The assessment of the quality of science education textbooks:
Conceptual framework and instruments for analysis

by

Sarita Swanepoel

submitted in accordance with the requirements
for the degree of

Doctor of Education

in the subject
Didactics

at the

UNIVERSITY OF SOUTH AFRICA

Supervisor: Prof. C.H. Swanepoel

April 2010

Summary

Science and technology are constantly transforming our day-to-day living. Science education has become of vital importance to prepare learners for this ever-changing world. Unfortunately, science education in South Africa is hampered by under-qualified and inexperienced teachers. Textbooks of good quality can assist teachers and learners and facilitate the development of science teachers. For this reason thorough assessment of textbooks is needed to inform the selection of good textbooks.

An investigation revealed that the available textbook evaluation instruments are not suitable for the evaluation of the physical science textbooks in the South African context. An instrument is needed that focusses on science education textbooks and which prescribes the criteria, weights, evaluation procedure and rating scheme that can ensure justifiable, transparent, reliable and valid evaluation results. This study utilised elements from the Analytic Hierarchy Process (AHP) to develop such an instrument and verified the reliability and validity of the instrument's evaluation results.

Development of the Instrument for the Evaluation of Science Education Textbooks started with the formulation of criteria. Characteristics that influence the quality of textbooks were identified from literature, existing evaluation instruments and stakeholders' concerns. In accordance with the AHP, these characteristics or criteria were divided into categories or branches to give a hierarchical structure. Subject experts verified the content validity of the hierarchy.

Expert science teachers compared the importance of different criteria. The data were used to derive weights for the different criteria with the Expert Choice computer application. A rubric was formulated to act as rating-scheme and score sheet. During the textbook evaluation process the ratings were transferred to a spreadsheet that computed the scores for the quality of a textbook as a whole as well as for the different categories.

The instrument was tested on small scale, adjusted and then applied on a larger scale. The results of different analysts were compared to verify the reliability of the instrument. Triangulation with the opinions of teachers who have used the textbooks confirmed the validity of the evaluation results obtained with the instrument. Future investigations on the evaluation instrument can include the use of different rating scales and limiting of criteria.

Key terms: science education, science textbooks, textbook evaluation, Analytic Hierarchy Process, evaluation instrument, criteria, weights, pairwise comparisons, rubric, textbook quality

I, Sarita Swanepoel, declare that “The assessment of the quality of science education textbooks: Conceptual framework and instruments for analysis” is my own work and that all the sources that I used or quoted have been indicated and acknowledged by means of complete references.

Sarita Swanepoel

Acknowledgements

I am heartily thankful to my supervisor, Prof C. H. Swanepoel, for his competent guidance and support, and to my colleagues who sacrificed their time to do the pair-wise comparisons and test the instrument.

I also want to express my gratitude to everybody who supported me in so many ways during the completion of this project.

Contents

List of Figures

List of Tables

Chapter 1

Orientation

1.1 Introduction

According to Section 29 of the South African Bill of Rights, every child has the right to education. The education provided should enable learners to realise their potential and prepare them for life, as expressed by the revised National Curriculum (Department of Education of South Africa, 2001a, 8):

> The curriculum aims to develop the full potential of each learner as a citizen of a democratic South Africa. It seeks to create a lifelong learner who is confident and independent, literate, numerate and multi-skilled, compassionate, with a respect for the environment and the ability to participate in society as a critical and active citizen.

Science education is an important component of education for all learners – not only for future scientists. Although only a small percentage of students are destined to follow scientific careers, every person needs some understanding of mathematics, science and technology to succeed in today's technologically oriented world (Lederman, 2008; Lisichkin, 2007; Jenkins, 2004; Singer and Tuomi, 2003; Department of Education of South Africa, 2003a; Leitte, 2002). Science education of high quality is, therefore, essential not only to prepare learners to enter scientific careers but to contribute to providing the country with a scientifically literate population that can "address the global challenges that humanity now faces" (Wieman, 2007, 9). In this regard science education aids the realisation of the learner's potential and contributes to the development of our country's human resources (Reddy, 2006; Department of Education of South Africa, 2001b).

Unfortunately, the 2003 Trends in International Mathematics and Science Study (TIMMS) shows that South Africa's performance in both science and mathematics is at the bottom of the list of 46 participating countries (Maree, Aldous, Hattingh, Swanepoel and van der Linde, 2006; Conzales, Guzman, Partelow, Pahlke, Jocelyn, Kastberg and Williams, 2004). One of the factors contributing to this unfortunate state of affairs is the fact that science education in South Africa is hampered by unqualified, under-qualified and inexperienced teachers (Taylor, 2008b; Bisseker, 2005; Muwanga-Zake, n.d.; De Beer, 2002; Department of Education of South Africa, 2001b). The plight of the teachers has been aggravated by the implementation of a policy of outcomes-based education (OBE) in 1997. This implied the introduction of a new, unfamiliar learner-centered activity-based approach. Teachers had to become facilitators of knowledge construction and designers of learning programmes and materials (Department of Education of South Africa, 2001a). They were expected to develop learning programmes by "designing back" from the outcomes stated by the National Curriculum (Killen, 2003). This placed a heavy burden on teachers. The OBE approach also required the recording of marks for all formative assessment tasks which resulted in increased work pressure on the part of OBE teachers (Aldridge, Laugksch and Fraser, 2006). Although this aspect of OBE will change significantly with the implementation of the new planning for 2010 and beyond (Department of Education of South Africa, 2009), teachers will still have to prepare and plan to guide learners to master the science content.

In order to address the problem of inadequately qualified teachers, teacher training and development has been identified by the National Education Department as a priority on the transformation agenda (Muwanga-Zake, n.d.). On 8 February 2004 the Sunday Times (Sukhraj, Mkhize and Govender, 2004) reported that the Education Labour Relations Council is spending R95-million in an effort to "upgrade" teachers. New norms and standards for teacher education were developed and accepted by the Committee on Teacher Education Policy (COTEP) (Department of Education of South Africa, 1997) and local universities accordingly designed appropriate courses for the training and retraining of teachers (Hattingh, 2009; Muwanga-Zake, n.d.). The shortage of science teachers, however, makes it impractical to simultaneously remove a substantial number of teachers from schools for long periods in order to retrain them. Training in the new curriculum was limited to in-service courses lasting three to five days (Robinson, 2002). Mentorship programmes have been launched to assist teachers in their own classrooms (Hattingh, 2009), and, where needed, targeted in-service training will be provided by the Department of Education from 2010 (Motshekga, 2009).

Textbooks can play an important role in addressing the problem of inadequately qualified teachers (Taylor, 2008a; Newton and Newton, 2006; Reddy, 2005; Huber and Moore, 2001). Research shows that textbooks are among the most cost-effective ways of improving classroom practice (Lubben, Campbell, Kasanda, Kapenda, Gaoseb and Kanjeo-Marenga, 2003; Asmal, 2002; Verspoor, 1991). Unfortunately the Department of Education did not initially embrace this as an aid in the transformation process. Initial teacher training actually indicated that teachers should create their own learning material. In their review of the implementation of Curriculum 2005 the Chisholm report indicated in 1999 that most teachers do not have the time, the resources or often the skills to be involved in the development of high quality learning programmes and materials (Department of Education of South Africa, 1999). The negative attitude towards textbooks persisted in varying degrees until 2009 when the Task Team for the Review of the Implementation of the National Curriculum Statement (2009) identified the use of textbooks as critical to educational success and the Department of Education undertook to emphasise the importance of textbooks from January 2010 (Motshekga, 2009; Department of Education of South Africa, 2009).

Textbooks can be used to address the problem of under-qualified and inexperienced teachers in a number of ways. The connection between learner achievement and the availability and utilization of high quality textbooks is well-established (Ogan-Bekiroglu, 2007; Singer and Tuomi, 2003; Lubben et al., 2003; Asmal, 2002; Mozambique Ministry of Education, 2002; Department of Education of South Africa, 2001b). Furthermore, research has shown that using well designed textbooks or curriculum material can positively influence teacher beliefs and practices, aiding curriculum implementation (Davis, 2009; Newton and Newton, 2006; Davis, 2003a; Izsak and Sherin, 2003; McKenney, 2001). The UNESCO Regional Bureau for Education in Africa (2003) even describes textbooks as the cornerstone for meeting the goals of education (UNESCO Regional Bureau for Education in Africa, 2003). The utilisation of high quality textbooks can contribute much in supporting learners in the learning process, as well as supporting teachers in their instruction and their professional development (Taylor, 2008a; Pandor, 2006; Singer and Tuomi, 2003; Ansary and Babaii, 2002; Goldsmith et al., 2000; Watts and Simon, 1999; Smith, Blakeslee and Anderson, 1993). Malcolm and Alant (2004, 72) formulate the role of textbooks in Science Education in South Africa as follows:

> In schools where teachers often have limited content knowledge and planning skills, and where students need to do considerable work by themselves, text-books serve as sources of science knowledge, curriculum planning and teaching ideas for teachers and students.

Although some teachers function well without the use of textbooks, studies worldwide show that textbooks are routinely used in classes (Lemmer, Edwards and Rapule, 2008; Ogan-Bekiroglu, 2007; Pepin and Haggerty, 2003; Kesidou and Roseman, 2002; Sitte, 1999) and that the majority of teachers use textbooks in their planning and presentation of instruction (Arriassecq and Greca, 2007; Klassen, 2006; Henson, 2004; Issitt, 2004; Pepin and Haggerty, 2003; Mozambique Ministry of Education, 2002). Some teachers even use a variety of textbooks to provide them with examples of high quality teaching strategies, activities and assessment tasks (Newton and Newton, 2006; Henson, 2004; Pepin and Haggerty, 2003; Izsak and Sherin, 2003). A qualified and experienced teacher will use, supplement and discard portions of the textbook according to his or her learners' needs (Tyson, 1997; Wong, 1991). A very good teacher can use almost any textbook to good advantage. The quality of textbooks is, however, more important when used by inexperienced or under-qualified teachers who follow these rigidly (Ogan-Bekiroglu, 2007; Van den Berg, 2004; Muwanga-Zake, n.d.; Kesidou and Roseman, 2002; Tyson, 1997; Richaudeau, 1980). In South Africa, with the number of unqualified science teachers, this is especially relevant (Malcolm and Alant, 2004). It has a major impact on beginner educators (Steyn, 2004). An inexperienced or unqualified teacher may use and overuse the textbook, because it provides a feeling of safety (Henson, 2004). The quality of the science textbooks used by these teachers is critical to their professional development and to the quality of the instruction their learners receive. It is, therefore, understandable that Malcolm and Alant (2004) identify research into the quality of texts as an important field for research in South Africa.

Any change in educational approach necessitates designing new textbooks (Tobin, Tippins and Gallard, 1994). The implementation of OBE in South Africa has provided the opportunity to provide teachers and learners with new science textbooks of good quality (Stoffels, 2005). Unfortunately submissions to the Chisholm Committee on the implementation of Curriculum 2005 described newly designed OBE textbooks as "superficial and essentially a re-issue of the old textbooks with the relevant outcomes annotated in the margins" (Department of Education of South Africa, 2001a). The Independent Projects Trust reports that large numbers of textbooks were still "outdated, irrelevant and inaccurate" (Independent Projects Trust, n.d.). The quality of textbooks has improved in recent years and although there are examples of bad textbooks, there are also many good textbooks (Taylor, 2008a).

In the light of the negative reports on the quality of some available textbooks, as well as the fact that the South African government annually spends over a

R1000 million on textbook (McCallum, 2004; Department of Education of South Africa, 2003b), it is important to be able to assess the quality of available science textbooks during the selection of textbooks for use in science classrooms. Only thorough assessment of the textbooks can enable teachers to select instructional materials that can support them as they translate outcomes into everyday teaching (Lemmer et al., 2008; Davis, 2003b). However, teachers are generally not offered any training in the evaluation of textbooks, either in their teacher training programmes or even as members of evaluation panels (Stein, Stuen, Carnine and Long, 2001; Johnsen, 1993).

