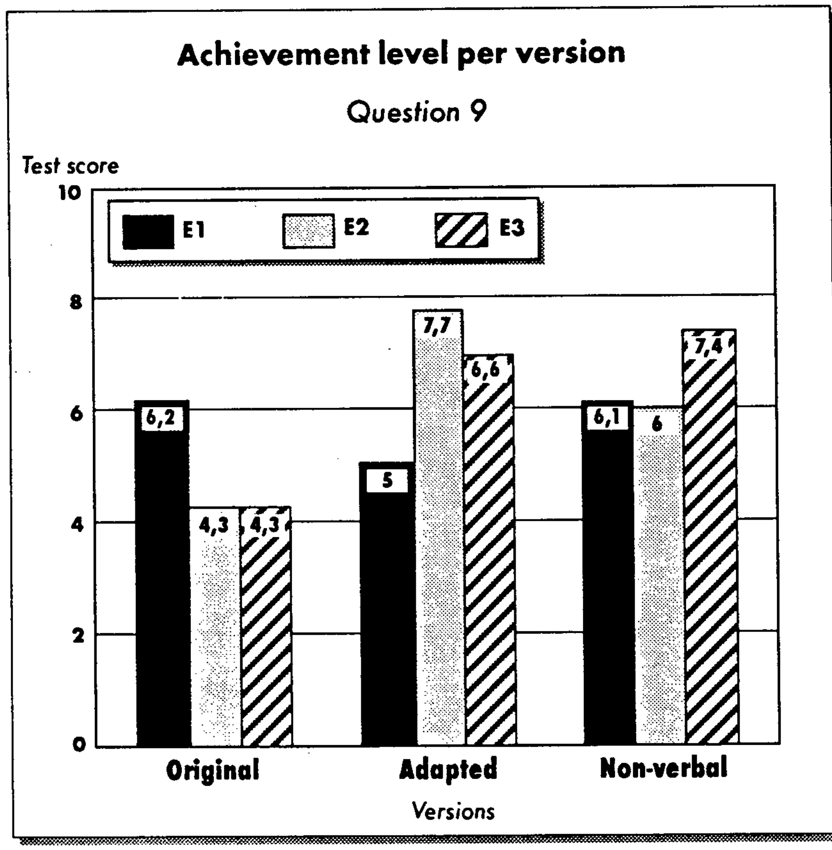
The unfamiliar behaviour within the context was not addressed in the adaptation which could be the reason for an insignificant improvement in test scores between the original and adapted versions. Many students were not acquainted with the phenomenon that cost and selling price per unit depend on the number of articles made and sold - a concept that many adult readers take for granted. This behaviour is not part of many students' real world yet, therefore the concept of optimalization *in this context* is unfamiliar to many. Test scores for the adapted version show the same pattern for these two questions: All groups gained by improved readability with the E2 group showing the greatest improvement. The reason could be the same as that mentioned in Question 3. All groups performed exceptionally well in the non-verbal versions. Question 7 does not require mathematization of basic functions so one would expect a smaller difference between the test scores of the verbal and non-verbal versions. However, this is not the case. Unfamiliar behaviour and artificial functions seem to be the main reasons for this phenomenon.

Figure 5.8




Improved readability considerably improved the test scores of all groups. Although the hypothesis was confirmed only for the E1 and E3 groups, the difference between the adapted and original versions for the E2 group nearly reached the significant level as well ($p = 0{,}1289$). Relative to the other groups, the E3 students performed excellently on all three versions. All groups performed equally well on the non-verbal version. If more readability problems could have been removed, the test scores of the adapted versions might have been even higher because this question required no mathematization. The non-verbal test scores prove that students found the necessary calculations quite easy.

Figure 5.9




It is clear that test scores of both second language groups have gained considerably by improved readability. The hypothesis was accepted for the E2 group. Although test scores of E3 students were improved by as much as 23% , a relatively large standard deviation prevented the difference between test scores of the original and adapted versions from being significant for this group.

Question 9 was the only question where a sketch was added to aid visualization. Other major adaptations were more explicit information and the reduction of proposition density per sentence and paragraph. These improvements did not seem to aid comprehension of first language readers noticeably. As a matter of fact, their average test score for the adapted version is weaker than for the original version. Information from their test papers and the protocol study gives two possible reasons for this strange phenomenon:

1. E1 readers were not so much affected by the readability problems of the original version as the two second language groups. Test scores of the original version proves this. Although E1 readers also had visualization difficulties, these were not as serious as those of their second language peers. Their mother tongue proficiency helped them to overcome the difficulties more quickly and they had enough time and cognitive energy left to construct their own sketch correctly. So even if the adapted version did improve their reading ease, it did not necessarily improve comprehension, which means that improved readability would not have affected achievement positively either.

2. The weaker test score for the adapted version could be attributed to the fact that three E1 students gained full marks for the original version. These exceptionally high scores were not repeated for the adapted version which was done by other students. As mentioned, the groups consisted of only twelve students each. Had the group been larger, the negative tendency might not have been reflected in the average test score for the adapted version.

The test scores of the non-verbal version are considerably higher than that of the original version for both second language groups, especially for the E3 group, who has the highest score. Question 9 was the only question where mathematization was required in the non-verbal version. This version of Question 9 requires students to apply their mathematics in a mathematics context, which is an abstract context to most students. Other questions that require the same skill in the non-verbal version, but without mathematization, are Questions 1 and 5. In all three questions (1, 5 and 9), the E3 group performed better in the three non-verbal versions whereas the E2 group tended to perform worse.

Bearing in mind the role of comprehension and the problem solving models discussed in 1.5, the results of Questions 1, 5 and 9 seem to suggest that the E3 readers were more able to *devise a plan* and *implement their strategies* when the information was presented in a mathematical context. More research is necessary to investigate this phenomenon. It could be related to teaching practice and the manner in which the relevant mathematical knowledge was developed.

## 5.3 The possible effect of improved readability on students' performance

The hypothesis that improved readability of the common language used in mathematics examination questions will improve achievement was confirmed in the majority of cases. This emphasizes the importance of the effect of readability on achievement in mathematics. Quantitative confirmation of the hypothesis came from seven of the nine questions, although not always from the same language group. This was in spite of the fact that in some cases the relatively small number of students made testing for significance more difficult. It has already been pointed out that improved readability does not always improve comprehension and if comprehension is not improved, achievement will not be improved either. In this respect it is quite surprising that the hypothesis was confirmed in so many cases.

Although improved readability is not always able to improve achievement significantly, there is always a gain in reading time and reading ease. Time and reading ease are two important factors in mathematics examinations. Although the effect of these factors may not be noticeable in the test scores of a specific question, it could have a marked effect on the next question or even on the whole examination paper.

Although improved readability may possibly improve comprehension, it cannot remove *all* barriers that prevent a student from improving his test scores. So the fact that the hypothesis was not confirmed in all cases, did not come as much of a surprise. Other important factors influencing achievement is the inherent difficulty of a verbal problem or the student's learning experience.

As mentioned, the fact that the hypothesis was not confirmed in all of the above cases does not mean that the readability problems of the other questions are not important. Analysis of the protocols showed that all nine questions presented serious comprehension problems that prevented students from effectively demonstrating their mathematical abilities. In this regard the *qualitative* information generated by the protocol experiment must be considered as most important and is processed in the concluding chapters.

The fact that all readability problems could not be addressed, as well as the fact that the tests were not time-limited, had a definite effect on the outcome of test scores. The protocol experiment proved that, given enough time, students are able to overcome certain readability problems effectively. Also, the tests contained a limited anxiety element because students knew the outcomes would not effect their final grades or even possibilities of bursaries or university selection. Had the factors of time and test anxiety been present, as is the case during authentic examinations, the effect of readability on achievement might have been even greater.

The main purpose of testing the hypothesis for each individual group was to ascertain what effect improved readability would have on the test scores of the three separate language groups. Comparing the test scores of the adapted and original versions, one can say that in general, all groups gained by improved readability. To be more specific, one could consider the *total* improvement of test scores per language group across all nine questions. The following table illustrates the average percentage score of all nine questions for the original and adapted versions as well as the *total* improvement in test scores due to improved readability. Differences in scores between the adapted and original versions were tested by once again subjecting the differences to the Mann-Whitney U-test. The corresponding p-values are also reported in Table 5.5.

**Table 5.5 Average percentage score of all nine questions for the original and adapted versions**

| Group | Average % for nine original versions | Average % for nine adapted versions | % Gain: | p-value |
|---|---|---|---|---|
| E1 (n=108) | *49* | *61* | *12* | $p < 0{,}02^{**}$ |
| E2 (n=108) | *51* | *70* | *19* | $p < 0{,}001^{**}$ |
| E3 (n=108) | *41* | *55* | *14* | $p < 0{,}002^{**}$ |
| E1+E2+E3 (n=324) | *47* | *62* | *15* | $p < 0{,}001^{**}$ |

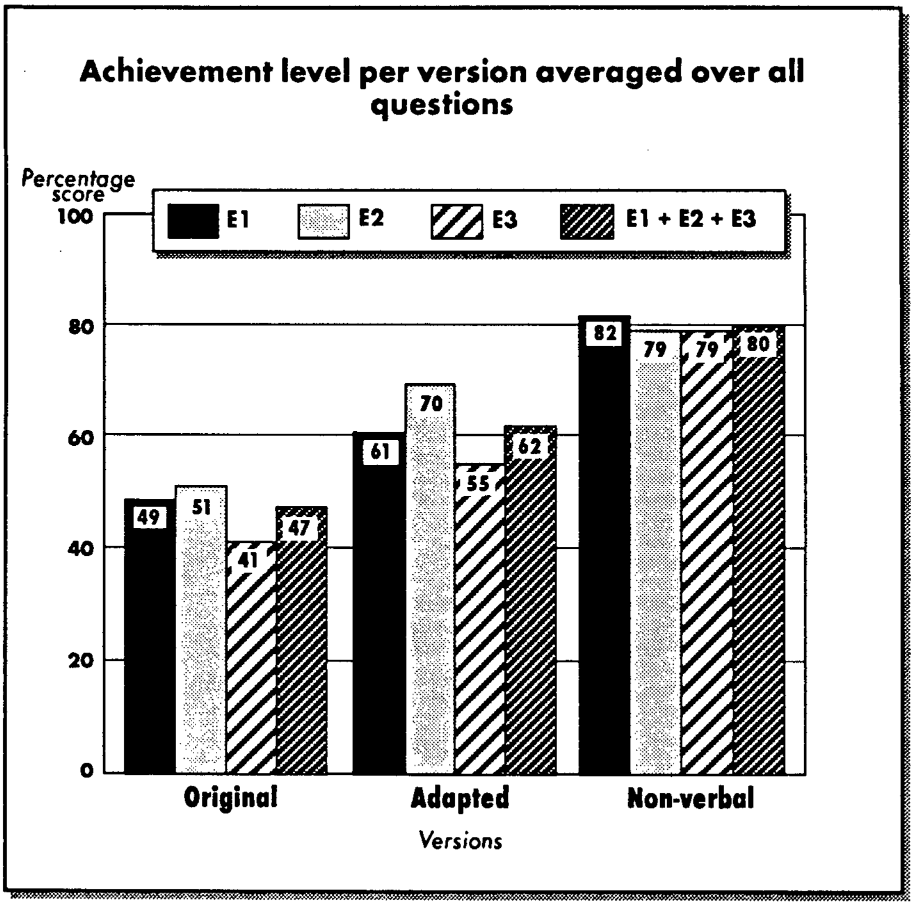
If one were to accept the above results of the original and adapted versions coming from two examination papers, an "original" and an "adapted" one, then one could conclude that students who had written the "adapted" paper performed significantly better than the other group. Overall, the two second language groups seem to have gained most by improved readability, with the E2 group gaining most. Improved readability improved the average achievement levels of E1, E2 and E3 students by 12% , 19% and 14% respectively.

Although the level of significance revealed in Table 5.5 confirms the hypothesis for all language groups, the influence of readability problems reaches much further than the quantitative results are able to show. Poor formulation could have far-reaching effects in authentic examination conditions. For example, just one serious readability problem could cause a serious enough emotional upset to influence an examinee negatively for the rest of the examination paper.

To form an idea of the practical significance of the results given in Table 5.5, a superficial analysis was done of five previous examination papers. The analysis shows that an average of 32% of the total marks could have been affected by readability factors. If one keeps in mind that improved readability has improved the test scores of the combined language group by 15%, then improved readability could improve the achievement level of students writing a complete paper by about 4,8%. However, as indicated above, the improvement could be even more than the suggested 4,8% in authentic conditions due to factors like loss of time and increased tension. A suggestion regarding this matter is given in 7.3.5.

A total graph illustrating the test scores reflected in Table 5.5 is shown in Figure 5.10. For the sake of completeness, the graph also includes the test scores for the non-verbal versions.

Figure 5.10




One of the salient results illustrated in Fig. 5.10 is the equal performance of all language groups on the non-verbal version. This has special significance for the E3 group. Whereas this language group had the weakest performance for the two verbal versions, it performed equally well on the non-verbal version. This indicates that these students have the same operational skills at their disposal as their E1 and E2 counterparts, but that factors related to language play a larger role in their performance compared to the other two language groups. These results should, however, not be interpreted to suggest that mathematics should rather be tested out of context to avoid less verbal language for the E3 group. This would jeopardize the rationale underlying realistic mathematics education (cf. 1.1). A suggestion to address this aspect is discussed in 7.1.4.

The fact that first language students also gained much by improved readability was surprising. Although the talk-alouds did reveal that E1 readers had experienced many readability problems, there was evidence that given time, they were able to overcome many of the problems due to their mother tongue language proficiency. Subjects who did the talk-aloud experiment were also exceptionally bright in mathematics, which could have helped them overcome the comprehension difficulties as well. The composite test on the other hand, was done by students in the 55% - 75% bracket. This study has indicated that, achievement-wise, improved readability has the greatest influence on this last group of students.