All evaluation research projects aim to prescribe changes and improve the situation (Walliman, 2005). In this instance, good assessment and selection procedures and practices can also suggest changes to improve textbooks and encourage publishers to implement the changes to produce better textbooks (Mozambique Ministry of Education, 2002; Chambliss, 1994; Britton, Gülgöz and Glynn, 1993). According to Tyson (1997, 10)

> The most powerful and direct way to draw forth better textbooks is to create and sustain a well-funded, unhurried, and thoughtful system of textbook evaluation.

1.2 Problem analysis

1.2.1 Orientating remarks

Textbooks are the most frequently used learning support material (LSM) and the availability of high quality textbooks is one of the critical factors in the successful implementation of educational reform (Department of Education of South Africa, 2009; Cocking, Mestre and Brown, 2000; Asmal, 1999). However, in reaction to some teachers' "overuse of the textbook" and traditional learner-passive teaching strategies, the introduction of activity-based OBE led to the view that teachers and learners do not need textbooks (Taylor, 2008a; Land, 2002). The Chisholm report states that the initial OBE training even supported this view. But the report also states unequivocally that this view is contrary to the policy of the Department of Education of South Africa (2001a) and government spending on textbooks confirms this (Ncokwane and Prabhala, 2006; Department of Education of the Western Cape, n.d.; Asmal, 2002; Land, 2002). To set the record straight the minister of Education admitted in 2009 that "there was a strange anomaly in our system in which the importance of textbooks was no longer appreciated"

(Motshekga, 2009) and the Department of Education stated in *Planning for 2010 and beyond* (Department of Education of South Africa, 2009, 6) that

> . . . textbooks play a vital part in teaching and learning. Textbooks must be used by teachers and learners to enhance their teaching and learning.

Since teachers use (and will in future be encouraged to use) textbooks as curriculum guides and sources for preparing lessons, the quality of textbooks will have a great impact on the quality of their instruction (Lemmer et al., 2008; Newton and Newton, 2006; Ogan-Bekiroglu, 2007; Reys and Reys, 2006; Brandt, 2005). It is therefore imperative that they should use the best science textbooks available (Roseman, Kulm and Shuttleworth, 2001). What Hubisz (2003, 2) says about textbooks and teachers in North Carolina and the rest of USA holds true for the teachers in South Africa:

> . . . often these teachers lack the appropriate academic qualifications to be teaching these courses. In fact, in many instances these materials form the teacher's own introduction to the subject. Naturally, with their limited backgrounds, they are heavily dependent on the materials from which they teach. It is especially important, therefore, that the textbooks and other materials that teachers and students are forced to use get it right.

Bernier (1996, 284) defines the concept "quality" in printed educational material as the "congruence between the desired learning outcomes as specified . . . and the actual learning outcomes achieved." Therefore, a good textbook is one that enables learners to achieve the intended learning outcomes (Bernier, 1996). To accomplish this a good textbook should fill the needs of the teacher and the student (Nitsche, 1992) and *enhance* a student's ability to deal effectively with the skills, concepts and content of the subject (Nitsche, 1992). The *quality* of a textbook can also be defined as it's "fitness for purpose" (Sursock, 2001). According to this definition the quality of a science textbook is indicated by it's ability to support the learner and teacher in attaining the goals of science education (Kesidou and Roseman, 2002).

With the introduction of OBE a system was initiated in South Africa, which allowed teachers to evaluate and select their own textbooks (Mahlaba, 2006). Since most teachers have never received any training in textbook evaluation, teachers found it difficult and the National Professional Teachers' Organization of South Africa (NAPTOSA) recommended that criteria for the selection of textbooks should be developed nationally. The evaluation of textbooks and the compilation of lists of approved textbooks for the General Education and Training band (GET) have now become the responsibility of the LTSM sub-directorate of every provincial education department (Gauteng Department of Education, 2006; Mahlaba, 2006; Western Cape Department of Education, 2005), while the list of approved textbooks or national catalogue of textbooks for the Further Education and Training band (FET) is developed by the National Department of Education (Department of Education of South Africa, 2009). Schools that receive their textbooks from the education departments have to choose from this catalogue of approved books. Evaluation of the textbooks is done on national and provincial level to compile the catalogue of approved textbooks and from this list teachers at school level choose the books most suited to their learners and resources (Motshekga, 2009; Mahlaba, 2006).

The concern about the quality of teaching materials is not limited to South Africa (Hsu and Yang, 2007; Dimopoulos, Koulaidis and Sklaveniti, 2005; Kesidou and Roseman, 2002; Ninnes, 2002; Kearsey and Turner, 1999). In the USA, for example, learners' poor achievement in the Trends in International Mathematics and Science Study (TIMMS) caused widespread concern about science education, and the "poor quality of teaching materials, especially textbooks" was considered as partly responsible for the state of affairs (Singer and Tuomi, 2003, 5). This concern stimulated projects in the assessment of science textbooks by the U.S. Department of Education, the National Science Foudation (NSF), American Association for the Advancement of Science (AAAS) and the National Research Council's Centre for Science, Mathematics and Engineering Education (CMSEE) (Singer and Tuomi, 2003; Jepson, 2002; Roseman et al., 2001; Kulm, Roseman and Treistman, 1999). These projects contributed greatly to our understanding of textbook assessment, although no consensus was reached regarding the best procedure for the assessment of science education textbooks. None of them considered the unique outcomes-based context in South Africa.

Various methods are used to assess the quality of textbooks. Some assessments are even made on the basis of *popularity*. This is usually justified by stating that a book that sells well "must be doing something right" (Ansary and Babaii, 2002, 4). The most obvious method of textbook assessment is *experimental investigations*. Experimental investigations measure how well learners reached the desired outcomes when using the textbook. This requires time and large scale research projects (Britton et al., 1993; Mikk, 2000). Usually, pre and post tests are completed by learners and compared with the data obtained from a control group using a different textbook or no textbook at all. This may seem the ideal method for assessing textbook quality, but experimental investigation into textbook effectivity can be hampered by the design of investigations that do not take into consideration other variables that influence learning. Even if the investigation is designed with care, the results are limited to the comparison between two textbooks or between learning with and without the textbook. Furthermore, investigations are usually limited to a chapter or even a shorter piece extracted from the textbook. For example, Britton et al. (1993) studied 62 investigations that compared two or more versions of texts, and reported quantitative, empirical measures of learning, but none of them referred to an entire textbook. The pre and post testing often focus on measuring recall of facts and seldom include measurement of the learners' levels of comprehension or their ability to apply the knowledge (Mikk, 2000).

Another method of assessing textbook quality comprises the use of *opinions* of learners, teachers and experts about the textbook. These opinions can range from respondents who have used the textbook to respondents who have only looked at the textbook in isolation. Some opinions of respondents who have not used the textbooks are based on thorough examination, but often respondents only skim through the table of contents (Ade-Ridder, 1989). Usually, each respondent bases his or her opinion on his or her own set of unspecified or even unformulated criteria for evaluating textbooks (Lemmer et al., 2008). This makes it difficult to summarise and compare their opinions. Consequently, researchers limit the required responses by posing a number of questions or categories to which respondents must react. At the utmost respondents are asked for further comments after answering specific questions. This limits and shapes the respondents' expression and involvement.

Experimental investigations as well as opinions of respondents who have used the textbook are used to evaluate textbooks *during and after use*. These assessments provide valuable information for the revision of textbooks and aid the design of future textbooks (Mikk, 2000). When an education provider or a teacher needs to select a textbook it is, however, impossible to use every available textbook before deciding which one to prescribe for the learners. In this instance, the *analysis of textbooks* provides an effective alternative to experimental investigations and respondent opinions. Textbook analysis is the evaluation of data about textbook characteristics according to explicitly stated criteria. Researchers determine which characteristics of textbooks contribute to learning. The presence of these characteristics are then formulated as criteria for all textbooks of good quality. The analysis of a textbook is often the only viable method of assessing a textbook's quality in isolation.

The number of criteria to be considered in the evaluation of a textbook can be daunting and many different analysis instruments have been developed to help ensure that every aspect is considered. The simplest analysis instruments are *checklists* of desirable criteria. Analysts consider each criterion and indicate whether it is present in the textbook or not. The number of criteria that are satisfied can be counted to give an idea of the quality of the book. Checklists are limited by the fact that they only indicate which criteria are met by a textbook and not *how well* the criteria are met. This can be addressed by adding a rating scale to provide a more informative *analysis instrument*. The rating scale can even include a benchmark against which the textbooks can be compared. It is also possible to attribute weights to criteria to make provision for the influence of more important criteria.

None of the above mentioned methods are perfect, but each has the potential to give us more information on the quality of a book. Mikk (2000) advises the use of two or more methods in conjunction to validate results. Textbook analysis has the best potential to assess textbook quality in a predictive situation, for example during textbook selection. Analysis instruments are available to assist teachers in the assessment of textbooks, but only a few of these instruments address the unique OBE context in which science education textbooks are utilised in South Africa.

1.2.2 Becoming aware of the problem

Teachers need access to high-quality learning materials to facilitate effective classroom learning. This is especially true for teachers with limited experience and teachers without the necessary qualifications who rely heavily on a textbook for support in the planning and presenting of their instruction (Lemmer et al., 2008). When teachers spend less time on lesson preparation due to lack of time, the quality of the textbook as main learning resource becomes crucial (Pepin and Haggerty, 2003, 97). Although it is necessary to assess the quality of textbooks during textbook selection, it is unrealistic to expect unqualified and inexperienced teachers, that have received no training in textbook evaluation, to assess the quality of a textbook without guidance (Ansary and Babaii, 2002). McKinney (2005, 39) remarks that in the South African context

> . . . it seems there are a wide range of practices in textbook 'selection' including a sponsor dictating texts, publishing sales representatives choosing for schools, provincial officials selecting books, and finally educators and heads of learning areas selecting texts themselves.

Some schools in South Africa receive their textbooks from the Departments of Education. These schools have to choose the textbooks from lists of approved textbooks. Even the lists of approved books are not without controversy. Books that are approved by one department are sometimes rejected by other departments as being totally inadequate (McKinney, 2005). Unfortunately the evaluation criteria used are often not available, and this detracts from the transparency of the assessment process. The teachers or administrators included in the evaluation panels do not always have the necessary level of knowledge, experience or training to be able to judge the quality of the textbooks (Pretorius, 2001). Furthermore, even if teachers are provided with a list of approved books the dilemma remains the same. They may not be able to choose from the approved list the textbooks that will suit their specific situations best. In addition, many of them do not know how to assess the quality of the material prepared by themselves or any textbook that was not evaluated by the Department.

Instruments or systems are available to guide the evaluation and selection of textbooks. Most of these instruments are, however, unsuitable for the analysis of science textbooks since it is generic or focusses on other subjects, especially history and language education (Mikk, 2000; Brown, 1998; Garinger, 2003; Ansary and Babaii, 2002). With the introduction of OBE the National Department of Education, for example, issued a very superficial checklist as generic guideline for the development of LSM for OBE (Goodwin-Davey and Davey, 2000).

Although a number of analysis instruments that focus on specific aspects of science textbooks exist (Good, 1993; Kesidou and Roseman, 2002; California Department of Education, n.d.), only a few accommodate the unique OBE approach that is required in science textbooks in South Africa (Western Cape Department of Education, 2005). Examples of instruments for science textbooks in the OBE context were developed by the Western Cape Department of Education (Western Cape Department of Education, 2005) and the Gauteng Department of Education (Mahlaba, 2006) and recently also by the National Department of Education (Department of Education of South Africa, n.d.). Each of these instruments contains a generic and a subject specific component. The three instruments will be discussed in greater detail in Section 3.9.4. Unfortunately the Western Cape and National instruments do not include rating schemes and can therefore not be used to compare textbooks. The Gauteng instrument (LOETA) is the most thorough and extensive evaluation instrument, but unfortunately it is not specifically developed for physical science textbooks in the FET band (Mahlaba, 2006).

The need exists for an analysis instrument that is appropriate for and will provide scientific justification and transparency to the analysis of textbooks to be used in science education in South Africa. The ideal instrument must be suitable to outcomes-based education, but also be adaptable to address future changes in education policy.

1.3 Statement of the problem

1.3.1 General problem statement

It is important to assess the quality or appropriateness of textbooks, but multiple criteria of differing importance make it a difficult and complex task. Suitable methods and instruments must be available to make the evaluation process easier, more transparent and where possible, more objective.

With textbooks of varying quality flooding the South African market, education providers and science teachers (especially unqualified, under-qualified or inexperienced teachers) need assistance in the assessment of textbook quality for textbook selection. Unfortunately no method or instrument exists for the assessment and comparison of the quality of science education textbooks that can be used in OBE and can be easily adapted to incorporate changes in educational policy. Therefore, the assessment process is worthy of investigation.

1.3.2 Research problem

In pursuance of an appropriate and adequate instrument to assist evaluators in the assessment and comparison of science textbooks, a number of questions arose that directed the focus of this investigation.

The first question that arose was: What is the most appropriate method or instrument for the evaluation of the quality of science education textbooks?
In order to identify the most suitable method or kind of instrument to guide the assessment process, all available methods and instruments used for the assessment of textbook quality were investigated. The different kinds of information provided by the different methods of assessment, the reliability and validity of the results they produce, as well as the ease of application had to be taken into consideration.

The next problem that was encountered was: What is the most scientifically justifiable approach to follow in developing an appropriate instrument to guide the assessment process?
Once the most suitable method or kind of assessment instrument was identified, a logical, systematic approach for the development of such a method or instrument had to be devised. Unfortunately most methods and instruments offered in the literature did not include documentation regarding the development process or justification for the choice of criteria. Where possible the different approaches for the development of different methods were investigated. The assessment of text-

book quality is a challenge that involves multiple criteria which do not contribute equally to the quality of the textbook. The investigation was, therefore, also extended to techniques used to handle multi-criteria problems in other fields such as Operations Research. Once an appropriate approach for developing a suitable method or instrument was identified and described, an instrument were developed utilising this approach.

Once an analysis instrument had been developed by utilising the chosen approach two further questions should be answered: Does the instrument provide reliable and valid information about the quality of textbooks? Does this approach provide a workable and transparent process for the development of analysis instruments? To answer these questions the instrument was implemented on a pilot basis and thereafter on a larger scale. The results of the different applications were analysed to determine the reliability of the instrument. To determine the instrument's validity the obtained analysis results were also compared to the results acquired using other assessment methods, such as the opinion and experience of teachers who have used the relevant textbooks.

1.4 Aims of the investigation

1.4.1 Immediate aims

The immediate aims of this investigation were to

- investigate learning and teaching strategies in science with special reference to the South African context;
- investigate the unique demands of science as subject on educational textbooks;
- investigate possible ways to measure the quality of textbooks;
- investigate different approaches to the development of textbook analysis instruments and identify an approach or approaches suitable to the SA context;
- develop a suitable instrument for the assessment of the quality of science education textbooks in the OBE context, which can also be adapted for use in other contexts;
- launch a pilot application of the instrument, investigate the application process, analyse the results that are obtained and modify the instrument;

- apply the instrument on a wider scale;

  and

- compare the analysis results with the opinions obtained from teachers who have used the relevant textbooks.

Through the immediate aims described in this section, the following distant aims can be realised.

1.4.2 Distant aims

A suitable instrument can be of great value in the *selection* of the most appropriate science education textbooks. Identifying the most appropriate textbooks will contribute to the quality of science education and thus indirectly support the development of the full potential of every learner: be it scientific literacy or a successful career in science.

Thorough assessment of textbook quality during textbook selection can lead to *improved quality* in future textbook editions. The assessment results can make publishers aware of faults or inadequacies in textbooks and encourage them to produce better textbooks, by not accepting books that are inadequate (Tyson, 1997; Chambliss, 1994). If publishers know how the textbooks will be evaluated and teachers and education authorities buy only books that receive good results in the evaluation, the publishers will try to improve their textbooks to receive better results in the evaluation, in order to sell more textbooks (McKinney, 2005, 50). As the Mozambique Ministry of Education (2002) summarises:

> A successful textbook evaluation system will ensure over a period of time continuous improvement in the quality and availability of learning and teaching materials.

Any investigation that focusses on teaching materials *contributes to our understanding* of its role in teaching and learning. According to Tomlinson (Ansary and Babaii, 2002, 5):

> The process of materials evaluation can be seen as a way of developing our understanding of the ways in which it works, and in doing so, of contributing to both acquisition theory and pedagogic practices. It can also be seen as one way of carrying out action research.

1.5 Research method and design

A literature survey was undertaken to obtain an overview of what science education as well as the role and functions of science education textbooks comprise and to discover what methods are used to assess textbook quality and what instruments are available to aid the assessment. The available literature on different approaches to the development of assessment instruments was considered.

Both qualitative and quantitative techniques were used in the development of the instrument, utilising teachers and experts in science education, instructional design, media and OBE.

The implementation phases incorporated quantitative statistical methods to compare the results obtained by different analysts. It also incorporated qualitative interviews to gather information on the analysts' experience of the instrument during the application process.

To provide the opportunity to determine the validity of the results of the analysis instrument, the application phase was followed by interviews with teachers who have used the relevant textbooks.

1.6 Concept elucidation

In this research the assessment of textbooks was investigated in the context of *science education*, where science education refers to Physical Science education or learning and teaching Physics and Chemistry. Science focusses on investigating physical and chemical phenomena in the world around us and tries to explain and predict events in our environment (Department of Education of South Africa, 2002a). Science education guides all learners in the development of scientific literacy as well as those who has aptitude and interest in their preparation for scientific careers. The focus will be on science education in South Africa's OBE system, with reference to relevant recent research results on Science learning.

The term "textbook" or "schoolbook" has been used in different contexts to refer to different concepts. Some authors use it to refer to any book used during education, but most limit it to books written specifically for classroom use (Mikk, 2000). Traditionally textbooks were print-based media and print-based media remain *economically* and *practically* the most viable option for the support of teachers in the design of learning programmes. According to UNESCO (2002) the production cost of audio and CD-ROM in 1998 was respectively 35 and 40 times that of print per student learning hour. It is not only the production cost that limits the use of other technologies but often the absence or unreliability of infrastructure like electricity or telephone lines (Monge-Najera, Rossi and Mendez-Estrada, 2001) and the expertise to maintain the necessary equipment, like video recorders and computers. Textbooks are portable, convenient and can be used as lead medium or as a supplementary resource (UNESCO, 2002).

The *future of textbooks* in education is often questioned when technological alternatives are considered. With audio cassettes, videos and a variety of computer-based technologies to choose from, print may seem a dull option, but studies worldwide show that textbooks are still widely used (Dimopoulos et al., 2005; Stern and Roseman, 2001; Sitte, 1999). Instead of eliminating textbooks from the future, technology has given us new characteristics for textbooks. It has changed the traditional printed textbook. In developed countries new software packages are being developed that will allow teachers to order textbooks *a la carte* with deletions and additions custom-designed for their Learner and Resource Context (LRC) (DeBolt, 1992). Technology can also change the substance of textbooks. In future textbooks might be available as e-books or even as documents on learners' laptops (Ruttimann, 2006). Textbooks on laptops can be instantly updated or supplemented by the teacher via a wireless network if there is a hotspot in the science laboratory or school. Through this network a teacher could, on a daily basis, add guidelines or resource material for learning activities to the textbooks on the laptops of learners without the hassle of cables. Libya is the first government to reach an agreement with the American nonprofit organization, One Laptop per Child, to provide all its schoolchildren (1,2 million) with inexpensive ($100) laptop computers (Foreign Policy, n.d.). This study will be limited to printed textbooks, but the approach to instrument development used can be replicated for other types of textbooks.

In Section 1.2.1 the concept "quality" in printed educational material is discussed. Two definitions are mentioned: the first being "achievement of the intended learning outcomes" and the second "fitness for purpose" (Bernier, 1996; Sursock, 2001). According to these definitions the quality of science textbooks lies in how well it is able to support the learner and teacher in attaining the goals of science education.

In the *assessment* of the quality of science textbooks the "quality" is regarded as the latent variable (Bernier, 1996) and it can be measured by

- comparing actual learning outcomes with stated goals in *experimental investigations*;

- determining the presence and adequacy of desired characteristics as "indicators of the quality" during *textbook analysis* (Mikk, 2000);
- gathering *respondent opinions* about the presence of desired characteristics in a textbook, the outcomes that have been attained by using the textbook or the support provided by the textbook in the learning process.

The next section is devoted to a discussion of the course of the study.

1.7 Scope and course of investigation

A general orientation and statement of the research problem is covered in Chapter 1. The next two chapters are dedicated to an investigation of the relevant knowledge domains, through a literature study.

If the *quality* of science textbooks is based on its effectiveness in supporting the learner and teacher in their attaining of the learning goals, it is imperative to investigate these goals and the nature of the support that the textbook must provide. In Chapter 2 education in general is discussed. The discussion is extended to science education as field of study, the goals of science education and specific teaching strategies. Special reference is made to OBE. Against this background the nature of science education textbooks is contemplated.

In Chapter 3 the concept "textbook quality" and the various methods and instruments used in textbook assessment are explored to identify the most appropriate method or instrument for the evaluation of science education textbooks utilised in the South African education system. Documented approaches to the development of assessment methods and instruments are investigated, as well as suitable methods and techniques developed in other fields like Operations Research.

The empirical study is reported in the next three chapters. Chapter 4 consists of a description of the design of the empirical study. The criteria used by experts in textbook assessment, that is, the characteristics they consider ought to be present in textbooks and the importance of each characteristic to the overall quality of the textbook, are determined. Thereafter, the design of a pilot implementation of the instrument is discussed, as well as the subsequent analysis of the data and adjustments to the instrument. It is followed by a description of the final implementation of the revised instrument on a larger sample.

In Chapter 5 the results of the empirical study are reported, interpreted and discussed.

Chapter 6 is devoted to a summary of the research, conclusions and possible recommendations.

Chapter 2

Education and science education: Teaching and learning

2.1 Education

Parents and other adults react to a child's inability to cope with life and through spontaneous interaction they facilitate the child's development in this primary education situation. In developed communities this primary education situation is inadequate to prepare the child for his or her life within a complex environment. Formal education institutions, like schools, were established to address this need by supplementing the primary education situation with a secondary, more formal and planned intervention. As Einstein was quoted (New York Times, 2007) in the New York Times of October 16, 1936:

> The school has always been the most important means of transferring the wealth of tradition from one generation to the next. This applies today in an even higher degree than in former times, for through modern development of economic life, the family as bearer of tradition and education has become weakened. The continuance and health of human society is therefore in a still higher degree dependent on school than formerly.

As planned intervention, formal education requires decisions on what to teach and how to teach it. There is no universal consensus on these issues. Actually, a variety of approaches to education exist. This is reflected in the words Aristotle wrote more than 2 300 years ago (Shaw, n.d., 1):

> In modern times there are opposing views about the practice of education. There is no general agreement about what the young should

> learn either in relation to virtue or in relation to the best life; nor is it clear whether their education ought to be directed more towards the intellect than towards the character of the soul . . . And it is not certain whether training should be directed at things useful in life, or at those conducive to virtue, or at non-essentials . . . And there is no agreement as to what in fact does tend towards virtue. Men do not all prize most highly the same virtue, so naturally they differ also about the proper training for it.

Everything we do is guided by our basic philosophy of what we consider to be true and valuable in life. Since this basic philosophy also determines our approach to education it deserves our attention. The following four *schools of philosophical thought* are most often mentioned:

- *Idealism*: Idealists view reality as spiritual, mental and unchanging (Kneller, 1964). They believe that knowledge is gained by re-thinking tried and true ideas (Arif, Smiley and Kulonda, 2005).

- *Realism*: Realism is based upon Idealism and views reality as objective and composed of measurable matter and natural laws. Consequently, knowledge is gained through the use of senses (Arif et al., 2005).