As mentioned before, the improvement of test scores was not the same for each question because questions and their corresponding difficulties were too different. (As the questions reveal, these varying difficulties were not only on the readability level). To compare achievement levels of *different* questions within the *different* groups, consider Table 5.6 which is a summary of test results for all questions and versions. Percentage gains between versions is given for each case.

**Table 5.6 Summary of test results of the nine questions for each of the three versions. Results expressed as percentages**

| LANGUAGE GROUPS | QUESTIONS | VERSIONS | | | | |
|---|---|---|---|---|---|---|
| | | O:Original | A:Adapted | | NV:Non-verbal | |
| | | Score | Score | % Gain O to A | Score | % Gain O to NV |
| E1 | 1 | *48* | *78* | *30** | *67* | *19* |
| | 2 | *50* | *52* | *2* | *75* | *25#* |
| | 3 | *22* | *57* | *35*** | *92* | *70#* |
| | 4 | *37* | *35* | *-2* | *72* | *35* |
| | 5 | *75* | *94* | *19*** | *91* | *16* |
| | 6 | *44* | *54* | *10* | *100* | *56#* |
| | 7 | *51* | *56* | *5* | *86* | *35#* |
| | 8 | *53* | *76* | *23** | *93* | *40#* |
| | 9 | *62* | *50* | *-12* | *61* | *-1* |
| **E1: Average for 9 questions** | | ***49*** | ***61*** | ***12**** | ***82*** | ***33*** |
| E2 | 1 | *53* | *69* | *16** | *63* | *10* |
| | 2 | *50* | *60* | *10* | *75* | *25#* |
| | 3 | *28* | *83* | *55*** | *83* | *55#* |
| | 4 | *42* | *38* | *-4* | *78* | *36* |
| | 5 | *71* | *88* | *17* | *87* | *16* |
| | 6 | *59* | *71* | *22* | *89* | *30#* |
| | 7 | *57* | *65* | *8* | *81* | *24#* |
| | 8 | *57* | *74* | *17* | *97* | *40#* |
| | 9 | *43* | *77* | *34*** | *60* | *17* |
| **E2: Average for 9 questions** | | ***51*** | ***70*** | ***19***** | ***79*** | ***28*** |
| E3 | 1 | *35* | *49* | *14** | *58* | *23* |
| | 2 | *40* | *51* | *11** | *74* | *34#* |
| | 3 | *22* | *49* | *27** | *75* | *53#* |
| | 4 | *22* | *19* | *-3* | *73* | *51* |
| | 5 | *50* | *63* | *13* | *81* | *31* |
| | 6 | *51* | *61* | *10* | *88* | *37#* |
| | 7 | *47* | *53* | *6* | *89* | *42#* |
| | 8 | *61* | *80* | *19** | *97* | *36#* |
| | 9 | *43* | *66* | *23* | *74* | *31* |
| **E3: Average for 9 questions** | | ***41*** | ***55*** | ***14***** | ***79*** | ***38*** |
| **E1+E2+E3: Total for 9 questions** | | ***47*** | ***62*** | ***15***** | ***80*** | ***33*** |

*: Hypothesis confirmed at a level, $p \leq 0{,}1$

**: Hypothesis confirmed at a level, $p \leq 0{,}01$

#: Percentage gain, O to NV, for questions needing no mathematization in the original version

African students reacted well to adaptations done essentially by E1 and E2 students. There are significant gains and the hypothesis is confirmed in four of the nine cases with considerable gains in other questions as well. Although the suggestions for adaptations were mainly dictated by white students, readability research reported in Chapters 2 and 3 guided the adaptations and linguistic and cultural features were taken into consideration as far as possible. This could be the reason why test scores of E3 students show a considerable improvement in spite of the fact that only a few of their suggestions were available when the final adaptations were done for the composite test.

The purpose of including a non-verbal version for each question was to ascertain the extent to which verbal problems prevent students from demonstrating their operational skills. In all nine questions, original and adapted, most marks were allocated for these skills. The exceptionally high test scores for the non-verbal versions make one realize that students were able to do most of the necessary calculations, but difficulties in the verbal versions prevented them from achieving higher scores. This study has clearly pointed out readability to be one of those difficulties. One does however realize that readability is not the only barrier preventing students from demonstrating what they know. Another barrier is the important mathematization skill. Whereas readability problems should be removed, mathematization is a skill which should be developed in classroom practice and accessed during mathematics examinations. However, the above table shows great differences between the test scores of the original and non-verbal versions *even for questions where no mathematization is required.* The differences between test scores of verbal and non-verbal versions have certain implications for assessment which will be discussed with other issues in the following chapter.

# CHAPTER SIX

# CONCLUSIONS AND IMPLICATIONS

## Rationale

*In the previous chapters, answers to various research questions were pursued. The following chapter is a summary of the conclusions drawn from readability research reported in Chapters 2 and 3 as well as the empirical studies discussed in Chapters 4 and 5. Results emerging from the literature and empirical research have implications for examiners in particular as well as examination practice in general.*

## Chapter contents

6.1 Broad outline of the present study

6.2 Conclusions and implications for writers of mathematics examination text

6.3 Readability checklist

6.4 General conclusions and implications for examination practice in a wider context

## 6.1 Broad outline of the present study

During the course of this study, the following questions were raised regarding the readability of verbal problems used during senior secondary mathematics examinations:

- What type of readability factors in the ordinary English of mathematics examination questions prevent a clear understanding of verbal problems?

- Are these factors restricted to second language readers only or do first language readers experience similar problems?

- If there are readability problems in the text, how do they influence achievement?

- What influence do these readability factors have on African readers whose home and school cultures are often quite different?

Answers to these questions were sought by approaching the readability of ordinary English from a cognitive and psycholinguistic point of view while linguistic and cultural issues referring to second language readers were also taken into account. Information generated by the literature study provided partial answers and generated guidelines for writing more readable mathematics text (cf. 2.5 and 3.4). However, new questions arose:

- Would linguistic factors that proved to influence the readability of ordinary English prose have similar effects on relatively short pieces of mathematics text - especially text written for examination purposes?

- Since improved readability did not always improve comprehension, would improved readability indeed have a meaningful effect on mathematical achievement?

- Do first and second language readers have different needs regarding the readability of mathematics text?

- Who would gain most by improved readability, first or second language students?

The above questions guided a protocol study to identify readability problems experienced by students while solving verbal problems in mathematics. The hypothesis that improved readability would improve test scores was tested for first and second language students.

Although not all research questions have been answered in full, this study has led to some definite conclusions and implications for examination practice in mathematics:

## 6.2 Conclusions and implications for writers of mathematics examination text

Conclusions are discussed by referring to readability factors in general (cf. 6.2.1 - 6.2.4) as well as to more specific, concrete examples (cf. 6.2.5 - 6.2.10). The influence of readability on achievement and the inability of readability formulae to guide examiners efficiently (cf. 6.2.11 - 6.2.14), lead to the suggested readability checklist in 6.3.

6.2.1 *The same readability factors causing comprehension problems in ordinary English occur in verbal problems of mathematics examination text. Therefore guidelines for writing more readable English could serve as guidelines for writing more readable verbal problems in mathematics as well.*

Analysis of the protocol study as well as the adaptations suggested by the students generated the same kind of readability factors that influence the readability of ordinary English. There was abundant evidence that students experienced the same type of problems in spite of the fact that the text of mathematical verbal problems is comparatively short and comprehension is aided by non-verbal information and mathematical cues. In a sense the results of this study were a confirmation, in the context of mathematics examination text, of results generated by contemporary readability research. The preliminary guidelines generated by the literature study and presented in 2.5 and 3.4 are therefore incorporated in the guidelines suggested for writers of mathematics examination text in 6.3.

6.2.2 *Students experienced more readability problems than were reported in the literature study of Chapters 2 and 3. Therefore writers of mathematics text should not only follow guidelines developed for the writing of ordinary English, but consider guidelines which are specific to mathematics examination text as well.*

Students experienced additional problems mainly because the text was focused on verbal problems in *mathematics* and more specifically, text used in examinations at senior secondary level. It has been mentioned before that not many research results are available on readability problems experienced at this level, especially where mathematics equations and formulae play an important role. (cf. 1.4). Most of the "additional" readability problems addressed by students were related to the non-verbal mode of the mathematics register or the overall structure of information (cf. 2.4.3.1; 4.6.2 and 5.2.3). Students' adaptations that were specific to verbal problems in mathematics text seem to suggest that a specific type of text necessitates its own level of readability (cf. Figure 2.3). Much depends on what the reader has to do with the information, how much time he has to do it in and what previous experience he has had of the type of text.

6.2.3 *Second language readers are even more dependent on highly accessible text than first language readers. Examiners should therefore take the special reading needs of these students into consideration when processing examination text.*

The protocol study proved that the effect of readability problems on second language readers was often more substantiated than on their first language peers (cf. 4.5). For example, core concepts and abstract information are not so easily recognized in a second language, making these readers more dependent on the overall structure of discourse. An important reason seems to be the more difficult access to a relatively weaker lexical network. Access to the lexical network is by means of sound and visual features (cf. 3.2.1). Low frequency words from a second language lexicon do not have the same activating power as words that are frequently used in the mother tongue. Furthermore, the important role of culture cannot be ignored (cf.

3.3). Certain second language readers, like the E2 group of this study, have the support of a related mother tongue and culture, whereas second language readers like the E3 group have neither. Moreover, many E3 students have not yet developed an efficient cognitive academic language proficiency (CALP) in English due to circumstances beyond their control (cf. 3.1). As indicated by Cummins & Swain (cf. 1986:156), it takes between five to seven years under optimal conditions to develop efficient second language academic skills. Examiners writing for E3 readers should therefore consider the E3 audience with even great care. Of course, not only the black South African student experiences the lack of support from a related language and culture. Any second language reader whose mother tongue is unrelated to the language of the text has the same problem. World-wide, more and more attention is being given to the special educational needs of E3 learners (cf. Bishop, 1988; Lagerwerf, 1992; Mulder, 1992).

Research on linguistic and cultural aspects of reading in a second language have identified important readability guidelines when writing verbal problems intended for second language examinees (cf. 3.4.1 and 3.4.2). Most of the issues were confirmed by the empirical studies reported in Chapters 4 and 5 (cf. 4.5 and Table 5.1). Although most of the readability guidelines apply to all types of readers, research has shown the following issues to be even more important when writing for second language readers:

1. Avoid misleading cognates, but use true cognates whenever possible (cf. 3.2.1).

2. Avoid long pieces of text in time-limited examinations. Rather alternate text and metatext (cf. 7.1.4).

3. Concentrate on the use of as much authentic language as possible i.e. language that the specific students use in *their* daily life. The principle of authentic language should cut across vocabulary, syntax and discourse.

4. Important information should be placed in superior positions of the text, for example in headings or at the beginning of sentences.

5. Text can be formatted in such a way that important information is accented strongly, for example by using bold letters or italics.

6. Use pictorials or other visual aids to make the information more concrete. If students are expected to draw a graph and the graph could aid comprehension, this instruction should be given as soon as possible in the metatext.

7. Contexts should be culture free and not privilege one group of students above another (cf. 4.5).

8. The hierarchical organization of discourse should aid logical thinking, therefore the structure of information input should support the construction of a contingent knowledge structure. (cf. 4.5).

9. Hypothetical constructions and hypothetical conditions should be avoided whenever possible (cf. Question 4; Dale & Cuevas et. al. 1987).

10. Readers with an unrelated culture or mother tongue, i.e. the E3 readers, need more explicit *and* more extensive information. .

11. E3 readers need more time to process written discourse in a second language, therefore more examination time should be allowed for these examinees.

This study has demonstrated the special reading needs of E3 readers. The recommendation for a suitable E3 writer to produce examination text for E3 readers is elucidated in 7.1.4.

One should keep in mind that there are first language students whose language proficiency is not so high and who also need a great deal of high frequency language. Likewise, one finds second language students who actually share the same cognitive academic language proficiency (CALP) with first language learners and are therefore similarly influenced by readability factors. Examiners in the South African context have a large range of readers to consider when preparing mathematics text which makes text production difficult in many ways.

6.2.5 *Inappropriate mathematical models encourage rote learning because the interpretation of these models leads to comprehension difficulties related to non-real behaviour. Therefore, examiners should take care that mathematical models generate behaviour that is true to real life.*

Five of the nine verbal problems had the mathematical models given and in four of the five cases there was a mismatch between the described situation and the given model (Questions 3, 6, 7 and 8). The fifth situation described a context that remained so abstract that one could not decide whether the described behaviour was true to life or not (Question 2).

Evidence from the protocol study has proved that students are able to score as much as 90% for a verbal problem even though they are baffled by information generated by an inappropriate function (cf. 4.6.3.4).. After a few unsuccessful interpretations, they look for mathematical cues or do what they think they ought to do and sometimes even manage to get correct answers. Inappropriate models do not encourage important mathematical skills like interpretation or reflection, but rather encourage students to operate according to fixed rules. Students will not persevere with reflection or the interpretation of a mathematical model if it only leads to comprehension difficulties or is time consuming, especially if they are often able to score well without subjecting themselves to these difficulties.