- *Pragmatism*: Pragmatists regard knowledge as worthwhile if it enables people to solve the problems they are likely to meet as adults in a democratic society (Tansey, n.d.). They see man as a social and intelligent being (Kneller, 1964) that can apply knowledge, make predictions and test them (Arif et al., 2005). Pragmatism is related to Instrumentalism (John Dewey), Functionalism and Experimentalism (Wikipedia, n.d.d).

- *Existentialism*: According to Existentialism reality is subjective and a multitude of realisties exist (Arif et al., 2005). Therefore, we should not accept any predetermined philosophical system. We should acknowledge our freedom, accept the responsibility for that freedom and try to define who we are (Wikipedia, n.d.c).

These basic philosophies determine a person's view of the world, which includes education. Therefore, all educational philosophies have their roots in these basic philosophies. In the next section the different educational philosophies, originating from these basic philosophies, are discussed.

2.1.1 Educational philosophies

Educational philosophies flow naturally from the basic philosophies of what is true and valuable, since education is an attempt to provide learners with the knowledge and skills that society regards as the most valuable aspects of culture (Pepin and Haggerty, 2003). To guide the planned educational intervention in the lives of learners, an *educational* philosophy should address questions like:

- What is the purpose of education?
- What outcomes should the learners reach?
- What is the role of the learner and the teacher?

The following four educational philosophies are generally mentioned try to answer these questions from different perspectives:

- *Perennialism*: Perennialists believe in teaching things that are of everlasting importance to all people everywhere (Kneller, 1964). It is based on Idealism and it considers rationality as man's highest attribute. Therefore, scientific reasoning should be taught, not facts. Furthermore, they believe that learners are people first, and workers second if at all, and consequently, liberal rather than vocational topics should be taught (Wikipedia, n.d.b).
- *Essentialism*: Essentialists contend that culture and core knowledge must be transmitted to each new generation. Some philosophers link it to Realism (Arif et al., 2005), while others consider it compatible with a number of different philosophical outlooks (Kneller, 1964). Essentialism is characterised by an emphasis on basic academics (the 3 Rs), respect for the existing power structure, and nurturing of middle-class values (Rojewski, 2002). Children are taught in a thorough and rigorous fashion, progressively, from less complex to more complex skills. Essentialism is often associated with William Bagely, who believed schools must prepare students for a harshly competitive world (Gross, Shaw and Shapiro, 2003).
- *Progressivism*: Progressivism is based on pragmatism (Reed and Johnson, 2000). Pragmatism values knowledge that is useful in solving problems, but the problems in life are constantly changing. It is, therefore, obvious that learning through problem solving should take precedence over repetition of subject matter and children should be taught how to think rather than what to think (Kneller, 1964). Progressivism is usually associated with John Dewey, who regarded humans as social beings who learn best in real-life activities with other people (Reed and Johnson, 2000). Education should

therefore be student-centered and collaborative (Gross et al., 2003), with the teacher in an advisory role (Reed and Johnson, 2000). A typical progressivist slogan is "Learn by Doing!". Children learn like scientists, similar to John Dewey's model of learning (Reed and Johnson, 2000):

1. Become aware of the problem.
2. Define the problem.
3. Propose hypotheses to solve it.
4. Evaluate the consequences of the hypotheses from one's past experience.
5. Test the most likely solution.

In 1957, the orbiting of Sputnik caused a panic in educational establishments as Americans and Europeans felt they had fallen behind the Soviet Union technologically. A rethinking of education theory followed that caused Progressivism to fall from favour. However, progressive education methods, such as hands-on activities and science experiments are still considered important in science education.

- *Social Reconstructionism*: According to Social Reconstructionism education should lead the way towards the creation of a "more equitable society" (Kneller, 1964, 61). To realise this goal, the means and ends of education must be completely re-fashioned to meet the demands of the present cultural crisis. The curricula must embrace social issues and local or global relations (Bailey, n.d.). The teacher must convince his or her pupils of the validity and urgency of the reconstructionist solution, but he or she must do so with scrupulous regard for democratic procedures. Consequently, education must be non-authoritarian, and must promote students' rights for individualised education.

The field of psychology provides us with *learning theories* that try to explain how children learn. These theories give guidance to teachers on how to teach (Sulaiman and Dwyer, 2002). The discussion on the different learning theories that follows is limited to the most prominent theories.

Behaviourism. Behaviourism assumes that all knowledge is enacted as behaviour. Behaviour is determined by stimuli and experiences, a combination of forces comprised of genetic factors and the environment. Therefore, Behaviourism is also known as stimulus-response-psychology and learning is equated with behaviour modification (Klassen, 2006; Carr, 2003). The desired target behaviour is identified and instruction is planned accordingly. The instruction aims at changing the learner's behaviour to the desired target behaviour through techniques such as positive and negative reinforcement. Consequently, behaviouristic education is characterised by features like practice and reinforcement, programmed learning, behavioural objectives, and competency-based education. *Cognitivism*: Cognitivismes focus on the thought process behind the behaviour. Changes in behaviour are observed, but only as an indicator of what is going on in the learner's mind. Knowledge is acquired from within the mind (Watkins, 2000) and cognitivism emphasises internal processes, information processing and memory.

Constructivism: Constructivism proposes that individuals learn by constructing meaning through interacting with and interpreting their environments (Klassen, 2006). Learners are not "empty vessels" ready to receive knowledge. They start with alternative conceptual frameworks for understanding phenomena (Halloun, 2007; DiSessa, 2006; Carey, 2000; Cocking et al., 2000; Ridgeway and Dunsion, 2000). Therefore, learning can be seen as *assimilation* or *conceptual expansion* (Hewson and Lemberger, 2000), a process of connecting new knowledge to existing knowledge (Renner, 2005; Ulerick, n.d.; Watkins, 2000). When students have misconceptions or conceptual frameworks that are at odds with scientific ideas, *accommodation* or *conceptual change* is necessary for meaningful learning of science (Lynch, Taymans, Watson, Ochsendorf, Pyke and Szesze, 2007; Costu, Ayas, Niaz, Unal and Calik, 2007; DiSessa, 2006; Kozma, 2001; Carey, 2000; Cocking et al., 2000; Hewson and Lemberger, 2000; Shiland, 1997). Smith et al. (1993) define conceptual change as the "realigning, reorganizing or replacing" of existing conceptions to accommodate new ideas. It can even be considered a *paradigm shift* (Halloun, 2007). According to Matthews (2002) the most important contributions of Constructivism to science education is that it alerted teachers to learners' prior knowledge and its role in learning.

Both conceptual expansion and conceptual change require active engagement of learners (Clough, 2006). Learners must be aware of their meta-cognition or the control they have over their own thoughts and learning (Hewson, 2004; Davies and Ward, 2003; Thiede, Anderson and Thierriault, 2003; Parkinson, 2002; Hynd, Holschuh and Nist, 2000; Dillemans, Lowyck, Van der Perre, Claeys and Elen, 1998; Tella, 1998). Learners have to know whether they understand or not (Otero and Campanario, 1990). White and Frederiksen (1998, 4) even plead for "curricula that help students to develop an awareness of their own inquiry process and an ability to reflect on it."

Social constructivism, described by Vygotsky, regards the process of knowledge construction as the result of social interaction and stresses the importance of the teacher's role as facilitator and the implementation of collaborative learning situations or group work opportunities for negotiating meaning (Van den Berg, 2004; Tella, 1998; Solomon, 1994a). Vygotskij defined the *zone of proximal development* (ZPD) (Klassen, 2006, 45). He contends that the learner is only able to reach a specific level of understanding by himself or herself. This level of success can be increased to an optimal level (ZPD) by communication with others or the use of a learning tool (Brändström, 2005).

Society is no longer a "stable entity" (Dillemans et al., 1998). The explosion in knowledge creation and technological development has accelerated to such an extent that it is no longer possible to foresee, and therefore equip learners with the knowledge and skills that will be necessary to succeed in the future world in which they will have to live and work. For example: in a relatively short time the use of cell phones and computers have become basic skills, while new knowledge on health and safety issues have changed the products on our shopping lists and even the way we behave at work, in public and at home.

Though it is impossible to predict exactly what the future holds, it can safely be assumed that our world will keep changing at a rapid pace and thus keep requiring new knowledge and skills (Ausiello, 2007). It has, therefore, become imperative for every learner to continue learning, even after their formal education, in order to be able to upgrade their knowledge and acquire skills that will become necessary at some time in their future. Throughout the world preparation for this *lifelong learning* has now become an essential component of education (Dillemans et al., 1998). This includes the skills and even the right attitude necessary for lifelong learning, without which people will find themselves excluded from job opportunities and participation in everyday life.

2.1.2 Outcomes-based education

All across the globe new curricula are being developed to provide learners with an education that will prepare them for an unpredictable future (Aldridge et al., 2006; Maree et al., 2006; Orpwood, 2001; Rogan and Grayson, 2003). According to Lisichkin (2007, 25)

> The traditional, historically established method of modernizing school programs is relatively simple. Its basis will be an old paradigm that, as a rule, has proved its worth but for some reason needs to be replaced.

In South Africa it was a new political dispensation that paved the way for a whole new education system (Jansen, 2003). Christian National education (CNE) was replaced with outcomes-based education (OBE) in a bid to provide learners, in diverse situations, with an education that better enables them to take part in economic and social life (Department of Education of South Africa, 2001a). OBE is an approach to education that is built around a belief system (or paradigm) that states that education should *strive toward an education where all students become successful learners*.

Most educators and education departments pursue an education approach that includes tenets from different, and sometimes even competing educational philosophies. This is also true with the regard to OBE in South Africa. According to Steyn and Wilkinson (1998) OBE is grounded in *Behaviourism, Pragmatism, Critical Theory and Social Constructivism*. It is also associated with Benjamin Bloom's mastery learning, Tyler's curriculum objectives and the competency educational models of the United Kingdom (Sarinjeve, n.d.).

True to the *behaviouristic* approach, all variations of OBE focus on the outcomes to be achieved by learners, and instruction is then planned to guide learners to achieve the desired outcomes (designing down principle). *Traditional OBE* emphasises subject academic outcomes, while *transitional OBE* accentuates interdisciplinary outcomes. South Africa opted for *transformational OBE* that focuses on what is essential for all learners to be able to do successfully at the end of their learning experiences (Spady, 1994; Killen, 2003). This choice of learning outcomes is the legacy of *Pragmatism* (discussed in Section 2.1). The chosen learning outcomes are knowledge, skills, attitudes etc. that can empower learners to solve the problems they are likely to meet as adults in a democratic society (Tansey, n.d.). OBE is related to *competency-based education* that was developed to ensure that learners, in addition to acquiring subject knowledge, develop the skills necessary to succeed in life (Van den Berg, 2004). This is reflected in South Africa's learning outcomes as stated in the National Curriculum Statements. The learning outcomes cover all the levels of Benjamin Bloom's taxonomy of educational objectives or outcomes (Killen, 2002). *Critical theory* holds that no set truth exists and it contributes to OBE by identifying "critical thinking" as a valued learning outcome (Van den Berg, 2004).

Curriculum reform is not only changing *what* is taught, but also *how* it is taught (Powell and Anderson, 2002). This does not only imply a change in the teacher's role. Anderson and Helms (2001) contend that changing what learners *do* is the "bottom line" of science education reform. It is the area where almost all reforms fall short. Carey (2000) advocates a change in the whole culture of the classroom. In accordance with the *behaviouristic* approach students in OBE are required to "show" that they have achieved the required learning outcomes (Department of

Education of South Africa, 2005). The assessment standards in the National Curriculum Statements state (in detail) how learners are expected to show that they have achieved the required learning outcomes (Department of Education of South Africa, 2005). Assessment is designed to measure observable actions that illustrate the achievement of the predetermined standards (outcomes) (Renner, 2005). The *social constructivist* influence on OBE can be observed in the attention given to collaborative learning situations or group work opportunities and the teacher's role as facilitator of knowledge construction (Van den Berg, 2004).