6.2.6 *Not only familiar contexts but also* ***familiar behaviour within the context*** *is important for total comprehension. Therefore examiners should be careful not to presuppose that students are acquainted with communicated behaviour.*

Research by Steffensen and her colleagues has emphasized the important role of familiar behaviour to support overall comprehension (cf. Steffensen et al. 1979; 3.3.2 ). In Questions 3 and 7 comprehension difficulties were mainly caused by the fact that many students were unacquainted with the presupposed behaviour (or conduct) of the variables i.e. the cost price, the selling price and the number of items sold. Apart from the non-real behaviour generated by the artificial functions, many students did not know that in real life it often happens that the more articles are produced, the less the cost price becomes *per article*, or that the selling price of one calculator can be influenced by the number of calculators sold. Decreasing profit in spite of increasing sales was even more unfamiliar. Instead of the real-life situation supporting mathematical thinking, the unfamiliar behaviour had a counter-productive effect. The necessary schemata could not be activated by the otherwise familiar context because it was not available - it had not been developed. *Interactive behaviour* between cost price, selling price and profit, all depending on the number of articles made and sold, was unfamiliar behaviour to many students.

Although an important objective of mathematics education is to extend the boundaries of a learner's known world, one cannot present new contexts and unfamiliar behaviour to a reader in examination conditions as if it were part of his known world. This would be unfair and a violation of the given-new contract (cf. 2.4.3.4). This finding is in line with the cooperative principle of Grice (1967). One of his maxims for successful communication is that one should only mention that which both parties concerned believe to be true and have adequate evidence for.

An example from a recent examination paper illustrates that the problem of unfamiliar behaviour is not only restricted to the questions used in the protocol study. Consider the following:

> *During an experiment the temperature, M Celsius changes with time, t seconds, according to the formula*
>
> $M = - t^2 - 6t + 49; \quad 0 \leq t \leq 10$
>
> *1. Determine the rate at which the temperature changes when t = 3*

2. *During which intervals will the temperature*

   *decrease?*

   *increase?*

3 *Determine*

3.1 *t if M is a maximum*

3.2 *the maximum temperature*

3.3 *the rate of change of the temperature at that moment.*

(House of Representatives, March, 1994).

Of course one can do the above question in a rote way if one looks for cues and follows the rules. The analysis of protocols proved that bright students are able to score almost 100% for questions like the above without really understanding what the context or the behaviour within the context is all about. But what if one does not know the rules or tries to *understand* the behaviour by asking a few questions like:

- *"What experiment is the examiner referring to?"*
- *"What kind of an experiment lasts only 10 seconds?"*
- *"What kind of experiment has such a peculiar temperature behaviour?"*
- *"What does 'rate of change of temperature' mean in this case?"*
- *"Why is it important to know what 'the rate of change of temperature' is?"*
- *"Why is the initial temperature so high? Room temperature is about 20 C... or not? Have I missed something?"*

If real-life verbal problems are supposed to

- relate education to the environment of *the learner*,
- illustrate to the learner how to use mathematics in *his* or *her* daily life,
- give meaning to various mathematical operations,
- be the source of a student's solution methods,

- develop mathematical skills like interpretation, reflection and a critical attitude,

then the above question fails totally. In fact, from a real-life perspective, the above verbal problem and most students' real-life experiences are totally different.

It seems that one of the main reasons for students' experiencing unfamiliar behaviour is the fact that the information in the text is too concise for real meaning. As in the case of inappropriate functions and insufficient time, students overcome the comprehension difficulties by reverting to "rules without meaning" - with apologies to Skemp (1976). Much can be done to overcome this problem by using questions in the metatext to elucidate the text (cf. 7.1.5).

6.2.7 *Words or phrases that are familiar to specialist readers like mathematics examiners, are not necessarily those frequently used by examinees. Writers should therefore be careful not to presuppose the familiarity of words or phrases, especially those that are crucial for the solving of the problem.*

This applies to ordinary English as well as the vocabulary of mathematics or other science registers. For example, the words *cost price* (Question 3), may be commonplace to examiners, but to students who in daily life have little contact with economic concepts, the word *cost* is more commonly used in the sense of, "What does the article *cost*?" and in the talk-aloud experiment evoked the concept of *selling* price. Not only contexts, but also words and phrases should be culture-free - they should not advantage one group above another.

This principle could be problematic to examiners. Words and phrased related to certain contexts are often culture-specific. If examiners and examinees differ in culture, this principle could leave examiners with few contexts to choose from. To overcome this problem, more extensive information about the topic as well as appropriate visuals could aid comprehension. Examiners should be sensitive to the possibility that words and phrases could be culture strange to some of their readers.

6.2.8 *Not just single words and phrases, but also the way in which language is used as a whole should be in line with the way students normally speak about everyday experiences. Examiners should therefore communicate everyday experiences in everyday language.*

Examiners are specialists in their field, so they speak the language of mathematics as if it were their native tongue (cf. Pimm, 1989). To them, speaking "mathematically" about an everyday experience has become a natural way of seeing the world. They would not find a sentence like the following strange: *The sum of the number of hours that the pumps operate per day must not exceed 20* (cf. Question 1). Furthermore, sentences mixing verbal and non-verbal information like, *The cost of producing n calculators is* $C = 100n + 200$*, and n calculators can be sold per week at a price* $P = 400 - 0.02n$ *per calculator*, have become part of their everyday language use. However, students at school level do not experience their world in this kind of language.

This does not imply that the language of mathematics should not be taught. As Gullat (1986) points out, control of the mathematics register should be one of the important objectives of mathematics education. It is, however, unnatural *and* unnecessary to use mathematics language when describing real-life events to a non-specialist reader. Verbal problems in mathematics should bring students' daily life into their school environment by means of everyday language. If not, comprehension is not optimized and everyday experiences might not be recognized as such.

6.2.9 *Not only the* ***text,*** *but also the* ***metatext*** *(the questions) as well as the* ***solution*** *of a verbal problem interact to determine the level of readability. Writers should therefore evaluate the readability of a verbal problem by ascertaining whether all three these components support total comprehension.*

The analysis of protocols proved that comprehension difficulties were not only caused by the text but also by the metatext (cf. Question 2.4) and the answers (cf. Question 6.2). Uncertainty experienced in any of these sections caused students to think they had missed important information and they were inclined to revert to comprehension strategies like rereading or reformulation, causing loss of time and increased anxiety. One can therefore not comment on the readability of a verbal problem if one has not worked through the metatext to ascertain whether the information communicated by the text, metatext *and* solution is meaningful to the reader. Even the number of marks allocated for a question can result in uncertainty. For example, if a relatively easy question counts many marks an examinee could think he has misunderstood something.

6.2.10 *The ability to apply mathematical knowledge to solve problems in real-life situations needs more time than normally allowed for in time-restricted examinations. Therefore examiners should find a more realistic time arrangement for these types of questions than is present practice.*

Lack of time was not so evident in the protocol experiment as it was in the experiment used to test the hypothesis. It must be remembered that the composite test was written by *all* Std 10 HG students of a specific school and not just by bright students as in the talk-aloud experiment. Although the tasks were not time-limited, they had to be completed in normal teaching periods. However, the arrangement allowed for more time than is normally allowed under examination conditions. In spite of more time than usual, there were many students who did not attempt all questions, nor did they have enough time to evaluate solutions. (The scores of students who did not attempt all questions were not used in the statistical analysis).

Second language learners, especially E3 students, seemed to have more of a time problem. Even in the protocol experiment, most E3 readers needed as much as twice the amount of time than even the E2 students. Initially, the schools of E3 students were asked to set aside as much time for the composite test as was necessary in the schools of E1 and E2 students. However, when the tests were returned from the first E3 school, all the tests were far from complete. The teacher responsible said that the time allocated had been far too short - she had not been aware that students were allowed to use more time if necessary. She said, "*The time was not enough. Students got so upset that they started to jump up and down and shout because of all the anxiety.*"

Lack of time is one of the readability factors discussed in 1.8.1. As soon as time starts running out, anxiety sets in and comprehension and creative thinking are seriously impaired (cf. Smith, 1982:12).

For verbal problems, additional time is needed to execute processes not normally followed in ordinary non-verbal problems. As indicated, the correct contingent knowledge structure must be constructed before the actual analysis of the verbal problem can take place. This relates to the dual purpose of mathematics reading discussed in 6.2.13. Limited time is one of the major reasons why examiners should write highly accessible text. Readers must be able to move into the text as quickly as possible. Furthermore, metacognitive processes like comprehension, interpretation, reflection and mathematization are time consuming. There are researchers in mathematics education who are convinced that the goals of realistic mathematics education cannot be realized effectively in time-limited examinations (De Lange, 1987:170). This matter is discussed further in 7.1.7.

Time-limited tests and examinations are more appropriate for the testing of *products* than *processes*. Like artificial functions, time-limited tests tend to encourage rote learning. As soon as students realize there is not enough time for proper thinking, they tend to look for solutions to this problem by depending on rules that can be applied in as little time as possible.

6.2.11 *Although improved readability improved test scores, achievement is not the only important area influenced by readability factors. Therefore writers should formulate as accessible a text as possible, even though achievement levels may not be significantly affected.*

In this study, improved readability improved test scores in all questions except Question 4. When considering the students as one language group the increase in test scores proved to be significant in six of the nine cases. On the other hand the study did reveal that test scores of exceptionally high or low achievers did not always benefit by improved readability. However, the necessity of highly accessible text is not only important because of the influence on test scores. Not all effects of readability problems can be quantified. The analysis of protocols gave enough evidence of readability problems upsetting the cognitive and emotional disposition of students without noticeably affecting their performance. Irrespective of comprehension difficulties, readability problems proved to

- increase reading time
- decrease reading ease
- increase anxiety
- encourage rote learning
- develop a wrong attitude towards mathematics
- question the credibility of mathematics.

One should remember that the tasks of the students were not time-limited, and that the test scores were not of great importance to the students. In authentic conditions the effect of low accessible text could be even more profound and needs to be researched. A suggestion in this regard is discussed in 7.3.5.

6.2.12 *The significant difference between test scores of the verbal and non-verbal versions necessitates a more positive approach towards testing.*

The main reason for including the non-verbal versions of the nine questions in the composite test was to ascertain the difference in achievement levels between verbal and non-verbal versions. The different versions of the same question required students to demonstrate the same manipulative skills. Of course the verbal versions required other mathematical skills like mathematization, interpretation and reflection, but in relation to the total marks allocated per question, most marks were given for the execution of manipulative skills. There was a need to know how the verbal form of a problem prevents a student from demonstrating other mathematical abilities.

Table 5.6 illustrates the marked difference between test scores of the verbal and non-verbal versions. The average achievement level for the non-verbal versions of the nine questions was 80% whereas the original and adapted verbal versions had an average of 47% and 62% respectively. Improved readability was able to improve the situation to a certain degree, but if one takes into consideration that students did not have to formulate an appropriate mathematical model for the real-life situations in Questions 2, 3, 6, 7 and 8, the differences in scores for these questions still remain alarmingly high. The average difference between the scores of the non-verbal and adapted verbal versions for these questions is as high as 22,8%.

The significant difference between the scores of verbal and non-verbal versions confirms the fact that even if one understands a verbal problem and can execute all the necessary calculations, this is no guarantee that one can solve the problem. There are more elements in the problem solving model than only *understanding the problem* and *executing the plan* (cf. 1.5).

It is clear therefore, that expecting students to apply their mathematical knowledge in real-life situations often has the unfortunate effect of preventing them from demonstrating their manipulative skills. If one takes into consideration that most of the marks in the verbal versions are allocated for these skills, one realizes that the problem evident from Table 5.6 needs to be addressed.

This aspect links up with one of the principles of testing advocated by Cockroft (1982:159) and De Lange (1987:180) which states that achievement tests should give students the opportunity to demonstrate what they know rather than what they do not know - the so-called positive approach to testing. Applying this principle to the assessment of manipulative skills, one could say that achievement tests should not make the execution of these skills dependent on the skills of mathematization. For example, if the goal of a question is to test the linear programming algorithm, one could not make the execution of this skill dependent on the restructuring of difficult discourse in order to formulate an appropriate mathematical model. In this case the necessary formulae should be given or the text should be structured in such a way that it aids not only the construction of a CKS, but also mathematization (cf. 4.6.2.1).

On the other hand, one should preferably not test manipulative skills out of context. The analysis of protocols showed that real-life contexts guide operational skills. Students are more able to pick up manipulative errors if they can interpret their results in a real-life situation. For example, an answer of -0.25 indicates an error if this number represents the number of books in a classroom. Research by Clarkson (1992) confirms this finding by concluding that students are able to correct many errors in a mathematical language context if given a fair chance to do so. A familiar real-life context provides students with such a chance. So if most marks are allocated for manipulative skills, as in Question 4, one could provide the mathematical models and describe the real-life context before students are expected to execute the necessary calculations.

6.2.13 *Examination text in mathematics must be processed for a dual purpose: for ordinary comprehension as well as for mathematical meaning. Therefore the necessity to produce highly accessible examination text is doubly important.*

Meaning is the common element in both reading and mathematics (McKenzie, 1990). A mathematics examination paper is an example of what Flower and her colleagues (1983) call a functional document - readers *have to do something* with the information, they do not read it just for the sake of reading. But unlike most other functional documents the *instructions* of *what to do* as well as *how to do* it, are not given with mathematical verbal problems. Before a student can start operating mathematically on a verbal problem he has to construct a knowledge structure of the written information. Before he can understand *what to do*, he first has to understand *what is given*. Exact reasoning calls for exact meaning. Correct solutions are therefore dependent on a correct understanding of the *structure* of the problem. That is the most likely reason why all the adaptations had one common aim: the removal of barriers preventing the quick and correct construction of a knowledge structure of the problem so that they could decide how to apply their mathematical knowledge to solve it. In a special sense writers of functional documents like mathematics examination papers must write doubly well: for ordinary comprehension but also for mathematical meaning.

6.2.14 *Readability formulae are not able to sufficiently guide examiners in producing more readable mathematics text. Therefore guidelines other than those suggested by readability formulae should be used when writing mathematics text for examination purposes.*

This study confirmed the issues raised by researchers like Selzer (1983) and Strother (1990) regarding the use of readability formulae. Most readability formulae concentrate on factors like short words and sentences and are unable to address factors related to overall comprehension of mathematics text, like proposition density (5.2.2), artificial functions (4.5.3) or entangled verbal/non-verbal information (4.5.2). They were not designed for this purpose. When solving mathematics problems, short words and sentences are not as important as a clear understanding of the problem. A short sentence with much mathematical information is more difficult to process than a longer sentence with practically no mathematical detail. For example, consider the first sentence of Question 1 in the original version. There are about ten propositions of which six are related to mathematical detail. This requires quite a high level of processing for a sentence of relatively average length.

- *Two* pumps
- X and Y : names of pumps as well as the suggestion that the variables are connected to some or other characteristic of the pumps.
- at least
- 60 000
- per day
- litres (not kilolitres)

Comprehension ease and not brevity should be the key word for the writing of verbal problems in mathematics. The following paragraph presents guidelines which could be considered for this purpose.

## 6.3 Readability checklist

The present study has generated a variety of readability guidelines. These guidelines have been used to compile a checklist which can be used to formulate mathematical verbal problems for assessment or other purposes.

The checklist was developed by using three main sources of information:

- the results of readability research reported in Chapters 2 and 3
- the data generated by the protocol study (cf. Chapter 4)
- the adaptations suggested by the students (cf. 5.1.2).

Irrespective of the checklist, the analysis of the empirical studies has pointed out that all writing should constantly be guided by the following important considerations:

- the aim of the question
- the mother tongue and culture of the examinees
- the time allowed for the solving of the problem
- the goals of the syllabus as realized in teaching practice.

The checklist, in the form of guidelines, cuts across all linguistic levels: lexis, syntax and discourse as well as non-verbal features. All guidelines marked by an asterisk (*) are specific to this study and were mainly identified during the talk-alouds and the analysis of the students' adaptations.

*Lexical features*

1. Avoid difficult vocabulary.
2. Avoid unfamiliar words or phrases be they in ordinary English or a relevant scientific register.
3. Use as much natural, authentic language as possible.
4. Consistently use the same words for the same concept.

*Syntactical features*

1. Curb the proposition density per sentence (cf. 2.4.3.4).
2. Use action-specific sentences (cf. 5.2.2).
3. Structure the sentence to aid the construction of a CKS (cf. 2.4.3.4).
4. Avoid hypothetical conditions and hypothetical constructions whenever possible.
5. Use the subject position of a sentence to re-instate important information in the short term memory of the reader (cf. 2.4.3.4).

6* Use separate sentences for text and metatext.

7* Avoid unnecessary parentheses (cf. 2.4.3.2).

*Discourse features*

1. Structure the discourse to aid the construction of a CKS (cf. discourse features in 5.1.2).
2. Curb the proposition density per paragraph (cf. 5.1.2).
3. Place important information in superior positions.
4. Use a topic sentence as an introductory sentence to introduce the context (cf. 2.4.3.4).
5. Use a heading for a verbal problem to act as an advance organiser (cf. Ausubel, 1963).
6. Use pictorials for every verbal problem to communicate the real-life context.

7* Information-sentences (text) should precede question-sentences (metatext), i.e. first synthesize a mental representation of the problem and then pose the question (cf. the given-new contract, 2.4.3.4).

8* Avoid long pieces of extensive discourse before posing the metatext.

9* Alternate text and metatext so as to communicate longer pieces of discourse more effectively (cf. 7.1.5).

10* Remove redundant information if teaching has not constantly subjected students to this type of information (cf. 4.6.3.2 and 5.1.2).

11* Communicate like information all in one paragraph (cf. adapted version, Question 1).

12* If students are expected to draw a graph, pose this question as soon as possible to aid visualization.

13* Use more explicit and extensive information if the language or the context is not firsthand.

14* Not only contexts but also behaviour within the context should be familiar to the readers. Elucidate less familiar contexts or foreign behaviour by using the suggestions in 7.1.5.

15* Avoid abstract contexts as in Question 2, for example.

16* Avoid senseless solutions.

17* Text, metatext and answers should all further readability.

*Non-verbal features*

1* Avoid misleading or ambiguous use of letter symbols (cf. 4.6.5.2).

2* Avoid inappropriate mathematical models (cf. Questions 3, 6 and 7).

3* Do not semi-mathematize the verbal information.

4* Keep verbal and non-verbal information separate (cf. Question 3, original and adapted version).

5* Do not presuppose that behaviour generated by mathematical models is familiar (cf. Question 7).

6* Use function notation where applicable.

All guidelines have one common aim and that is improved readability. Writers should therefore not apply the guidelines as rules, but use them with discretion. The readability checklist is not a finished product, but the beginning of a process. As new insights into the problems of readability emerge, the checklist can be adapted accordingly.

## 6.4 General conclusions and implications for examination practice in a wider context

Readability and how this influences the fairness of examination papers also relates to a variety of other factors not formally researched in this study. However, certain results of the empirical study touched upon some of these issues and point to certain implications. Although the reported research clearly underlines the need for measures as those suggested below, the suggestions are not offered as research based solutions.

Two main issues are presented for discussion. They are related to the congruency between teaching and examination practice. Although congruency is advocated, this does not imply that examinations should endorse textbook shortcomings.

6.4.1 *Readability problems are often the result of a mismatch between assessment and teaching practice. Examiners should therefore ensure that examination questions are in line with instructional objectives as reflected in the instructional material of textbooks.*

Vague syllabus objectives often cause a mismatch between examination and teaching practice. It could happen that the method of a textbook and the way the examiner interprets the syllabus differ and various types of readability problems could occur during examinations. For example, during the protocol analysis and the tests written by students, it was clear that most students were not able automatically to link what they had been taught about *average rate of change* to *average speed of A over the first two hours* or to *how much the circulation will change during the 2nd year.* (cf. Questions 6.1 and 8.4, original versions). When looking into a commonly used textbook (De Jager, et. al. 1986) one can understand why. The concept of calculus is developed in a purely mathematical context. Sixteen pages of explanation with more than 70 questions are posed in a graphical context *before* rate of change is related to applications in everyday life. A meagre two pages and only six rather pointless examples are used to illustrate the application of calculus to the rate of change in daily life. Even then, *rate of change relative to time* is not referred to as *speed*, a word that students use frequently in their daily lives. Nowhere is the concept, ***average** rate of change* taught in terms of contexts relevant to the real life of students. One can therefore understand why Questions 6.1 and 8.4 were so poorly answered. If textbooks and, consequently, teaching stress the *control* of techniques and the corresponding algorithmic skills at the expense of *insight* into the phenomenon of calculus in the context of real- life experiences, it is unfair to test this insight in time-restricted examinations. An issue closely related to this matter is discussed in 6.4.2.

In South Africa, objectives for std 10 HG mathematics are mainly concretised in the method which textbooks use to present the relevant instructional material. Subsequently teachers as well as students depend on the method of the textbook to interpret the rather non-specific aims set out in the std 10 HG Mathematics Syllabus (cf. Syllabus for Mathematics Higher Grade, 1985; 1991). Examiners have the responsibility to develop examination questions that measure clearly defined learning outcomes that are in line with instructional objectives, a principle of testing advocated by many leading educators (cf. 2.1). This is a difficult principle to realize if thousands of examinees are taught in hundreds of different schools spread across large areas of land as in South Africa. Actually, the external examiner of a std 10 HG mathematics examination has no other source but the textbook to determine the nature and content of contemporary teaching practice in a specific area.

To be able to draw up a fair examination paper, examiners should carefully go through all relevant textbooks to ascertain what methods have been taught. Then a method common to *all* textbooks should be accepted as basis for examination practice so as to give *all* students a fair chance. Teaching practice should guide assessment, not the other way around. If teaching objectives are not reflected in textbooks the way examiners have in mind, an examination paper is not the correct place to address the issue. Shortcomings and vague syllabus objectives should be exposed differently. A more honourable way would be to use an examination programme as a means of communicating any concern (cf. 7.1.2). Examiners should keep in mind that the main aim of the final examination is not to improve learning, but to assess grades. These grades are extremely important to students as they are the key to further education facilities or job opportunities. In a final examination all students have the right to demonstrate what they have been taught, not what they have not been taught. No matter how "readable" an examiner may try to write, if the problem is a mismatch between the objectives pursued in teaching and examination practice, the access to examination text will remain difficult and scores will be affected accordingly.

Suggestions to realize a better congruency between examination and teaching practice are given in the next chapter. These suggestions not only address the issue of readability, but also teaching and examination practice as a whole.

6.4.2 *If mathematical knowledge is assessed in real-life contexts, but was developed in a purely mathematical situation, readability problems are likely to occur. Therefore, if students are to relate what they have learnt to their real-life environment, mathematical knowledge should be developed within the context of real-life situations.*

This aspect relates to the previous one, a mismatch between teaching and examination practice.

From a reading perspective one could say that if mathematical knowledge is developed in a *mathematical* context, then language communicating a *real-life* context does not have a good chance of activating the necessary mathematical schemata in examination conditions. Students will rather associate *mathematical* concepts with *mathematical* situations because that was the way the concepts were developed. When reading verbal problems, semantic contexts as well as mathematical concepts must be activated. If real life and mathematics were not successfully linked during teaching practice, the link does not automatically realize during assessment.

It has been pointed out in 6.4.1 that the quoted textbook extensively develops the concept of calculus in a purely mathematical context *before* applying the mathematical knowledge to real life in only a few examples. Even then *rate of change relative to time* refers to geometric figures like cubes, circles and squares and consequently the real-life attempt remains non-real to many. As Freudenthal says, *What is worth being taught? In order to be taught it must be applicable.....what does this mean? Teaching as much mathematics as science teachers pretend they need? Or after a block of compulsory algebra and calculus a few choice subjects like probability, numerical methods, linear programming or mechanics? Everybody knows that it does not work. The right perspective is primarily from environment towards mathematics rather than the other way round. Not: first mathematics and then back to the real world, but first the real world and then mathematics* (1983:5).

A good example to confirm Freudenthal's words that *first mathematics and then the real world does not work,* can be found in the low test scores generated by some of the sub-sections of the verbal versions of Questions 2, 3, 6, 7 and 8. This was surprising because the necessary mathematical formulae were given and test scores of the non-verbal version proved that students had the necessary operational skills at their disposal. Test scores of the original, versions of the above questions were, on average, 40% lower than those of the non-verbal versions.

On the other hand, the test scores of Questions 1 and 5 demonstrate how much more successfully mathematical knowledge can be applied in real-life contexts if the knowledge has been developed in this manner. Questions 1 and 5 test the students' knowledge of linear programming, one of the few sections in the Std. 10 HG syllabus where the underlying mathematical concepts are mainly developed in real-life contexts. In spite of the fact that mathematization is required in these two questions, the average difference between the original and non-verbal versions is only 19% - less than half the average difference measured in Questions 2, 3, 6, 7 and 8 where the mathematical formulae were given. Furthermore, improved readability was able to improve the test scores of Questions 1 and 5 in such a way that there was only a difference of 2% and 5% respectively between the test scores of the adapted verbal and non-verbal versions (cf. Table 5.3). Question 4 also required students to formulate the necessary mathematical model as for Questions 1 and 5. However, in Question 4 students had to apply their knowledge of simultaneous equations in a real- life situation, but the knowledge had not been developed in this manner. There is a difference of 40% between the test scores of the original and the non-verbal versions. Although other factors contributed to this situation as well (cf. 5.2.3.1), the manner in which the necessary mathematical knowledge was developed could have aggravated the situation.

An important contributing reason for the low test scores of the verbal versions of Questions 2, 3, 6, 7 and 8 could therefore be the fact that the concept of calculus had not been successfully integrated with the real-life experiences of students. The teaching method did not allow for this so the assimilation of mathematical knowledge in the schematic network of real-life experiences was poor. It could be that to many students, calculus and real life had very little in common. When reading the above questions the familiar real-world situations were therefore unable to activate the necessary mathematical schemata effectively (cf. 2.3.1 and 2.3.2). If a familiar real-life context is activated, but the necessary mathematical concept cannot be associated with real life, the concept is not available for problem solving.

One could therefore conclude that if mathematics knowledge is developed in a *non-real* setting, examination text communicating a *real-life* setting cannot effectively activate the necessary mathematical concepts, even though the schemata of the concepts have been developed by teaching practice. The implication following this conclusion is clear: If one wishes to realize the objective of developing students' abilities to apply their knowledge of mathematics to solve relevant real-world problems as the std 10 HG mathematics syllabus (1991) suggests, one should *develop* this knowledge in the context of the real world. This approach will not only make the text more accessible to students, but also links up with the rationale underlying realistic mathematics education (cf. 1.1).

It has been demonstrated that the readability of senior secondary examination papers is a multifaceted problem that needs to be addressed in many ways. The important role of comprehension in problem solving as advocated in Chapters 1, 2 and 3 has been confirmed, therefore the concluding chapter will present further suggestions for examination practice and readability research.

# CHAPTER 7

# SUGGESTIONS FOR EXAMINATION PRACTICE AND FURTHER RESEARCH

**Rationale**

*To conclude the study, various suggestions are made for examination practice as well as further research. Although not all suggestions are research-based, they are informed by results generated by the present study.*

**Chapter content**

## 7.1 Suggestions for examination practice

Conclusions discussed in the previous chapter have pointed to needs in a variety of assessment-related areas. Suggestions are made to address these needs by referring to possible arrangements that could be made within departments of education (cf. 7.1.1 - 7.1.4, i.e. macro-level arrangements) as well as recommendations referring more specifically to examiners (cf. 7.1.5 - 7.1.8, i.e. micro-level recommendations).