The South African version of OBE was derived from US and Australian models (Jansen, 2007; Sarinjeve, n.d.; Pryor and Lubisi, 2002). Curriculum 2005 was South Africa's first attempt to implement the paradigm, precepts and principles of transformational OBE and according to Jansen (2007, 38) the Australian model

> Translated into the African context with enormous complexities added-in, OBE became completely unworkable in the impoverished settings of a developing country, and was radically revised three years ago – at great cost to the country.

The original Curriculum 2005 has been replaced by the Revised National Curriculum Statements (RNCS) (Aldridge et al., 2006; Coetzer, 2001). Twelve critical and developmental outcomes (goals) have been defined. These are cross-field outcomes that apply to every phase of education and to all the learning areas. The critical outcomes require learners to be able to:

- identify and solve problems and make decisions using critical and creative thinking;
- work effectively with others as members of a team, group, organisation and community;
- organise and manage themselves and their activities responsibly and effectively;
- collect, analyse, organise and critically evaluate information;
- communicate effectively using visual, symbolic and/or language skills in various modes;
- use science and technology effectively and critically showing responsibility towards the environment and the health of others; and
- demonstrate an understanding of the world as a set of related systems by recognising that problem solving contexts do not exist in isolation.

The developmental outcomes require learners to be able to:

- reflect on and explore a variety of strategies to learn more effectively;

- participate as responsible citizens in the life of local, national and global communities;
- be culturally and aesthetically sensitive across a range of social contexts;
- explore education and career opportunities; and
- develop entrepreneurial opportunities.

The NCS specifies learning outcomes for every learning area. These learning outcomes apply to specific knowledge and skills in that learning area and will contribute to realising the critical outcomes (Olivier, 1997; Spady and Schlebusch, 1999). The learning outcomes state the outcomes (skills etc.) that must be reached, as well as the core or minimum content knowledge that must be included (per phase), although it is the teachers' responsibility to match specific content to specific outcomes. Teachers have to design the learning programme, by starting from the learning outcomes stated in the national curriculum and designing learning activities that will enable learners to reach the outcomes. Consideration is also given to cross-field outcomes and progression in complexity with grades.

The OBE principle of high expectations is in accordance with Spady's philosophy that "All children can learn given the necessary support" (Department of Education of South Africa, 2002a, 14). Although Spady assumes that all learners can reach the desired outcomes, he admits that they will not necessarily all reach it on the same day or in the same way (Lubisi, Wedekind and Parker, 1997). Accordingly, teachers in South Africa are required to supply postmodern learner-centred and activity-based learning programmes with a variety of activities that will enable every learner to reach the stated outcomes in the way that suit his or her learning style (Robinson, 2002; Department of Education of South Africa, 2001a). This is the OBE principle of expanded opportunities and illustrates OBE's kinship to *mastery learning* that aims to provide learners (with different learning styles) with the time and opportunities they need to master the content (Van den Berg, 2004).

Many characteristics of OBE in South Africa are not unique to OBE. According to Vosniadou (2001b) curricula and instruction world wide are changing as a result of research that has offered us new insights into the learning process and the development of knowledge in many subject-matter areas. Curricula and instruction are becoming more student-centred than teacher-centred in an attempt to connect the content to real-life situations, and focussing on understanding and thinking rather than on memorization, drill and practice (Vogler, 2006). Teaching strategies can include activities from behaviouristic repetition exercises to constructive conceptual change activities. Ruttimann (2006, 1) gives the following as an example of a science lesson:

> Toxic chemicals leak into a lake and only you – a doctor, environmental scientist or government official – can stop it.

The challenge, especially in OBE, is to choose teaching strategies that will facilitate the desired outcomes. Killen (2002) advises teachers to use the taxonomy for learning, teaching and assessment that was developed by Anderson and Krathwohl to align outcomes with teaching strategies. This taxonomy was developed as a revision of Bloom's taxonomy of educational objectives, from the perspective of recent theories of cognitive development and can be used to link outcomes with corresponding teaching strategies.

The implementation of OBE in South Africa started from an emotional political introduction of the concept in Curriculum 2005 in 1997. This was replaced by the less complex National Curriculum Statements in 2000. In 2009 Minister Angie Motshekga (Motshekga, 2009) declared OBE in South Africa "dead" (Motshekga, 2009), but it was not replaced by a new education policy. The existing policy was only amended to reduce the burden on teachers, by reducing the number of policy documents, reducing the number of projects, ending the recording of formative assessment marks, and emphasising the use of textbooks (Motshekga, 2009; Department of Education of South Africa, 2009).

2.2 Science education

Early inventors did not have a basic training in scientific fundamentals, but since the beginning of the 19th century scientific understanding was the driving force for the development of technology. Since then new technologies have provided new tools for research, which in turn has increased scientific understanding. This snowball effect is accelerating the pace of "science-driven change" in our society (Lederman, 2008) and science as educational subject has become more important than ever before. According to Lederman (2008, 1)

> ... we have arrived at a point in history when there must be a major increase in the capability of ordinary people to cope with the scientific and technological culture that is shaping their lives and the lives of their children.

Science education requires an approach to education that is consistent with the unique nature and content of science as subject (Aleixandre, 1994). It is, therefore, imperative to consider the nature of scientific knowledge when learning and teaching science is explored. The nature of science is discussed briefly in the next section, to provide the background for the reflection on science education in the remainder of Section 2.2, where the goals of science education and the learning and teaching strategies employed to reach those goals will be considered. In Section 2.3 the textbook and its role in science education will be discussed.

2.2.1 Nature of science as field of study

The nature of science has been the focus of continuous philosophical contemplation. Many scientists and philosophers have considered what science is (ontology) and how scientific knowledge is produced or acquired (epistemology) (Ladyman, 2002). Views on the essence of science are continually changing with developments in various scientific disciplines, but a very definitive shift in paradigm occurred around 1960 (Klassen, 2006).

Prior to this shift the logical empiricists' ontological and epistemological view of science knowledge dominated the scene. This view, which is still accepted by many scientists, regards scientific knowledge as existing truths or facts about phenomena of nature. Hence, scientific knowledge can be acquired through objective observation, induction and deduction according to *the scientific method* (Walliman, 2005; Abd-El-Khalick and Lederman, 2000).

In his watershed book, *Structures of Scientific Revolutions*, Thomas Kuhn coined the concept "paradigm" to describe the generally accepted perspective of a particular discipline at a given time (Brackenridge, 1989). Kuhn described scientific knowledge as existing within a specific paradigm and within this paradigm scientific knowledge grows by refining and expanding the existing knowledge. According to Kuhn the existing paradigm was both determined by and reflected by the textbooks of the time (Guisasola, Almudi and Furio, 2005). Major developments in science, like the development of quantum theory, cause new paradigms to replace the old ones (and new textbooks to replace the old ones with the scientific knowledge rewritten from the perspective of the new paradigm) (Kindi, 2005). Kuhn referred to these as *revolutions in science* (Brackenridge, 1989; Abd-El-Khalick and Lederman, 2000).

According to Rudolph (2000) the accepted view of the nature of science has since 1960 shattered into an assorted array of different perspectives. Some general characteristics are, however, accepted by most post-modern scientists and philosophers. Scientific knowledge is not seen only as phenomena of nature. It also includes constructs (theories and models) that have been invented in attempts to interpret and explain these phenomena (Halloun, 2007; Harrison, 2001; Hodson and Prophet, 1994; Driver, Asoko, Leach, Mortimer and Scott, 1994). In 1938 Einstein and Infield explained it in their book, *The Evolution of Physics*, (Driver, 1994, 41) stating that:

> Science is not just a collection of laws, a catalogue of facts, it is a creation of the human mind with its freely invented ideas and concepts. Physical theories try to form a picture of reality and to establish its connections with the wide world of sense impressions.

Wenham (1995) contends that a scientific theory is the best explanation for the specific phenomenon that scientists have developed up to the present. Scientific knowledge is therefore tentative and subject to change as new evidence or better explanations become available, although core ideas that have survived many attempts to disprove it are more durable (Guisasola et al., 2005; Abd-El-Khalick and Lederman, 2000). Consider as an example the following contention of Asimov (1981) on the "truth" of the law of momentum conservation:

> All we can say is that at no time under any condition have we observed the law violated . . . Furthermore, all the consequences we deduce on the assumption that the law is true seem to make sense and to fit with what is observed. Scientists therefore feel they have ample right to *assume* (always pending evidence to the contrary) that the law of momentum conservation is a 'law of nature' that holds universally through all of space and time and under all conditions. (Asimov, 1981, 22)

More recent perceptions of science argue that observations are never objective; they are always theory-driven. There is not one scientific method; scientific knowledge is acquired by a multitude of different methods and creative, logical explanations of the observations with validation through further observation and social negotiation (Osborne, 2002; Driver et al., 1994; Sandoval, 2003). According to Halloun (2007, 657) scientific conceptions correspond to physical realities, but

> this does not ignore the fact that scientists invent their conceptions (just like ordinary people do) in order to reconstruct, in a convenient way, what they represent in the real world.

The view of science preferred by the South African National Department of Education is reflected in its definition of Physical Science:

> Physical Science investigates physical and chemical phenomena. This is done through scientific inquiry, application of scientific models, the-

ories and laws in order to explain and predict events in the physical environment. (Department of Education of South Africa, 2005, 7)

This corresponds with the idea that all scientific explanations should meet the criteria of concurring with experimental and observational evidence and should make accurate predictions, when appropriate (Nola and Irzik, 2005; Hand, Prain, Lawrence and Yore, 1999). These criteria distinguish scientific explanations from explanations based on myths, personal beliefs, religious values, mystical inspiration and superstition.

This discussion on what science is and how it is produced brings us to the dilemma of what learners should be taught in science education. In next section various perspectives on this question will be cited and discussed.

2.2.2 Science education goals

The *goals of science education* is the subject of a continuing debate between different stakeholders that have different interests and, therefore, prefer different curricula (Osborne, Ratcliffe, Collins and Duschl, 2006; Knain, 2001; Sleeter and Grant, 1991). The essence of the debate is the balance in the science education curriculum, between *scientific literacy for all* on the one hand and the *preparation of scientists for scientific careers* on the other (Walford, 1985; Abd-El-Khalick and Lederman, 2000; Osborne, 2002). A whole spectrum of opinions exists regarding this balance. Science education has its origin in one extreme of the spectrum: preparation of future scientists. But the balance has shifted. The rapid development of scientific knowledge and its application in our daily lives have escalated the need for scientific literacy (Ausiello, 2007; Leitte, 2002). Presently, most educators recognise scientific literacy as a perennial goal in science education (Lisichkin, 2007; Wieman, 2007; McEneaney, 2003; Murphy, Beggs, Hickey, O'Meara and Sweeney, 2001; Brickhouse, Dagher, Shipman and Letts, 2000; Abd-El-Khalick and Lederman, 2000; Rudolph, 2000), but usually not to the exclusion of preparation of future scientists. Humanists like Osborne (2002), who would prefer science education to aim exclusively at the scientifically literacy extreme of the balance, are in the minority.

Most science education policies contain a combination of the two goals, with different ratios with regard to time and importance. South Africa's National Education Department is no exception. It refers to both goals. The Senior phase (Grades 7–9) of the GET band focuses on scientific literacy as goal (Department of Education of South Africa, 2001a, 22), while the Curriculum for the Further Education and Training band (FET, Grades 10–12) also aspire to "provide the learner with expertise and special knowledge and skills to join Higher Education and training or career pathways" (Department of Education of South Africa, 2002a, 12).

As an important goal of science education, *scientific literacy* has received a good deal of attention from researchers and educators (McEneaney, 2003; Gräber, Nentwig, Becker, Sumflecth, Pitton, Wollweber and Jorde, 2001; Abd-El-Khalick and Lederman, 2000; Rudolph, 2000; Shen, 1975). Scientific literacy implies an understanding of the nature of scientific knowledge (Klassen, 2006) and the ability to use scientific principles and processes to make their own decisions and participate in discussions and decisions with regard to scientific issues (Gräber et al., 2001, 61). The scientific literacy can be of a practical, civic and cultural nature and it can vary from planning a nutritional dinner to appreciating the laws of physics (Shen, 1975). Knowing and understanding the language of science is an essential component of scientific literacy (Hsu and Yang, 2007; Osborne, 2002). According to Holliday, Yore and Alvermann (1994, 878):

> Scientific literacy involves the location and comprehension of scientific information, the adoption of a contemporary view of science, the development of informed conceptions, opinions, and beliefs, and the ability to communicate these ideas and persuade others of their veracity.