7.1.1 *A communication network between all members of the educational team could be developed to realize a better congruency between teaching and examination practice.*

Better communication between various members of the educational team should lead to a better congruency between teaching and examination practice. The aspects discussed in 6.4.1 and 6.4.2 demonstrate the need for more effective co-operation between *all* parties concerned with the curriculum. Better co-operation will not only serve to improve the readability of examination text, but should also contribute to better mathematics education in general. The efficiency of the network depends to a great extent on the way the curriculum is initially developed. A more successful mesh between teaching and examination practice could be achieved if all members of the educational team, i.e. teachers, students, researchers, examiners, textbook writers and curriculum planners developed the curriculum in close co-operation *from the very beginning.* This could alleviate the problems addressed in this study. Definite answers to crucial aspects like the following could be pursued:

- what are the aims? (clarifying the goals)
- what is to be taught? (clarifying the content)
- how much is possible? (clarifying the range )
- how should it be taught? (clarifying the methodology and instructional material)
- what is to be assessed? (clarifying the content)

- how should it be assessed? (clarifying the goals and methodology of assessment)

If all partners work together on these aspects from the very beginning, there should inevitably be a better match between teaching and examination practice. All parties then share a common body of curriculum knowledge. In the process of collaborative planning, teaching practice should guide examination practice. What happens in the classroom should show the way to what objectives could possibly be realised, time-wise and cognitive-wise. This suggests that students become equal partners of the curriculum team and the classroom becomes an essential laboratory in the educational development programme.

*7.1.2 A document known as an "examination programme" could serve the purpose of linking teaching and examination practice more effectively.*

In South Africa there is at present very little real contact between the different persons responsible for the various aspects of the mathematics curriculum. Curriculum planners, authors of textbooks, examiners, teachers and students have no effective, common forum to realize better coherence between what each one is doing. A large gap often exists between the intended and delivered curriculum and consequently many problems occur due to a mismatch between teaching and examination practice, as this study has shown (cf. 6.4.1). The above suggestion (cf. 7.1.1) could do much to bridge the gap. In the meantime, while hoping for a more satisfactory form of educational development, an examination programme could be compiled to act as a link between the demands of the classroom and the examination paper (cf. the Dutch "Examenprogramma", 1989).

*All* members of the curriculum team, especially examiners, teachers and textbook writers should be involved to identify issues concerning the interpretation of the syllabus and range of the examination. Examination programmes could be updated if necessary. This would give all members of the educational team the opportunity to interact on matters regarding the intended curriculum. For instance if examiners are not satisfied with the method of a specific textbook, a discussion can take place and the outcome could be incorporated into an updated examination programme. If writers do not update their textbooks accordingly, schools could be informed of the shortcomings or deficiencies.

Some of the matters that need to be addressed in such an examination programme are the following:

- **The aim of the syllabus and the examination.** For example: *the syllabus has a general, formative character and is aimed at the use of mathematics in society. This implies that the examination will have the same aim.*

- **A *general* description of knowledge and skills needed to reach the above-mentioned goals.** For example: *The student must learn to analyze a newspaper article containing mathematical presentations critically,* or: *The student must learn to use computer programmes to solve problems.*

- **A *specific* description of knowledge and skills related to subject matter.** In this section it is important to state explicitly what the students should be able to do in each subject area. Note that this is not the content of the subject matter, but a description of the *specific* knowledge and skills that are needed in each subject area to reach the initial aim of the curriculum programme. For example, referring to tables: *The student must be able to relate a table to a graph, formula or text,* or: *The student must be able to set up a table using other tables, a formula or text.*

- **Examples of the way in which the subject matter can be assessed.** In this section examples of examination questions should be given for each subject area to illustrate in what way the examination questions are in line with clearly defined learning outcomes.

- **Limitations**. Limitations should be set for each subject area. This is an important aspect of the examination programme, because it guides the examiner, teacher and textbook writer. While a teacher is free to teach *beyond* the limits of the examination programme, the examiner is strictly bound by them. If he exceeds the prescribed limitations, any of the readability problems mentioned in 4.6 could occur. Of course, problems other than those related to readability are just as liable to occur. Limitations curb the range *and* difficulty level of subject matter.

*7.1.3 Examination papers could be given to a body of screeners to check for various assessment problems, especially the time factor.*

It is often not possible to remove all readability factors causing comprehension problems. The large corps of examinees has many individual differences and therefore what is clear to one examinee may be totally confusing to the next. However, the protocol study proved that given enough time, students are able to overcome many comprehension problems themselves.

To aid examiners in their quest for a fair examination paper, a number of competent people, say five in total, could act as screeners to check for various aspects of the examination paper and intercept possible problems before the examination paper is finalized. These people act as a test group. The examiner uses them to pre-test the preliminary paper. But they also watch over the interests of the students. They would get the preliminary examination paper *without a memorandum* and work through all the questions, independently of one another, with instructions like the following:

- Answer the whole examination paper in the same way you would expect examinees to answer it.
- Make a note of how long it took to work through each question as well as the total time used for the whole paper. (If screeners use more than half the time allocated to examinees, the examination paper is too long.)
- Check for readability problems in the text, metatext and answers. (The readability checklist in 6.3 could guide the screening for readability if the *guidelines* are changed into *questions*.)
- Check for congruency between teaching and examination practice.
- Identify problem areas.

Examiners could use the individual reports to adjust the examination paper where necessary before the final paper is handed over to the moderators for approval.

7.1.4 *Examiners with a deep understanding of the culture and language of the examinees should be involved in writing examination papers to prevent unnecessary comprehension barriers.*

The test scores for the non-verbal versions of all questions were relatively high and practically the same for all language groups. However, the relatively lower test scores of the E3 students for the adapted versions suggest that language and cultural differences had not been addressed thoroughly enough. Although E3 readers responded exceptionally well to the adaptations done mainly by E1 and E2 readers the average difference between the adapted and non-verbal versions is the highest for E3 students, namely 24% (cf. Table 5.6). The fact that scores of E3 students for the non-verbal versions are just as high as those of the other two groups makes one realize that the E3 students had all the manipulative skills available to perform just as well as the E1 and E2 groups on the adapted versions.

Of course, factors other than readability could be responsible for the relatively weaker performance on the adapted versions. A factor usually mentioned in verbal problems is mathematization. However, if one were to omit the questions where mathematization was required, the E3 students still performed less successfully on the verbal versions than the other second language readers. Table 7.1 is an extract of Table 5.6 and demonstrates the situation:

**Table 7.1** **Percentage test scores of E2 and E3 groups for questions needing no mathematization**

| Question no | Original version | | Adapted version | | Non-verbal version | |
|---|---|---|---|---|---|---|
| | E2 | E3 | E2 | E3 | E2 | E3 |
| 2 | *50* | *40* | *60* | *51* | *75* | *73* |
| 3 | *28* | *22* | *83* | *49* | *83* | *75* |
| 6 | *59* | *51* | *71* | *61* | *89* | *88* |
| 7 | *57* | *47* | *65* | *53* | *81* | *89* |
| 8 | *57* | *61* | *74* | *80* | *97* | *97* |
| **Average** | ***50*** | ***44*** | ***71*** | ***59*** | ***85*** | ***84*** |

If one considers the memoranda of the adapted and non-verbal versions of these questions, they are virtually the same. Marks were mainly allocated for manipulative skills which the E3 readers had on the same level as the E2 students. The main differences between the two versions are related to language and contexts.

Three suggestions are made to address this issue:

1 Examination papers should be written by examiners who share the same culture and language as the examinees. Only then can *subtle* linguistic and cultural differences that could possibly cause readability problems be prevented. If this is not possible, the examination text should at least be checked for possible readability problems by a person sharing the same language and culture as the examinees.

2 E3 examinees could receive a bilingual examination paper as is present practice for E1 and E2 examinees. Examination papers could be made available in the language of instruction as well as the students' mother tongue which is usually one of the official languages of the region.

3 This study has demonstrated the distinctly different reading needs of the E3 reader - a second language reader with an unrelated mother tongue and culture. Therefore, without changing the mathematical demand, separate examination papers should be compiled for distinctly different cultural groups with respect to context and language.

7.1.5 *Examination text could be made more accessible to students if the metatext were used to elucidate the text.*

One of the readability problems which is more difficult to address is that caused by less familiar behaviour or conduct of variables within the context (cf. 6.2.6). Students who are unaccustomed to certain cultural behaviours need more extensive information to understand the communicated situation. This could result in longer pieces of discourse which could cause other readability problems in time-restricted examinations - especially to second language readers.

One way of solving this problem could be to use the metatext to generate more information about the chosen context. Introductory questions could be used to make a verbal problem come alive to the reader. Elucidation of text by the metatext could therefore aid examiners in their quest for shorter pieces of text without the risk of being too concise.

For example, in Question 1, introductory questions could help students to identify the number of hours as the independent variable. The following introductory questions illustrate the suggestion:

- *Both pumps work for 8 hours a day.*
  *How many litres of water are pumped to the dam?*

- *The electric pump works for 7 hours a day and the diesel one for 14 hours a day.*
  *Is enough water pumped to the dam?*
  *Have the pumps kept to the working restrictions?*
  *Explain your answer by referring to the various restrictions.*

Another way of using the metatext could be to alternate short pieces of text with metatext. In this way students could better understand and experience the behaviour within the context. Much information could be communicated without students getting lost in long pieces of text before the metatext is posed.

The suggested *introductory* questions of the metatext should be concrete. Students should first be allowed to experience the real-life situation concretely before moving to the more abstract mathematics. It is therefore suggested that just as in teaching practice, examination questions should follow the order: from concrete experiences in the real-world to the abstract world of mathematics, *"first the real world and then mathematising"* (Freudenthal, 1983; cf. 6.4.2). This implies that the hierarchy of the metatext serves the purpose of guiding the student from the concrete level of his real-life experiences to the abstract world of mathematics. This could be especially important to second language readers who find it difficult to comprehend abstract concepts in a non-native language (cf. Hamers & Blanc, 1989:95; 3.2.1). Moving from the concrete to the abstract could be of great help to these students.

The hierarchy of the metatext should also allow for information to be generated in the same order as necessary to form an image of what is happening within the context. The metatext actually continues the construction of the CKS initiated by the text. Simultaneously the metatext helps to augment the initial problem space as more becomes known about the problem (cf. Newell & Simon, 1972; 3.2.3).

7.1.6 *Examination papers could be compiled in such a way that students operate within fewer contexts than is present practice.*

In South Africa, students writing a std 10 HG mathematics examination are expected to move quite rapidly from one context to the other. For various reasons this does not seem to be desirable. For readability this implies that examinees have barely constructed one knowledge structure before having to move out of the context and constructing the next. Although the ability to move rapidly from one context to the next might be a worthwhile objective to pursue, this could be problematic in time-restricted conditions. The analysis of a recent examination paper illustrates this phenomenon:

Examinees were asked to apply their mathematical knowledge in *thirteen* different contexts (CED, 1993). Five were real-life and eight were purely mathematical. Apart from this, *fifteen* other questions tested mathematical knowledge out of context. The composition of this examination paper is typical of most other assessment programmes. It causes a rather rushed movement from context to context to no context at all. *Students therefore cannot benefit optimally from the intended support given by the context or structure of the problem.* In the process one of the fundamental reasons for realistic mathematics education is ignored (cf. 1.1). Before students can really experience what is happening within a context, they have to move on to the next question - be it in or out of context. They have very little time to experience the situation to their own benefit.

A means of addressing this problem could be to opt for a thematic approach during assessment. Fewer contexts are chosen, but more mathematical knowledge from various parts of the syllabus is tested within the same context. The context acts as the binding factor. Another advantage of a thematic approach is to interrelate different parts of mathematical knowledge, thus illustrating the coherence within the syllabus.

7.1.7 *When time is not limited, other means of assessment could be advantageous to the readability of verbal problems.*

If time is limited, readability problems limit the time even more and the corresponding increase in anxiety has the snowball effect of making text even less accessible (cf. 1.8.1). Unfortunately, verbal communication will always have the potential of being misunderstood. Whereas a non-verbal problem can be posed in an international, precise mathematics language (a language that students are expected to learn), a verbal problem has to be set in a language and context that takes the linguistic and cultural aspects of the reading audience into consideration. When writing for a large and diverse audience this is difficult. It is virtually impossible to presuppose the linguistic and cultural framework of every reader and unforeseen readability problems are likely to occur. Alternative testing methods, *where time is not limited*, could be advantageous not only in the area of readability, but in other areas as well. The following can be considered as main advantages:

1. Students have the opportunity to ask for clarification of text. Apart from removing any uncertainty or anxiety, this could further socialization and verbalization.

2. It gives the examiner a better chance to design tasks which develop skills like mathematization, interpretation, reflection, creativity and critical thinking - all possible elements of the problem solving model (cf. 1.5). Tasks like these are time consuming, but provide activities that are essential if one wishes to realise worthwhile goals in mathematics education.

3. Tasks can be directed to include group work which could further socialization and verbalization even more.

4. Research skills could be developed by setting up tasks that require students to look for help outside their textbooks, for example in the school library.

5. More realistic problems could be attempted.

6. More extensive text could not only communicate the information more clearly, but also improve reading ability.

In Scotland the problem of time-restricted senior secondary examinations is addressed in an interesting way. Students are given a number of challenging tasks to solve in their own time. Tasks are designed according to the aims of the syllabus and because time is not a problem, students become involved in activities that develop important mathematical skills. After a few weeks students write a time-restricted examination and one of the challenging tasks is part of the examination paper. Of course students do not know which one, so they have to prepare all of them. In this way, students get the opportunity of reporting their solution. Tasks are set up in such a way that it becomes virtually impossible to report back without understanding (Hutchon, 1991).