The description of scientific literacy given by (Holliday et al., 1994) links up with the themes of scientific literacy that Chiappetta, Sethna and Fillman (1993) identified:

- science as body of knowledge,
- science as way of investigating,
- science as way of thinking and
- interaction among science, technology, and society (STS).

These themes of scientific literacy correspond closely with the learning outcomes stated in the South African National Curriculum Statement for Grades 10–12 Physical Science (Department of Education of South Africa, 2003a), which are:

- scientific inquiry and problem-solving skills,
- constructing and applying scientific knowledge, and
- the nature of scientific knowledge and its relationships to technology, society and the environment.

These statements illustrate how science education aims at guiding learners to understanding and applying knowledge, rather than memorising facts, laws and theories. An adequate understanding of the nature of science is considered to be a central component of scientific literacy (Abd-El-Khalick and Lederman, 2000). Furthermore, Rudolph (2000) claimed that changing learners' conceptions of science requires an explicit approach. The learning outcomes mentioned above exhibit a definite effort by the South African Department of Education to ensure an explicit facilitation of the learners' construction of adequate concepts of scientific knowledge and scientific endeavour.

Huitt, Hummel and Kaeck (2000) provide the following classification of the scientific knowledge mentioned in the second outcome:

- facts – ideas or actions that can be verified
- concepts – rules that allow for categorization of events, places, people, ideas, et cetera
- principles – relationship(s) between/among facts and/or concepts
- hypotheses – educated guesses about relationships (principles)
- theories – sets of facts, concepts, and principles that allow description and explanation
- laws – firmly established, thoroughly tested, principles or theories

The third learning outcome mimics researchers who consider the "development of learners' scientific habits of mind and attitudes towards science" as an important component of scientific literacy. These researchers argue that learners' attitudes towards science should include a willingness to apply scientific habits, as well as willingness to contribute to public discussion about the application of scientific principles (Hand et al., 1999).

According to Maarschalk and McFarlane (1987) science education prior to 1960 was characterised by a "preps and props" approach – it was concerned with the appearance, preparation and characteristics of substances. Science was depicted as a *product* that learners had to acquire, consisting of concepts, models, theories and laws (Knain, 2001). Since 1960 science education has tried to foster the image of science as a *process* or a mixture of *product and process* (Ulerick, n.d.; Boulter and Gilbert, 1996). There has been a move from teaching "the scientific method" to recognizing a whole set of discrete processes employed in the creation of scientific knowledge (Millar, 1994). Emphasis was also placed on the skills necessary to produce new scientific knowledge and to understand and apply existing scientific knowledge. The skills which are involved in science education comprise of intellectual, motor and cognitive skills (Hudson, 1994). The South African National Department of Education refers to the following skills as examples of some skills that are relevant for the study of Physical Science (Department of Education of South Africa, 2002b, 13):

- observation and comparison
- measuring
- classifying
- inferring
- predicting
- communicating
- hypothesising
- designing an experiment
- controlling variables
- interpreting
- formulating models and
- reflective skills

Learning science as a *product* implies memorisation and limited understanding (lower-order thinking skills according to Bloom's taxonomy), while learning science as both *product and process* also demands higher-order thinking skills in science.

To assist the learner to reach these goals teachers need both *scientific knowledge*, *pedagogical knowledge* and *pedagogical content knowledge* (PCK) (Hattingh, 2009; Crawford, 2007; Wilen, Hutchinson, Bosse and Kindsvatter, 2004; Davies, 2003; Iszak and Sherin, 2003; Parkinson, 2002; Flores, Lopez, Gallegos and Barojas, 2000; Abd-El-Khalick and Lederman, 2000; Watts and Simon, 1999). Scientific knowledge is knowledge of the content that teachers must facilitate. Pedagogical knowledge is knowledge of how children learn (Sherin, 2002). Together they form the basis for the pedagogical content knowledge that will be discussed in Section 2.2.4.

The next section will consider the pedagogical knowledge of how children learn and specifically what the learning of science entails.

2.2.3 How children learn science

Educational researchers are continually producing results that contribute to our knowledge about how learners develop an understanding of complex domains such as science. (Holliday, 2003; Lesh and Lovitts, 2000; Simon, 2000; Kelly and Lesh, 2000). Cocking et al. (2000, 2) even call the current bloom in research on learning a "cognitive revolution."

Views of how children learn have shifted from behaviourist learning theories to more constructivist views (Klassen, 2006; Ulerick, n.d.). This was discussed in Section 2.1.1. In science education, presently, Constructivism and especially conceptual change and the alternative conceptual notions held by students continue to dominate science educators' views of science learning and teaching (Lynch et al., 2007; Holliday, 2003). Research on conceptual change has produced results that form the basis for educational practise. Posner, Strike, Hewson, and Gertzog (DiSessa, 2006; Hewson and Lemberger, 2000; Posner, Strike, Hewson and Gertzog, 1982) developed a theory of conceptual change that suggests the following four conditions necessary for accommodation or conceptual change to occur:

- dissatisfaction with existing conceptions
- intelligibility of new conception
- perceiving the new conception as plausible
- considering the new conception as possibly fruitful (additional applications)

Research on conceptual change, as reflected in Posner, Strike, Hewson and Gertzog's theory of conceptual change, has provided the incentive for the development of suitable learning and teaching strategies (Lynch et al., 2007; Nurtaç, Tacettin, Samih and Omer, 2006; Hynd et al., 2000; Ridgeway and Dunsion, 2000; Shiland, 1997; Smith et al., 1993), some of which are discussed in the next section.

2.2.4 Learning or teaching strategies

Teachers need *scientific knowledge*, *pedagogical knowledge* (Wilen et al., 2004; Davies, 2003; Iszak and Sherin, 2003; Parkinson, 2002; Flores et al., 2000; Abd-El-Khalick and Lederman, 2000) and *pedagogical content knowledge* (PCK) to facilitate science learning. Scientific knowledge or content knowledge and pedagogical knowledge were discussed in the previous section. Together they form the basis for the pedagogical content knowledge, that is all the forms of content representation and formulating, like the models, analogies, illustrations, examples, explanations, and demonstrations that teachers use to make the content comprehensible for learners (Erduran, Bravo and Naaman, 2007; Harrison, 2001) and even the limitations of the analogies and metaphors (Harrison, 2001; Glynn and Takahashi, 1998; Alexander and Kulikuwich, 1994). PCK also includes knowledge of the management of the learning environment (small groups, peer review, having access to data) and resources (such as textbooks, readings, laboratory equipment, maps and other media) (Hewson, 2004). In this regard, Halloun (2007, 655) predicated that:

> For efficient learning, teachers need to put together and forth coherent and systematic lesson plans that are flexible enough to accommodate both scientific rigor and pedagogical concerns. They need to put such plans into effect with active student participation, and mediate learning activities in a variety of modes so as to help individual students continuously evaluate and gradually regulate their personal knowledge.

The pedagogical content knowledge includes knowledge of the implementation of a variety of strategies drawn from different learning theories (e.g. Behaviourism, Cognitivism and Constructivism), chosen according to the specified learning outcomes, the content (knowledge and skills) and the learners involved (Holliday, 2003; Lesh and Lovitts, 2000; Simon, 2000; Department of Education of South Africa, 1997). Some strategies on their own have limited use, for example extreme radical constructivist teaching that refuses to convey any knowledge to learners, but instead only creates an environment in which learners can construct their own knowledge. Such an approach is insufficient for the learning of scientific knowledge which has been developed over hundreds of years (Matthews, 2002). Concepts (like atoms, velocity and force) that are used to describe and model phenomena are not revealed to learners by observing physical phenomena, and teacher (or alternatively, textbook) intervention is essential (Driver et al., 1994).

Since conceptual expansion and change are viewed as essential in science education, strategies that promote conceptual change should be developed and implemented (Nurtaç et al., 2006; DiSessa, 2006; Hewson and Lemberger, 2000). As learning can only occur when learners are actively engaged and motivated to learn, science education teaching strategies should also address these aspects (Coetzer, 2001; Sulaiman and Dwyer, 2002; Parkinson, 2002). Furthermore, the very nature of science as subject necessitates the inclusion of laboratory work in science teaching (Bennett and Kennedy, 2001; Parkinson, 2002) and attention to mathematical content (Basson, 2002; Huberman and Middlebrooks, 2000), but also activities that incorporate communication opportunities (Cocking et al., 2000; Hynd et al., 2000; Osborne, 2002; Unsworth, 2001). Lastly, every stage of the learning process requires strategies for assessment of learning outcomes to inform learners and teachers of learners' progress (Hargraves, 2005; Orpwood, 2001; Tamir, 1999; Department of Education of South Africa, 1997). These aspects of teaching strategies in science education will be discussed in greater detail in the next sections.

2.2.4.1 Strategies for conceptual change

Since conceptual change may be a prerequisite for understanding science concepts, learning strategies should provide the conditions necessary for conceptual change, identified by Posner, Strike, Hewson and Gertzog's theory (described in the previous section). Studies, like those done by Nurtaç et al. (2006), have indeed confirmed that a conceptual change approach is more effective than traditional instruction in facilitating acquisition of understanding of science concepts.

The first condition in the conceptual change model comprises *dissatisfaction with existing conceptions*. Research indicated that confronting learners' preconceptions and inducing conflict is effective in learning science (Lynch et al., 2007; Ridgeway and Dunsion, 2000; Cocking et al., 2000; Hynd et al., 2000). All learning and teaching strategies should, therefore, include opportunities for learners to identify their prior knowledge or alternative frameworks (Carey, 2000; Mikkilä-Erdmann, 2002; Driver et al., 1994). This can be accomplished through "pre-teaching strategies" like class or group discussions or writing activities that require learners to explain concepts or predict phenomena (Ridgeway and Dunsion, 2000). Once learners are aware of their preconceptions, conceptual dissonance can be induced through activities that make learners aware that their conceptions are inadequate or incorrect. A wide range of activities (often termed *discrepant events*) can be employed for this purpose, including demonstrations, laboratory experiences, discussions or debates and concept mapping (Ridgeway and Dunsion, 2000; Smith et al., 1993).

The other conditions that enable conceptual change can be met through well planned presentation of new concepts. The use of examples, analogies, metaphors and illustrations can help to establish *understanding* and illustrate the *plausibility* of new knowledge (Orgill and Bodner, 2006; Newton, 2003; Yerrick, Doster, Nugent, Parke and Crawley, 2003; Smith et al., 1993). An example is the analogy between a difference in electrical potential energy in an electrical circuit and a situation where there is a difference in gravitational potential energy, such as a hill or inclined surface (Department of Education of South Africa, 2006). Learners can be persuaded that new conceptions are *fruitful* by providing opportunities for the application of new concepts in a variety of unfamiliar and increasingly complex situations that can also satisfy critical cross-field outcomes (Smith et al., 1993).

One example of an instructional strategy that have been developed to aid conceptual change is the P-IC-D-A strategy developed by Ridgeway and Dunsion (2000). The strategy identifies learners' <u>P</u>reconceptions; <u>I</u>nduces <u>C</u>onflict, Discusses the concept and Amends misconceptions. Ridgeway and Dunsion (2000) stress the fact that the science content should guide the teacher's choice of strategies for each step in the process.

The CORE-strategy also targets conceptual change. The strategy aims to Connect to students' prior knowledge of new content, Reflect on what has been learned and Extend by transferring to new contexts (Hong Kong Department of Education, 2002).