The Hewet Project (cf. Kindt, 1987) also identified time-restricted tests as a barrier to reaching the goals of the new *Mathematics A* curriculum in the Netherlands. Research on this matter was done by one of the members of the Hewet team and culminated in his doctoral thesis (De Lange, 1987). Four kinds of alternative tasks were developed and are discussed in his book. Likewise, the Cockroft report (1982), giving directions for assessment at the 16+ level, advocates the necessity to assess as many skills as possible, including those that need other means of assessment than time-limited tests. Thus, research makes it quite clear: If time is not a limiting factor, more than just readability can be addressed more effectively.

On the other hand, time-limited examinations have a role to play and are common practice in many countries, also in South Africa. The aim of mentioning the above issues is to illustrate that time-limited examinations cause definite barriers and should not be *the only* means of assessment. If, however, time *is* limited, verbal problems must be designed with great care and the text should be as accessible as possible.

Part of the solution could be found by allowing more examination time for higher grade students. At present students taking mathematics on the higher grade have a double hurdle. Not only is the examination paper of a much higher standard (as can be expected), but the allocated time is noticeably less than that of the standard grade paper.

7.1.8 *Highly accessible text could be even more necessary if students are expected to apply their mathematical knowledge in a* ***purely mathematical*** *situation. Writers of such examination questions could therefore be assisted by similar guidelines as suggested in the checklist above.*

A familiar context plays an important role when students solve a verbal problem posed in a real-life situation. Not only does the context aid reading comprehension, but the structure of the real-life situation is often the source of students' solution methods. In short, all problem solving phases are supported by familiar, real-life contexts. To a great extent this support falls away in a purely mathematical situation. To most students a mathematical context is not a familiar one. Readability problems due to factors like unstructured text or a high proposition density may cause even more comprehension problems than in a real-life setting. Consider the following two questions taken from std 10 HG examination papers. The first question is from the multiple-choice section.

***Example 1*** *What is the gradient of the tangent to the curve of* $y = (2 - x)(x + 3)x$ *where* $x = -1$

(National Examinations, March 1994)

***Example 2*** *Determine the coordinates of the point where the tangent to the curve of the function given by* $g(x) = (x - 3)^2$ *is perpendicular to the line given by* $y = 4 - 3x$

(CED, 1993).

Both questions are examples of questions where students are expected to apply their mathematical knowledge in a purely mathematical situation. There is no support of a real-life context. Readability factors that could influence comprehension are the following:

- A time factor. The allocated marks suggest that the first question be completed in less than three minutes and the second in less than five.

- There are no introductory questions or sketches to elucidate the text or support comprehension.

- Both questions have a high proposition density per sentence and paragraph.

- The information input is not in the same order as is necessary for constructing a CKS and the principle of the given-new contract is not followed (cf. 2.4.3.4). In brief: the information input is unstructured. For time-limited examinations this is an important issue.

- In both problems, the sentence starts with a question. Text and metatext are given in one sentence with metatext before text. Students therefore have to keep this question (metatext) in their short term memory (STM) while they try to assemble the propositions or pieces of given information into a meaningful whole. This usually does not work because the capacity of STM is too small to retain the question. Before one can start giving meaning to the question, one first has to synthesize the context and this uses its own share of cognitive space (Frederiksen, 1990). The position of the question therefore interferes with the processing of information.

- The form of the first function is apt to cause problems. The x at the end of the expression is important, but the uncommon and unimportant position given to this factor will most probably cause some students not to "see" it. One of the possible answers given to this question in the multiple-choice section was based precisely on this blunder.

A more readable version of the first question could be as follows:

*A function, g, is defined by*

$g(x) = x(2 - x)(x + 3)$

*1. Draw a sketch graph of g. Clearly show the intercepts with the x-axis.*

*2. A tangent is drawn to the curve at the point where* $x = -1$.
*Find the gradient of the tangent.*

A few adaptations have been made:

- The metatext and text have been separated.
- The information is more structured. The order of information input aids the construction of a knowledge structure.
- The form of the function is more familiar.
- There is a relief in proposition density.

Although the influence of readability problems in mathematical situations was not researched in this study, the above examples suggest that similar problems will be experienced as was the case in real-life contexts. Suggestions for further research in this regard are given in 7.3.3.

## 7.2 Critical overview

The present research has generated some insight into the type of readability problems students experience when solving verbal problems in mathematics as well as the influence these problems have on achievement. Much more was uncovered than was originally anticipated. From this point of view the study could make a contribution to the already existing body of knowledge on readability research. However, the study could have revealed even more, had the following issues been considered:

- The number of students participating in the composite test could have been larger. This was especially true when the hypothesis was tested for each group separately. For example, the results of Question 6 for the E2 group and Question 9 for the E3 group indicate that improved readability improved test scores by as much as 22% and yet the differences were not measured to be significant (cf. Table 5.6). This was possibly caused by a high standard deviation amongst the relatively small number of available scores, only twelve per version per language group.

- Another factor relates to the adapted versions. If the adapted questions had been pre-tested and the results discussed with the students concerned *before* the actual composite test, the readability could probably have been improved even more and problems like the ones experienced in Question 4 could have been intercepted.

- As the study progressed, it became clear that first and second language groups often had different reading needs. This became more evident after the protocols of the E3 students had been analyzed. For example, more explicit and extensive information for the E3 group was not necessarily the prime need of E1 readers. This could mean that improved readability for one group could possibly even have been experienced as awkward or unnecessary by the other. First and second language readers should therefore have been subjected to different versions of improved readability. It seems that by using the same adaptations for all three groups, not one group was helped optimally. In this respect the E3 group was affected most because the adaptations were mainly inspired by the suggestions of E1 and E2 readers. Unfortunately, the delay of data from the E3 protocols was beyond control.

- To speak loudly into a tape recorder when solving verbal problems in mathematics was a new experience to all students. One can assume, with quite a high degree of certainty, that not all thoughts were verbalized. Besides, to do this in a second language must have been even more strange, especially since the subject was mathematics. Normally students do not find it natural to talk about mathematics. In a second language this is even more difficult. In spite of this, the E2 group managed quite well, although their reports were not as fluent and extensive as those of first language readers. The E3 group found the verbalization more difficult. They did not complain, but the analysis of protocols revealed their problems. With hindsight one realizes this is obvious. The E3 students who did the protocol study do not speak as much English as the E2 group. Their medium of communication at home and school is mainly their mother tongue, a language unrelated to English. Had the students been told to do the verbal reports in their mother tongue, much more data would probably have been available, especially from the E3 group. One would have had more insight into their comprehension problems and this could have guided the adaptations more effectively. An authority in the field of the relevant African language could have translated the reports and prepared them for the analysis. Should the protocol study be repeated, both second language groups should be allowed to verbalize their thoughts in their mother tongue.

- A few of the nine verbal problems used in this study were known to some students. Therefore, many possible readability problems, or other mathematical difficulties, had already been overcome during a previous attempt. Improving the readability did not help these students much. Had more care been taken to prevent this situation, the improvement in test scores would probably have been higher. One of the criteria when choosing the nine verbal problems was that the verbal problems had to be authentic - they had to represent *real* examination questions (cf. 4.3.2). One realized that students could possibly have been given these problems before and therefore each question paper had an additional question, asking students whether they had already done any of the questions. The data of those students who answered affirmatively was not used. There is, however, good reason to believe that not all students were honest or had perhaps forgotten that they had done the problem on a previous occasion.

Some of the above shortcomings have been addressed in the suggestions for further research.

## 7.3 Suggestions for further research

7.3.1 *Protocol analysis played a very important role in this study and is highly recommended for research in mathematics education.*

Although students did not verbalize all their thoughts, a large amount of verbal data nevertheless became available which would otherwise have remained unknown. Apart from being used in formal research, protocol analysis could also be utilized in classroom practice to guide teaching. It seems to be an excellent way of revealing misconceptions.

7.3.2 *This study could be repeated, but the research method should be improved by removing the points of criticism mentioned in 7.2*

This means the following round of research would:

- allow students to verbalize in their mother tongue during the talk-aloud experiment
- draw up separate test papers for each language group when testing the hypothesis
- use an African (E3) mathematics teacher to help with the analysis of the protocols and the adaptations of the E3 group.
- use more students for the composite test
- pre-test the adapted versions and discuss them with the relevant language group
- use only examination questions students have not seen before.

7.3.3 *This study could be repeated for verbal problems in a mathematical context.*

It would be important to know what readability factors are encountered in a mathematics context and what the effect of improved readability would be on students' performance in this case. There is a surmise that improved readability could improve test scores even more because in a mathematical context the support of real-life experiences falls away. In this regard the work of Nederpelt (1987) as well as Schouwstra (1988) need to be mentioned. Nederpelt's research is an in-depth study on the language of mathematics while Schouwstra investigated the readability of mathematical formulae. Both resources could be valuable for possible research on the influence of readability factors on problem solving if mathematical knowledge is applied in a purely mathematical situation.

7.3.4 *A study on the influence of readability on the performance of E3 examinees could be done by an E3 mathematics researcher.*

Language and culture are highly sophisticated features of human beings. Only a person with the same mother tongue and culture as the E3 examinee can really identify and address the readability problems this group encounters in a second language. It would be important to determine the difference in performance between the original and adapted versions if the E3 group were to write a test paper adapted by an E3 writer.

7.3.5 *The effect of improved readability on achievement should be researched in* ***authentic*** *conditions.*

While marking the composite tests one was always aware of the fact that the effect of improved readability could not be measured effectively while the conditions were not authentic. For example, two factors that play a crucial role during examinations are time and anxiety. Both these factors were virtually missing when students wrote the tests.

It is suggested that the influence of readability on achievement be researched by using the authentic setting of a final std 10 HG mathematics examination. This would be an ideal condition and generate invaluable data. An authority in the field of readability could be asked to improve the readability of the examination paper after the content has been approved by the moderator. There would then be two examination papers, an original and an adapted one, differing only in readability. A few schools could be identified to participate in the experiment. These schools would write the "adapted" paper. The advantages of research under these conditions are the following:

1. The conditions are authentic so the influence of factors like time and examination tension are present.

2. A large number of students are involved, which is beneficial for the statistical analysis.

3. By manipulating the correct variables one could incorporate a number of research projects within one experiment. For example, one could:

- Determine the influence of improved readability on the overall performance of students *on the whole paper.*
- Determine the influence of a question with readability problems on the performance of the question directly after that one.
- Determine the influence of improved readability if students were expected to apply their mathematical knowledge in a purely mathematical context.
- Determine the influence on performance of E3 students if the original examination paper (not the adapted one) were given to these readers in English as well as their mother tongue.
- Determine the influence on performance of E3 students if they were given an English examination paper with improved readability executed by an E3 writer.
- Determine the influence on performance of E3 students if they were given an English examination paper with improved readability done by an E1 writer.

To conclude, one could say that research on the readability of mathematics text can be considered as part of developmental research and as such

*"covers longitudinal developments*
*with a view on long-term learning processes*
*and is in itself*
*a long-term learning process"* (Freudenthal, 1991:162).

# BIBLIOGRAPHY

Adetula, L. (1990). Language factor: Does it affect children's performance on word problems? *Educational Studies in Mathematics*, (21).

Ahlfors, L.V. et al. (1962). On the mathematics curriculum of the high school, *Mathematics Teacher*, (55)3.

Akmajian, A., Dewers, R.A. & Harnish, R.M. (1984). *Linguistics - An introduction to language and communication*. London: The MIT Press.

*American heritage dictionary of the English language*, (1981). Boston: Houghton.

Arter, J.A. & Clinton, L. (1974). Time and error consequences of irrelevant data and question placement in arithmetic word problems. *Journal of Educational Research*, (68).

Artzt, A. & Amour-Thomas, E. (1992). Development of a cognitive-metacognitive framework for protocol-analysis of mathematical problem solving in small groups. *Cognition and Instruction*, (9)2.

Au, Terry Kit-Fong (1983). Chinese and English counterfactuals: The Sapir-Whorf Hypothesis revisited. *Cognition*, (15).

Austin, J. & Howson, G. (1979). Language and mathematical educaton. *Educational Studies in Mathematics*, (10).

Ausubel, D. (1960). The use of advanced organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology*, (51).

Ausubel, D. (1963). *The psychology of meaningful verbal learning*. New York: Grune and Stratton.

Baarda, D.B. & Goede, M.P.M. de (1991). *Basis boek SPSS/PC+: Eenvoudige introductie tot gebruik en beschrijving van commando's*. Leiden: Stenfert Kroes.

Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bartlett, F.C. (1964). *Remembering*. Cambridge: Cambridge University Press.

Bechtel, W. & Abrahamsen, A. (1991). *Connectionism and the mind: an introduction to parallel processing in networks*. London: Blackwell.

*Beeld*, (1989) 18 January. Probleme voorkeursaak vir Departement.

Berry, J.W. (1985). Learning mathematics in a second language - some cross-cultural issues. *For the Learning of Mathematics*, (5)2.

Bishop, A. (1979). Visualising and mathematics in a pre-technological culture. *Educational Studies in Mathematics*, (10).

Bishop, A. (1988). *Mathematics enculturation: a cultural perspective on mathematics*. Dordrecht: Kluwer Academic Publishers.

Bishop, A. (1993). Cultural conflicts in mathematics learning. Developing a research agenda for linking cognitive and affective issues. *Proceedings of 17th International conference for Psychology of Mathematics Education*, Hirabayashi, I. et al. (Eds.), Japan, (3).

Block, E. (1986). The comprehension strategies of second language readers. *TESOL Quarterly*, (20)3.

Block, E.L. (1992). See how they read. $L_1$ and $L_2$ reading comprehension monitoring. *TESOL Quarterly*, (26)3.

Bobrow, R.J. & Brown, J.S. (1975). Systematic understanding. *Representation and understanding: Studies in cognitive science*, D.G. Bobrow & A. Collins (Eds.). New York: Academic Press.