The Conceptual Change Model (CCM) is another teaching or learning model that substantially provides this opportunity for conceptual change. The model is based on the following six stages (Costu et al., 2007, 525):

1. Students become aware of their own conceptions in the beginning of the instruction by thinking about it and making predictions before the activity begins.
2. Students expose their views by sharing them in small groups.
3. Students confront their views by checking and discussing them in groups.
4. Students work to resolve conflicts between their ideas and their observations, thereby accommodating the new concept
5. Students extend the concept by trying to make connections between the concept learned in the classroom and other situations, including their daily lives.
6. Students are encouraged to go ahead, pursuing extra questions and problems related to the concept.

Stinner (1992; 1995) proposed the LEP model of science teaching that satisfies all the conditions necessary for conceptual change by providing activities that interconnect the *logical* plane (laws, models, theories and facts); evidential plane (support for the logical products) and the psychological plane (prior knowledge or intuitive knowledge). The model is represented in Figure 2.1.

2.2.4.2 Active engagement and motivation to learn

For any learning and teaching strategy to succeed, learners need to be actively engaged (Darabi, Nelson and Paas, 2007; Blumenfeld, Kempler and Krajcik, 2006; Clough, 2006; Henson, 2004; Sulaiman and Dwyer, 2002; Coetzer, 2001). Therefore, teaching strategies must motivate learners to become and to stay actively involved in the learning activities. Science education is often considered difficult. A great amount of effort is required for learners to effect conceptual change and achieve the desired outcomes. For this reason *motivation* or learners'

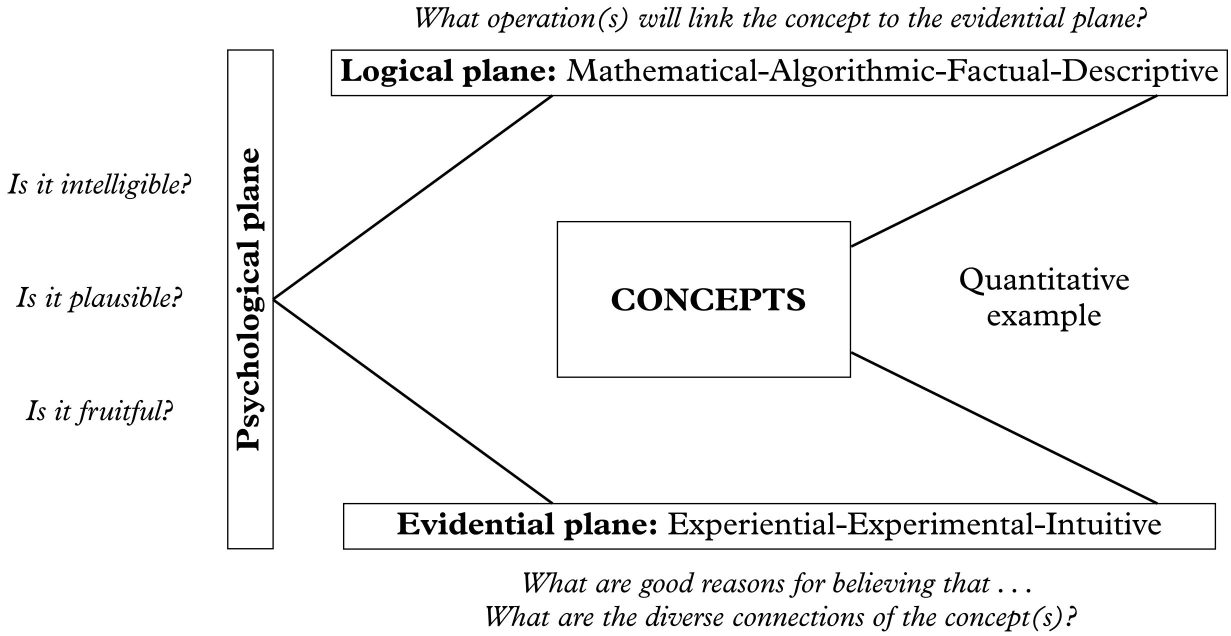


Figure 2.1: The LEP model for concept development in science education (Stinner, 1995)

willingness to invest mental effort (Dillemans et al., 1998) is especially important (Wilen et al., 2004; Hynd et al., 2000). In the words of Blumenfeld et al. (2006, 476):

> Motivation sets the stage for cognitive engagement. Motivation leads to achievement by increasing the quality of cognitive engagement. That is, content understanding and skill capabilities are enhanced when students are committed to building knowledge and employing deeper learning strategies.

The different learning theories have different perspectives on motivation. Behaviourism, for example, attributes motivation to learners' "prior experiences with the environment in response to present stimuli" (Darabi et al., 2007, 60). Thus, motivation to become cognitively engaged can be the direct result of interest, that can be *learner interest*, a permanent personal interest, or *situational interest*, that is a temporary sensation caused by something in our environment (Alexander, 1998). Learner-centred approaches, like OBE, endeavours to use strategies that foster learner interest as well as situational interest to motivate learners to become actively engaged in the learning activities (Flowerday and Schraw, 2000; Department of Education of South Africa, 1997). Cognitive learning theory employs strategies that include the attributes of motivation identified by Paris and Turner (Hynd et al., 2000, 26–27): *choice*, *challenge*, *control* and *collaboration* or applying the constructs of McCoombs' model for motivation: *skill*, *will* and *social support* (Hynd et al., 2000).

Unfortunately, research has shown that not all the strategies aimed at motivating students have the desired effect. Well-known strategies to motivate learners are to project the learning of science as "fun" and getting students to learn without "realizing that they are learning" (Appelbaum and Clark, 2001, 584). This is not aligned to conceptual change learning that emphasises meta-cognition (see Section 2.1.1).

World wide education reform has included a shift from predominantly teacher-centered strategies to learner-centered and activity-based strategies. Learners can easily become or remain uninvolved in the activities of a teacher-centered nature such as lectures, lectures combined with discussion, and direct instruction where learners have a passive role. In contrast, a learner-centered approach to education utilises learner-centered activity-based strategies to ensure learner engagement. This includes activities like project-based assignments, writing assignments, computers software and discussion groups. Effective teachers use both learner-based and teacher-based methods, depending upon the needs of their students and objectives of each lesson (Vogler, 2006, 48). Some activities can even include both approaches. This colligates with the idea of Cronjé (2006) that objectivist and constructivist approaches are complimentary rather than oppositional and learning activities can exhibit elements associated with both approaches.

The new role envisaged for learners in OBE is active involvement in the learning activities and demonstration of their mastering of knowledge, skills and processes in their development towards the specific and critical outcomes (Coetzer, 2001; Olivier, 1997). With the introduction of the OBE approach in South Africa the focus was exclusively on the outcomes and learner-centred and activity-based strategies to gain active learner engagement. This created the impression that subject content is unimportant. According to Taylor (2008a, 1):

> The problem is that, when these OBE devices are foregrounded to the extent that they obscure the subject dissipline, they distract teacher and learners and hinder access to the powerful knowledge that is the birthright of all citizens.

Involvement of learners should not comprise merely doing more hands-on activities (Yore, 2000). Activities should be chosen for the learning opportunities they present, which in turn require active *cognitive* engagement of students (Ridgeway and Dunsion, 2000; Simon, 2000; Boostrom, 2001; Driver et al., 1994; Smith et al., 1993).

2.2.4.3 Laboratory work

The empirical nature of science necessitates the inclusion of *laboratory work* in science education and most countries give prominence to practical work in science education, but no consensus exists over the nature and purpose of activities classed as "practical work" (Klassen, 2006; Jenkins, 2004; Bennett and Kennedy, 2001). The aims of practical work include physical, intellectual and emotional aspects. Bennett and Kennedy (2001) classified the various aims of practical work into eight categories:

- to develop manipulative skills and techniques;
- to encourage accurate observation and description;
- to discover or illustrate a concept, law or principle;
- to experience scientific phenomena;
- to motivate by stimulating interest and enjoyment;
- to develop certain 'scientific attitudes' such as open-mindedness and objectivity;
- to develop an understanding of experimental procedures and evidence and
- to get a "feel" for what it is like to be a problem-solving scientist.

Much of the laboratory work done in classrooms serve only to demonstrate or confirm the existence or correctness of previously discussed concepts (Mumba, Chabalengula and Hunter, 2007; Solomon, 1994b; Osborne, 2002) and most laboratory work is done according to step-by-step instructions (Mumba et al., 2007). Following step-by-step instructions can facilitate the development of physical, manipulative skills and techniques, but its potential to aid cognitive engagement or conceptual change is limited (Watts and Simon, 1999). Cheung (2007, 107) even asserts that many students see laboratory work as "manipulating equipment but not manipulating ideas."

The intellectual demand of any practical activity can be raised by the way it is presented and by the questions that are posed (Parkinson, 2002). When laboratory work is combined with some form of communication it has the potential to provide the dissatisfaction with existing conceptions required by the conceptual change model (Yore, 2000). Lynch et al. (2007, 204) explains as follows:

> Direct experiences with physical phenomena, alongside of the opportunity to make sense of conflicting ideas, reflect on their experiences, and discuss their views with other students, create fertile conditions for conceptual change.

Laboratory work can convey new concepts to learners that even persuade them that the new conceptions are intelligible, plausible and fruitful (the other conditions necessary for conceptual change). However, it must be kept in mind that, unless handled carefully by the teachers, inexperienced learners can obtain erroneous and confusing results that do not enhance understanding (Cheung, 2007; Hodson and Prophet, 1994).

The cognitive level of laboratory work can be increased by incorporating it in an *inquiry* or *problem-based* learning context (Klassen, 2006; Furtak, 2006). Students can then participate in "identifying problems, posing relevant questions, performing efficient and effective experiments, and making judgements on alternative hypotheses and interpretation of data" (Aladejana and Aderibigbe, 2007, 500). This can aid the development of scientific reasoning and scientific habits of mind (Gerber, Cavallo and Marek, 2001; Council of State Science Supervisors, n.d.). Teaching science as *inquiry* consists of two separate parts: teaching science *as* inquiry (an image of science as investigative process) and teaching *by* inquiry (inquiry as the means by which students gain knowledge) (Crawford, 2007; Abruscato, 2004; McNally, 2000; Eltinge and Roberts, 1993). The second part is the most difficult. Teachers find changing to teaching by inquiry difficult, time-consuming (Cheung, 2007; Furtak, 2006; Huberman and Middlebrooks, 2000) and difficult to assess (Eltinge and Roberts, 1993). It often result in "dilution" or watering down of the original inquiry design (Furtak, 2006; Huberman and Middlebrooks, 2000) limiting the activities to "confirmation," the lowest of the four inquiry levels: confirmation, structured, guided, and open inquiry, referred to by Mumba et al. (2007). With regard to the assessment, Bennett and Kennedy (2001) developed a helpful model for assessment of practical work. Even if teachers try to avoid teaching by inquiry, the assessment requirements of the education departments in South Africa and elsewhere will induce teachers to start using it (McNally, 2000).

Laboratory work that forms part of *problem-based learning* bases inquiry or investigative science on real world problems. Traditional teaching decontextualises instruction that can result in learners' lack of motivation. Problem-based or problem-solving learning was, therefore, designed to address this problem (Jonsson, Gustafsson and Enghag, 2007; Hewson, 2004; Cordova and Lepper, 1996). The analogy of student as novice researcher requires that the school experience should contain real-life elements. It is possible to vary the amount of structure built into activities and the degree to which learners have to ask their own questions, design their own investigations and develop their own explanations, depending on their previous experience and understanding (Council of State Science Supervisors, n.d.).

Discovery learning mirrors the scientific process and lets students work on interesting situations to discover concepts and connections for themselves (Cocking et al., 2000; Simon, 2000). To be effective it must guide students and offer opportunities for discovery at a pace that sustains their interest (Henson, 2004; Cocking et al., 2000). The length of time that a person will stay engaged on a task depends, inter alia, on the complexity of the stimulus. If the stimulus is too simple it will become boring and if it is too complex it will appear chaotic. What is needed is knowledge gain at frequent intervals (Simon, 2000). Learners may fail to discover what was intended (Millar, 1994) and teacher intervention may be essential to lead students to develop scientific concepts and models (Driver et al., 1994).

Laboratory work can be complemented by the use of thought experiments (Stinner, 1992) and other "non-laboratory-based pedagogies" (Jenkins, 2004, 172). Special attention should also be given to the development of learners' visual thinking (Simon, 2000).