Bormuth, J.R. (1967). New developments in readability research. *Elementary English*, (44).

Bransford, J. & Johnson, M. (1972). Contextual prerequisites for understanding: Some investigations of comprehension an recall. *Journal of Verbal Learning and Verbal Behaviour*, (11).

Brodie, K. (1989). Education in a second language: Its effects on mathematics learning. *Pythagoras*, (19).

Bruner, J.S. (1964). The course of cognitive growth, *American Psychologist*, (19).

Carrell, P.L. (1981). Culture specific schemata in $L_2$ comprehension. *Selected papers from the 9th Illinois TESOL/BE Annual Convention*, Orem, R. & Haskell, J., (Eds.), Chicago: Illinois TESOL/BE.

Clarke, M.A. (1980). The short circuit hypothesis of ESL reading - or when language competence interferes with reading performance. *Modern Language Journal*, (64).

Clarkson, P.C. (1992). Unknown/careless errors in a mathematical language context. *Focus on Learning Problems in Mathematics*, (14)4.

Clyne, M. (1981). Culture and discourse structure. *Journal of Pragmatics*, (5).

Cockroft, Sir W. (1982). *Mathematics counts*. London: H.M.S.O.

Cohen, S.A. & Stover, G. (1980). Effects of teaching sixth-grade students to modify format variables of math word problems. *Reading Research Quarterly*, (16)2.

Coleman, E.B. (1962). Improving comprehensibility by shortening sentences. *Journal of Applied Psychology*, (46).

Cummins, D. et al. (1988). The role of understanding in solving word problems. *Cognitive Psychology*, (20).

Cummins, J. (1980). The cross-lingual dimensions of language proficiency: implications for bilingual education and the optimal age issue. *TESOL Quarterley*, (14).

Cummins, J. (1984). *Bilingualism and special education: Issues in assessment and pedagogy.* California: College Hill Press.

Cummins, J. & Swain, M. (1986). *Bilingualism in education*. London: Longman.

Dale, J. & Cuevas, G. (1987). *Integrating language and mathematics learning in ESL through content-area instruction*. Crandall, J. (Ed). Englewood Cliffs: Prentice Hall.

Dawe, L. (1983). Bilingualism and mathematical reasoning in English as a second language. *Educational Studies in Mathematics*, (14).

De Corte, E., Verschaffel, L. & De Win, L. (1985). Influence of rewording verbal problems on children's problem representations and solution. *Journal of Educational Psychology*, (77).

De Jager, C. et al. (1985). *Just mathematics*. Cape Town: Maskew Miller.

De Lange, J. (1987). *Mathematics, insight and meaning*. Utrecht: Rijksuniversiteit.

DET report (1993). Unpublished report. South Africa: Department of Education and Training.

Dijkhuis, J.H. et al. (1988). *Getal en ruimte voor die vierde klas VWO*. Culemborg: Educaboek.

Ericsson K.A. & Simon, H.A. (1980). Verbal reports as data. *Psychological Review*, (87)3.

Examenprogramma Wiskunde B Havo, (1989). Unpublished examination programme of Mathematics B syllabus in the Netherlands.

Faigley, L. & Witte, S. (1983). Topical focus in technical writing. *New essays in technical and scientific communication: Research, theory, practice*. Anderson P.V. et al. (Eds.). New York: Baywood Publishing Co. Inc.

Fehr, H.F. (1961). *New thinkings in school mathematics*. Paris: O.E.E.C.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, (32)3.

Flower, L. et al. (1983). Revising functional documents: The scenario principle. *New essays in technical and scientific communication: Research, theory, practice*. Anderson P.V. et al. (Eds.). New York: Baywood Publishing Co. Inc.

Frederiksen, N. (1990). Measuring skills in problem solving. *Cognitive assessment of language and math outcomes*. Legg, S. & Algina, J. (Eds.). New Jersey: Ablex Publishing Corporation.

Frendenthal, H. (1983). Major problems of mathematics education. *Proceedings of ICME 4*. Zweng, M. et al. (Eds.). Boston: Birkhäuser.

Freudenthal, H. (1991). *Revisiting mathematics education*. Dordrecht: Kluwer Academic Publishers.

Garofalo, J. & Lester, F. (1985). Metacognition, cognitive monitoring and mathematical performance. *Journal of Research in Mathematics Education*, (16).

Grice, H.P. (1967). Logic and conversation. William James lectures, Harvard University, 1967. Lecture 2 published in P. Cole & J.L. Morgan (Eds.), *Studies in Syntax*, Vol. III, New York: Seminar Press, 1975.

Gronlund, N.E.: *Constructing achievement tests*. Englewood Cliffs: Prentice Hall.

Gullat, D.E. (1986). Help your students read mathematics. *Arithmetic Teacher*.

Hall, E.T. & Hall, M.R. (1989). *Understanding cultural differences: Germans, French and Americans*. Yarmouth, ME: Intercultural Press, Inc.

Hamers, J. & Blanc, M. (1989). *Bilinguality and bilingualism*. Cambridge: Cambridge University Press.

Harris, M. & Van Deventer E. (1990). Overcoming the confusion of reading mathematics. *Focus on Learning Problems in Mathematics*, (12).

Hart, K.M. (1981). *Children's understanding of mathematics, 11-16*. London: John Murray.

Haviland, S.E. & Clark, H.H. (1974). What's new? Acquiring new information as a process of comprehension. *Journal of Verbal Learning & Verbal Behaviour*, (13).

Haviland, S.E. & Clark, H.H. (1977). Comprehension and the given-new contract. *Discourse processes: Advances in research and theory*. Freedle, R.O. (Ed.). New Jersey: Ablex Publishing Corporation.

Hofstede, G. (1991). *Cultures and organizations. Software of the mind*. England: McGraw-Hill.

Hofstede, G. (1992). *Allemaal andersdenkenden: Omgaan met cultuurverschillen*. London: McGraw-Hill.

Holyoak, K.J. (1974). The role of imagery in the evaluation of sentences: Imagery or semantic factors? *Journal of Verbal Learning and Verbal Behaviour*, (13).

Horn, J. & Dollinger, S. (1989). Effects of test anxiety, tests, and sleep on children's performance. *Journal of School Psychology*, (27).

Huckin, T. N. (1983). A cognitive approach to readability. *New essays in technical and scientific communication*. P.V. Anderson, C.R. Miller & J. Brockman (Eds.). New York: Baywood Press.

Huckin, T. N. (1987). *Surprise value in scientific discourse*. Unpublished conference paper. Atlanta.

Huckin, T.N. & Olsen, L.A. (1991). *Technical writing and professional communication*. New York: McGraw-Hill.

Hutchon, K. (1991). Personal interview. Examination Board, Dalkeith, Scottland.

Inkeles, A. & Levison, D. (1969). National character: the study of modal personality and sociocultural systems. *The handbook of social psychology*, 2nd Edition, vol. 4, G. Linzey & E. Aronson (Eds.), Reading MA: Addison-Wesley.

Jerman, M.E. & Rees, R. (1972). Predicting the relative difficulty of verbal arithmetic problems. *Educational Studies in Mathematics*, (4).

Jerman, M.E. (1973). Problem length as a structural variable in verbal arithmetic problems. *Educational Studies in Mathematics*, (5).

Jerman, M.E. & Mirman, S. (1974). Linguistic and computational variables in problem solving in elementary mathematics. *Educational Studies in Mathematics*, (5).

Johnson, H.C. (1944). The effect of instruction in mathematical vocabulary upon problem solving in arithmetic. *Journal of Educational Research*, (38)2.

Johnson, P. (1981). Effects of reading comprehension on language complexity and cultural background of a text. *TESOL* Quarterly, (15)2.

Joubert, G. (1985). *Die rol van wiskundige leesbekwaamheid ten opsigte van akademiese prestasie van eerstejaar ingenieur studente*. Unpublished D.Ed thesis. South Africa: University of Stellenbosch.

Just, M.A. & Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4).

Kane, R.B. (1970). The readability of mathematics text books revisited. *The Mathematics Teacher*, (November).

Kant, E. & Newell, A. (1984). Problem solving techniques for the designs of algorithms. *Information Processing and Management*, (20).

Kaplan, R.B. (1980). Cultural thought patterns in inter-cultural education. *Readings on English as a Second Language*. Croft, K. (Ed.). 2nd Edition. Massachussetts: Winthrop Publishers.

Kaplan, R.B. & Connor, U. (1987). *Writing across languages: analysis of $L_2$ text.* California: Addison-Wesley Publishing Company, Inc.

Kempen, G. & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, (11).

Kieras, D. (1978). Good and bad structure in simple paragraphs: Effects on apparent theme, reading time and recall. *Journal of Verbal Learning and Verbal Behaviour*, (17).

Kieras, D. (1980). Initial mention as a signal to thematic content in technical passages. *Memory & Cognition*, (8)4.

Kindt, M. (1987). *Hewet Project - Eindverslag.* Utrecht: Rijksuniversiteit.

Kintsch, W. & Keenan, J. (1973). Reading rate and retention as a function of the number of propositions in the base structure of sentences. *Cognitive Psychology*, (5).

Kintsch, W. et al. (1975). Comprehension and recall of text as a function of content variables. *Journal of Verbal Learning and Verbal Behaviour*, (14).

Kintsch, W. & Green E. (1978). The role of cultural specific schemata in the comprehension and recall of stories. *Discourse processes: Advances in research and theory.* Freedle, R.O. (Ed.). New Jersey: Ablex Publishing Corporation.

Kintsch, W. & Van Dijk, (1978). Toward a model of text comprehension and production. *Psychological Review*, (85)5.

Kirschner, M., Wexler, C. & Spector-Cohen, E. (1992). Avoiding obstacles to student comprehension of test questions. *TESOL Quarterly*, (26)3.

Klare, G.R. (1963). *The measurement of readability.* Iowa: Iowa State University Press.

Klare, G.R. (1976). A second look at the validity of readability formulas. *Journal of Reading Behaviour*, (8).

Klare, G.R. (1977). Readable technical writing: Some observations. *Technical Communication*, 24(2).

Kouba, V.L. (1989). Common and uncommon ground in mathematics and science terminology. *School Science and Mathematics*, (89).

Kuijk, L. et al. (1993). *Realistische wiskunde voor de basis vorming*. The Netherlands: Malmberg.

Lagerwerf, B. (1992). Taalproblemen. *Nieuwe Wiskrant*, July 1992.

Lambert, W. & Rawlings, C. (1969). Bilingual processing of mixed-language associative networks. *Journal of Verbal Learning and Verbal Behaviour*, (8).

Laridon, P. et al. (1986). *Classroom mathematics std. 10*. Johannesburg: McGraw-Hill.

Lean, G. et al. (1990). Linguistic and pedagogical factors affecting children's understanding of arithmetic word problems: a comparative study. *Educational Studies in Mathematics*, (21).

Lepik, M. (1990). Algebraic word problems: role of linguistic and structural variables. *Educational Studies in Mathematics*, (21).

Macdonald, C.A. (1990). *Crossing the threshold into standard three in black education*. Consolidated main report of the Threshhold Project. Pretoria: HSRC Publishers.

Mackenzie, W.S. (1990). Meaning: The common element in both reading and mathematics. *Ontario Mathematics Gazette*, (28).

Mägiste, E. (1985). Development of intra-and interlingual interference in bilinguals, *Journal of Psycholinguistic Research*, (14)2.

Markel, M.H. (1992). *Technical writing. Situations and strategies*. New York: St Martin's Press.

Miller, G. (1956). The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychological Review*, (63)2.

Mulder, F. (1991). Een uur heen, twee uur terug, *Nieuwe Wiskrant*, (11).

Mulder, F. (1992). *Wiskunde een wereldvak*. Utrecht: Freudenthal Institute.

Nas, G. (1983). Visual recognition in bilinguals. *Journal of Verbal Learning and Verbal Behavior*, (22).

*National Education Policy Investigation (NEPI): Language*, (1992). Cape Town: Oxford University Press.

Nederpelt, R.P. (1987). *De taal van de wiskunde*. Almere: Versluys Uitgewerij.

Nesher, P. & Teubal, E. (1975). Verbal cues as an interfering factor in verbal problem solving. *Educational Studies in Mathematics*, (6).

Newell, A. & Simon, H.A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart and Winston, Inc.

Pimm, D. (1987). *Speaking mathematically: Communication in mathematics classrooms*. London: Routledge.

Polya, G. (1946). *How to solve it*. New Jersey: Princeton University Press.

Rakow, S. & Gee, T. (1988). What reading techniques do mathematics teachers find valuable? *School Science and Mathematics*, (88)3.

Samuels, M.S. (1988). Technical reading processes and collaborative learning: New approaches to structuring readable texts. *Technical Communication,* First Quarter.

Sapir, E. (1949). *Culture, language and personality*. Berkeley: Univeristy of California Press.

Schoenfeld, A.H. (1985). Making sense of "out loud" problem-solving protocols. *The Journal of Mathematical behaviour*, (4).

Schouwstra, P.C. (1988). *Leesbaarheid van wiskundige formules, een kwalitatief model*. Unpublished M.Sc. thesis. The Netherlands: Technical University, Eindhoven.

Selzer, J. (1983). What constitutes a readable technical style? *Essays in technical and scientific communication: Research, theory, practice*. Andersen, P.V. et al. (Eds.). New York: Baywood Publishing Co. Inc.

Shohamy, E. (1984). Does the testing method make a difference? The case of reading comprehension. *Language Testing*, (1).

Skemp, R. (1976). Relational understanding and instrumental understanding. *Mathematics Teaching*, (77).

Skiba, A. (1990). Reviewing an old subject: math anxiety. *Mathematics Teacher* (March).

Slotboom, A. (1987). *Statistiek in woorden*. Groningen: Wolters-Nordhoff.

Smith, F. (1982). *Understanding reading. A psycholinguistic analysis of reading and learning to read*. New York: CBS College Publishing.

Southey, P.(1992). Junior primary language policy. *Language Projects Review*, (5)3.

Statman, S. (1988). Ask a clear question and get a clear answer. *System*, (16)3.

Steffensen, M.S., Joag-Dev, C. & Anderson, R. (1979). A cross-cultural perspective on reading comprehension. *Reading Research Quarterly*, (15).

Steffensen, M.S. (1987). The effect of context and culture on children's $L_2$ reading: a review. In Devine, J., Carrell, P. & Eskey, D. (Eds.), *Research in reading in English as a second language*. Washington : TESOL.

Stenhouse, L. (1967). *Culture and education*. London: Thomas Nelson (Printers) Ltd.

Stevenson, D.K. (1981). 'All of the above' - on problems in the testing of FL reading. *System*, (9).

Strauss, J.P. et al. (1993). *Education and manpower development*, (13). Bloemfontein: Universiteitsdrukkery.

Strother, J.B. (1990). *Syntax in the ST Register. Effect on writers' choices and readers' comprehension*. Published doctoral thesis. The Netherlands: Technical University, Eindhoven.

Swanepoel, P.H. (1989). *Special language issues in black education*. Pretoria: Unisa.

---- Syllabus for Mathematics Higher Grade (1991). Senior Secondary Course. South Africa: Cape Education Department.

----Syllabus for Mathematics Higher Grade (1985). Senior Secondary Course. South Africa: Cape Education Department.

Tagatz, G. (1976). *Child development and individually guided education*. London: Addison-Wesley Publishing Company.

Terlouw, P. (1991). *Wiskunde in intercultureel perspektief*. Leeuwaarden: Educatief.

Thorndyke, P. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, (9).

Treacy, J.P. (1944). The relationship of reading skills to the ability to solve arithmetic problems. *Journal of Educational Research*, (38).

Tudor, I. (1988). Pre-reading and the use of advance organizers in $L_2$ reading instruction. *ITL Review of Applied Linguistics*, (81-82).

Ulijn, J.M. & Strother, J.B. (1990). The effect of syntactic simplification on reading EST texts as $L_1$ and $L_2$. *Journal of Research in Reading*, (13)1.

Ulijn, J.M. & Strother, J.B. (1995). *Communicating in business and technology: from psycholinguistic theory to international practice*. Frankfurt: Verlag Peter Lang.

Ursunier, J. (1993). *International marketing: A cultural approach*. United Kingdom: Prentice Hall International Ltd.

Van den Heuwel-Panhuizen, M. (1993). New forms of assessment but don't forget the problems. *Proceedings of the 17th International Conference for the Psychology of Mathematics Education*. Hirabayashi, I. et al. (Eds.), Japan.

Vygotsky, L.S. (1962). *Thought and language*. New York: MIT Press.

Waern, Y. (1979). Thinking aloud during reading: A descriptive model. University of Stockholm, Department of Psychology, *Report 546*.

Walters, P. (1993). Access to std. 6 (model C): Some linguistic barriers for pupils from DET background. *CRUX*, (May).

Washburn, S.L. (1974). Tools and human evolution. *Scientific technology and social change*. Rochlin, G.I., (Ed.). San Francisco: Freeman.

Weerman, F. (1994). Over talen - en wiskunde knobbels. *Euclides*, March 1994.

Whorf, B.L. (1956). *Language, thought and reality*. London: Chapham & Hall Ltd.

Wierzbicka, A. (1991). *Cross-cultural pragmatics*. New York: Mouton de Gruyter.

Zimmler, K. & Keenan, J. (1983). Imagery in the congenitally blind: How visual are visual images? *Journal of Experimental Psychology: Learning, Memory and Cognition*, (9)2.

# APPENDIX A

## 9 QUESTIONS FOR THE TALK-ALOUD EXPERIMENT

1. **STARTING TIME:** ....................................

Two pumps, X and Y, are used to supply a minimum of 60 000 litres of water per day to a reservoir. X pumps 5 000 litres per hour and the running costs are 90 cents per hour. Y pumps 2 000 litres per hour and the running costs are 60 cents per hour. The sum of the number of hours that the pumps operate per day must not exceed 20. The number of hours that X operates per day may not be more than twice that of Y.

1.1 Represent the above information as a system of inequalities. (5)

1.2 Use graph paper to represent these inequalities graphically and clearly indicate the feasible region. (5)

1.3 Use your graph and determine the number of hours that each pump will have to operate per day in order to minimize the total running costs. (6)

**STARTING TIME:** ....................................

**FINISHING TIME:** ....................................

---

What did you find difficult to understand?

.................................................................................................................

.................................................................................................................

.................................................................................................................

Remember

to

TALK

ALOUD

when

thinking

(Tvl. 1990)

2. **STARTING TIME:** ..................................

A particle moves along a straight line so that, t seconds after observations have commenced, its distance, s metres, from a fixed point 0, is given by

$$s = \frac{1}{2}t^3 - \frac{11}{2}t^2 + 19t - 20 \quad (t \geq 0)$$

CALCULATE:

2.1 how far the particle will be from the fixed point 0 at the moment when observations commence (2)

2.2 the values of t for which the particle is momentarily at rest(5)

2.3 the speed of the particle at $t = 2$ (2)

2.4 when, and how many times, the particle will pass the fixed point 0 (9)

2.5 the time intervals when the speed of the particle decreases while it is moving away from 0 (4)

2.6 the distance covered by the particle during the fourth second. (2)

**STARTING TIME:** ..................................

**FINISHING TIME:** ..................................

---

What did you find difficult to understand?

..................................................................................................................

..................................................................................................................

..................................................................................................................

Remember to TALK ALOUD when thinking

(CED, 1990)

3. **STARTING TIME:** ..................................

A manufacturer has the capacity to produce 8 000 calculators of a certain type per week. The cost of producing n calculators is $C = 100n + 200$, and n calculators can be sold per week at a price $P = 400 - 0{,}02n$ per calculator.

Determine the value of n which maximizes the profit. (6)

**STARTING TIME:** ..................................

Remember

to

TALK

ALOUD

when

thinking

**FINISHING TIME:** ..................................

What did you find difficult to understand?

..................................................................................................................

..................................................................................................................

..................................................................................................................

(House of Delegates, 1993)

4. **STARTING TIME:** ....................................

The number of pupils in a class is x and there are exactly enough exercise books to supply each pupil with y books. If there had been 5 more pupils each would have received 2 books less and there would have been 6 books left over. If however there had been 3 pupils less each would have received 1 book more and there would have been 11 books left over.

How many pupils are there in the class and how many books are available? (12)

**STARTING TIME:** ....................................

**FINISHING TIME:** ....................................

---

What did you find difficult to understand?

......................................................................................................................

......................................................................................................................

......................................................................................................................

Remember

to

TALK

ALOUD

when

thinking

(House of Delegates, 1993)

5. **STARTING TIME:** ..................................

In a factory x units of a product A and y units of a product B are processed through two departments. $D_1$ and $D_2$. A requires 2 hours per unit in $D_1$ and 4 hours per unit in $D_2$. B requires 3 hours per unit in $D_1$ and 2 hours per unit in $D_2$. $D_1$ and $D_2$ have 60 hours and 80 hours respectively, available each week.

5.1 Use the above information to write down the set of constraints in terms of x and y. (4)

5.2 Represent these inequalities on the graph sheet provided and shade the feasible region. (6)

5.3 If the profit margins are R30 and R40 for A and B respectively write down the objective function. (2)

5.4 Draw the optimal search line and hence write down the recommended product - mix that will maximize profit. (3)

5.5 Determine the maximum profit. (2)

5.6 What percentage of daily capacity will be utilized in each department? (2)

**STARTING TIME:** ..................................

Remember to TALK ALOUD when thinking

**FINISHING TIME:** ..................................

---

What did you find difficult to understand?

..........................................................................................................................

..........................................................................................................................

..........................................................................................................................

(House of Delegates, 1990)

6. **STARTING TIME:** ....................................

Two cyclists A and B travel from the same place and leave simultaneously. The distance S, in kilometres, that both of them are from the starting point after t hours, is respectively given by the formulae.

$S_A = t^3 - 12t^2 + 36t$ and

$S_B = 24t - 4t^2$

Both are back at the starting point simultaneously.

6.1 Find the average speed of A over the first 2 hours. (5)

6.2 Find the speed of B 4 hours after he has left. Interpret your answer. (4)

6.3 After how many hours do their speeds differ by 7 km/h? (7)

6.4 After how many hours has each begun to return to the starting point? Interpret your answer. (5)

**STARTING TIME:** ....................................

**FINISHING TIME:** ....................................

---

What did you find difficult to understand?

..........................................................................................................................

..........................................................................................................................

..........................................................................................................................

Remember to TALK ALOUD when thinking

(House of Representatives, 1990)

7. **STARTING TIME:** ...................................

A manufacturer of hi-fi sets determines that in order to sell x units of a new hi-fi, its price per unit must be p = (1000 - x) Rand. He also determines that the total cost, c, of producing x units is given by c = (3000 + 20x) Rand.

7.1 Deduce an expression for the total profit in terms of x. (4)

7.2 How many units must be produced and sold in order to maximise the profit? (2)

**STARTING TIME:** ...................................

Remember

to

TALK

ALOUD

when

thinking

**FINISHING TIME:** ...................................

---

What did you find difficult to understand?

.......................................................................................................................

.......................................................................................................................

.......................................................................................................................

(CED, 1990)

8. **STARTING TIME:** ....................................

It is estimated that t years from now the circulation of a magazine will be given by the function

$$m(t) = 50t^2 - 200t + 3000$$

8.1 At the time of first publication, what was the circulation of the magazine? (2)

8.2 Derive an expression for the rate at which the circulation will be changing t years from now. (2)

8.3 At what rate will the circulation be changing 3 years from now? (2)

8.4 Determine by how much the circulation will change during the 2nd year. Will it be increasing or decreasing? (4)

**STARTING TIME:** ....................................

**FINISHING TIME:** ....................................

---

What did you find difficult to understand?

..................................................................................................................

..................................................................................................................

..................................................................................................................

Remember

to

TALK

ALOUD

when

thinking

(House of Delegates, 1993)

9. **STARTING TIME:** ..................................

Two straight roads intersect perpendicularly at 0. P is a point on one road such that OP = 10 km. Two persons at P and 0 respectively, start to walk simultaneously, the one at P in the direction of 0 along the one road at 3 km/h, and the other at 0 along the other road away from 0 at 4 km/h. After t hours they reach the positions A and B on the two roads.

9.1 Find an expression for $AB^2$. (4)

9.2 After how many hours are the persons nearest to one another? (2)

9.3 What is the shortest distance between them? (3)

**STARTING TIME:** ..................................

Remember to TALK ALOUD when thinking

**FINISHING TIME:** ..................................

---

What did you find difficult to understand?

..................................................................................................................

..................................................................................................................

..................................................................................................................

(House of Representatives, 1990)

# APPENDIX B

# INSTRUCTIONS FOR ADAPTATIONS

## INSTRUCTIONS FOR ADAPTATIONS

The text must be changed to two forms:

1. As **verbal** as possible (as little non-verbal as possible).
2. As **non-verbal** as possible (as little verbal as possible).

The aim of changing the text to the extreme verbal form is to help with a clearer understanding of the question. The formulation must be as pupil-friendly as possible.

**To change to the extreme verbal form**: Start by getting rid of as many non-verbal interferences as possible. You may change anything as long as:

- you have in mind the aim of the adaptation: to make the question clearer to understand (easier to visualize) so that the mathematical process can get started in the brains of the pupils
- you keep the meaning the same
- you don't change the mathematical demand.

Here is an **example** of an original question with the two adaptations:

## Original question:

The length of a steel ingot passing through a rolling mill is given by $L = -0{,}01t^3 + 0{,}15t^2 + 1$ (L in m, t in min.). When the rate of increase of L has become zero the bar has cooled down and the rolling stops.

Find:

1. the initial length of the ingot
2. the rate of increase of the length of the bar when $t = 2$ min.
3. the time at which the rate of increase of length is a maximum.
4. the length of time the bar is rolled for, and its final length.

**Extreme verbal version** [my effort - you could do even better!]

A bar of steel passes through a rolling mill. The length of the steel bar is given by the formula:

The length of the steel bar is given by the formula:

Length of steel bar $= -0{,}01t^3 + 0{,}15t^2 + 1$ metres.

The time (t), is measured in minutes.

The rolling stops as soon as the bar has cooled down. Then the rate at which the bar's length increases, is zero.

Find:

1. the length of the steel bar just before the rolling begins.
2. the rate at which the length increases after two minutes.
3. the time when the rate at which the length increases, is a maximum.
4. how many minutes the steel is rolled.
5. the final length of the bar.

## Non-verbal version

Consider the function:

$$L = -0{,}01t^3 + 0{,}15t^2 + 1$$

Find

1. $L(0)$.
2. $L'(2)$.
3. $t$ if $L'(t)$ is a maximum
4. $t$ if $L'(t) = 0$
5. $L(t)$ for the value of $t$ found in 4.

# APPENDIX C

## NON-VERBAL VERSIONS

1. **STARTING TIME:** ....................................

Consider the following inequalities of a linear programming problem.

$x + y \leq 20$
$5000x + 2000y \geq 60000$
$x \leq 2y$
$x \geq 0; y \geq 0$

1.1 Represent these inequalities on a graph paper. Shade the feasible region. (6)

1.2 The objective function is

$P = 60x + 90y$

1.2.1 Determine the value of x and y for a maximum value of P. (3)

1.2.2 Determine the maximum value of P. (2)

[11]

**STARTING TIME:** ....................................

**FINISHING TIME:** ....................................