2.2.4.4 Mathematics in science education

The use of mathematics in science adds to the complexity of the science learning process (Alexander and Kulikuwich, 1994), since learners need to transfer mathematical skills and knowledge to science concepts (Basson, 2002). In response, both science and mathematics education have in the last two decades started to focus on linkage across subjects (Huberman and Middlebrooks, 2000).

A number of mathematical concerns influence science teaching. For example, attention must be given to sequencing instruction to ensure that learners have the requisite mathematical prior knowledge to understand and apply new concepts (Basson, 2002). Another example of a mathematical concern in science education is the use of qualitative reasoning in conjunction with quantitative calculations to improve understanding (Smith, Maclin, Grosslight and Davis, 1997; Alexander and Kulikuwich, 1994)

2.2.4.5 Language and communication in science education

One teaching strategy often mentioned is providing learners with the opportunity to *do what scientists do* (Klassen, 2006; Furtak, 2006; Murphy et al., 2001). This strategy aims to engage students in the same thinking processes and activities of practicing scientists in which they might develop scientific concepts and ideas from their own experience (Furtak, 2006). But this analogy of the student as novice scientist requires a clear image of the actions of researchers. Osborne (2002) and Scantlebury, Tal and Rahm (2007) contend that the image of a scientist as an individual surrounded by test tubes and equipment, ceaselessly exploring the material world, is a delusion. Real scientists spend most of their time modelling or theorising, or evaluating competing theories (Halloun, 2007). Most of their work is not done in the laboratory, but in the writing and reading of papers, e-mail messages and faxes that "fly between institutions" and "in the presentations and arguments engaged in at conferences" (Osborne, 2002, 206). He is supported in this line of thought by Norris and Phillips (n.d., 2) who claim that "constructing, interpreting, selecting and critiquing texts" are as much a part of what scientists do as collecting, interpreting and challenging data.

Communication is thus a pivotal point in establishing knowledge claims and should be an important component in science education (Butler and Nesbit, 2008; Huang, 2006; Black and Hughes, 2003; Moje, Collazo, Carrilo and Marx, 2001; Cocking et al., 2000; Hynd et al., 2000). As such, apprenticeship to the characteristic scientific language is also vital to students' science learning (Galili and Tseitlin, 2003; Unsworth, 2001). Consequently, language and communication skills are emphasised by several teaching strategies (Jonsson et al., 2007; Cocking et al., 2000). In the South African OBE communication is addressed by one of the critical outcomes, which states that learners must be able to communicate effectively (Department of Education of South Africa, 1997) and the then South African Minister of Education, Naledi Pandor, admitted in 2006 that all subjects in the new curriculum require "extensive reading and extended writing" (Pandor, 2006).

Communication can play a vital role in strategies aimed at negotiating meaning for understanding (Butler and Nesbit, 2008; Cocking et al., 2000; Appelbaum and Clark, 2001; Driver et al., 1994). This can be accomplished through student interaction with their peers, textbooks or other printed materials, and with the teacher (Jonsson et al., 2007; Hand and Keys, 1999).

Talking, listening, reading and writing each have a role to play in science education. *Verbal communication* between the teacher and the learners is the most obvious component of communication. It can help teachers determine learners' prior knowledge, monitor conceptual change and plays an important part in providing the conditions necessary for conceptual change discussed in Section 2.2.3. Verbal communication in the classroom can include discussions and presentations by learners. The preparation for a presentation can provide students with the opportunity for reflecting on and organising concepts, while the presentation can provide opportunities for assessment (Ratcliffe, 1999).

The activities of scientists include *reading* and learning from their reading. In the same way students can obtain useful knowledge from textbooks (Ulerick, n.d.). Early research considered reading as a "passive, text-driven, meaning-taking process", but more recent research regard science reading as an "interactive-constructive" process (Hsu and Yang, 2007; Norris and Phillips, n.d.; Penney, Norris and Clark, 2003; Holliday et al., 1994). This process of text comprehension and knowledge construction will be discussed in Section 2.3.1. In order to ensure meaningful learning, teachers should mediate the interaction of students and text, providing meaningful purposes for reading and strategies for meta-cognition (Thiede et al., 2003; Ulerick, n.d.). One example of the strategic use of text is the reading of refutational and explicit text, that has been found to be effective in bringing about conceptual change (Hynd et al., 2000).

Although *writing* as a form of communication is sometimes spurned as time consuming, it helps students refine and articulate ideas (Kesidou and Roseman, 2002; Cocking et al., 2000), provides time for integration of thoughts and requires deeper understanding (Gunel, Hand and Prain, 2007; Cicourel, 1985). The choice of written activity is, however, important. Holliday et al. (1994) draw attention to the fact that some writing activities, like filling in blanks and giving short responses, lack a degree of mental engagement that promotes meaningful learning. In contrast, formulating writing activities in interesting ways can also stimulate motivation and cognitive engagement (Taylor, 2008a; Gunel et al., 2007; Parkinson, 2002).

Communication implies *social interaction*. It may be indirect interaction via written text or direct interaction in lectures, discussions and presentations. Collaborative learning situations or group work opportunities can be utilised for negotiating meaning and for developing the ability to function in a group (Coetzer, 2001; Solomon, 1994a). In South Africa working effectively with others as members of a team, group, organisation and community is also a critical outcome in OBE (Department of Education of South Africa, 2002b, 7). Cooperative learning is an umbrella term for a variety of approaches, methods and techniques that capitalise on the principles of cooperation and group dynamics (Tella, 1998). Although most researchers mention the potential of collaborative learning situations in education, some researchers have also found that collaborative learning is not effective in facilitating counter-intuitive conceptual change (Hynd et al., 2000; Yerrick et al., 2003). Students may reinforce each other's intuitive but non-scientific ideas or cause confusion (Hynd et al., 2000). In his investigation of the relationships between instructional activities and science achievement in Hong Kong, House (2000) found that, although working together in small groups is usually beneficial for learning science, it is not true when utilised to learn a new topic. Furthermore, it can be very difficult to implement group learning in groups of diverse students. Yerrick et al. (2003) and Cocking et al. (2000) found that many teachers fall back on the lecture format simply to have some control over a large group. Parkinson (2002, 73) advises the use of very clear guidelines and even the provision of resources to help structure group discussions. This may include a series of questions or "stimulus material, e.g. pictures, apparatus, descriptions of scenarios or video clips."

All communication provide assessment possibilities (Butler and Nesbit, 2008; Ratcliffe, 1999). One example is students writing reports and records of experiments to demonstrate that they understand the concepts and can use the terminology (Hand et al., 1999). Assessment strategies will be discussed in the next section.

2.2.4.6 Assessment

Changes in science curriculum must be accompanied by corresponding changes in assessment (Orpwood, 2001). Tamir (1999, 401) even calls the present era the "student assessment reform era." If science education is directed at "understanding" as one of its goals then the assessment must measure the learners' understanding. Unfortunately understanding can not be observed directly and assessment activities should probe learners' understanding (Millar and Hames, 2006).

Traditional assessment opportunities (tests and exams) have been augmented by assessment strategies like project work, presentations, demonstrations, oral work, group work, interviews, learner journals, written reports and learner portfolios (Hargraves, 2005; Coetzer, 2001). Another change that has been observed in assessment is a shift from norm-referenced assessment, where students compete with other students, to criterion-referenced assessment that is based on a set of standards of mastery (Henson, 2004; Cowie and Bell, 1999).

Learners' achievements are assessed at different times, by different persons, with different aims (Segers and Dochy, 2006). Different types of assessment can serve different purposes at different stages in the learning process:

- *Summative assessment.* Traditionally, summative assessment was the only planned assessment in the classroom (Brooks, 2002). It was done at the end of a learning unit or course to provide information on the level of achievement reached by the learner (Parkinson, 2002; Skevington, 1994). The summative assessments can be used to ration access to higher levels of the educational system (Pryor and Lubisi, 2002). Internal and external examinations are typical forms of summative assessment. Formal continuous assessment can provide marks or grades which can be incorporated into the summative assessment results (Pryor and Lubisi, 2002).

- *Formative assessment.* This type of assessment has recently become the focus of attention as an important component of learning (Hargraves, 2005). It is a continuous process and is specifically designed to promote learning (Bell, 2000). It gathers feedback about the learners' progress towards the learning outcomes and is used diagnostically (Lumby and Foskett, 2005; Hargraves, 2005; Chatterji, 2003; Brooks, 2002; Cowie and Bell, 1999; Skevington, 1994), to identify the nature and cause of barriers to learning that learners experience and suggest modifications to the learning experience (Department of Education of South Africa, 2002b). Formative assessment is especially important in teaching for conceptual change and conceptual development, since it informs the learners "how their existing concepts relate to the scientifically accepted ones" (Bell, 2000, 49). The marks for formative assessment should not be included in summative assessment marks, since it provides evidence of the process and not of the final outcomes that are reached. In South Africa the recording of these marks added to the teachers' burden since the introduction of OBE, but from 2010 it will no longer be recorded in a portfolio (Department of Education of South Africa, 2009).
- *Baseline assessment.* It measures learners' prior knowledge at the start of a new unit. The teacher can use this information to plan the learning activities. It can be considered to be the first component of formative assessment (Department of Education of South Africa, 2002b).
- *Self-assessment.* Good learning strategies include habits of meta-cognition, like reflection and self-assessment during learning. Henson (2004, 117) points out that students need the opportunity to test their knowledge without penalty, to determine if they have reached the desired outcome or have to continue the learning process. He even recommends the use of "take-home tests." With lifelong learning as educational goal the development of students' learning strategies are receiving more and more attention.

Research has shown that their perception of evaluation demands influences learners' learning and teachers' teaching strategies (Costa, Caldeira, Gallástegui and Otero, 2000; Anagnostopoulos, 2005). For example, students inevitably learn what previous experience predict will be assessed (Anagnostopoulos, 2005; Alexander and Kulikuwich, 1994). This can contribute to or hamper learning. If assessment is limited to lower-order thinking skills and facts, the higher-order thinking skills and application of knowledge can not be reached. The scientific knowledge gained by the learners is then “inert.” It can only be accessed and applied in a restricted set of situations (Vosniadou, 2001a). Teachers with limited training or experience may find it difficult to develop assessment activities and Muwanga-Zake (n.d.) foresees that teachers in South Africa may experience difficulty with recognising and measuring abstract concepts such as critical thinking.

Teachers’ instruction can also be influenced by external assessment. Inadequate national exams, for example, can influence teachers to limit their teaching to what they consider necessary in a bid for high scores (irrespective of the learners’ skills and understanding) (Anagnostopoulos, 2005). This can be due to the fact that the testing of student learning outcomes can be regarded by some teachers as a tool for indirect testing of teaching effectiveness (Smith and Moore, 2005).

2.3 Textbooks

Thousands of years ago, learners in Egypt and Mesopotamia copied texts for practice and learned maxims on ethics and morality in doing so (Newton, 1990). Since then many texts and textbooks have paved the way for learners in mastering many subjects (Bensaude-Vincent, 2006). Some textbooks like Euclid’s *Elements of Geometry* (written in 300 B.C.), Lavoisier’s *Elementary Treatise on Chemistry* (written in 1789) , and Newton’s *Mathematical Principles of Natural Philosophy* (published in 1687) had become famous (Bensaude-Vincent, 2006; Galili and Tseitlin, 2003; Newton, 1990). Kuhn (Klassen, 2006) even contends that all “normal science” is conducted through science textbooks and that it is a necessary requirement for the production of scientists.

Textbooks form part of the larger group of educational media and more specifically printed media. Printed media include, among others, textbooks, educator guides, learner guides, learner workbooks, readers, atlases, dictionaries, magazines, newspapers, charts and posters (Mahlaba, 2006; Stoffels, 2005). Wilson (1997, 6) summarises the positive characteristics of textbooks, saying that

> . . . the book is **portable**, it has **random access** to its contents, especially if the book has an index; the book can also be a **multimedia object**, in that it may contain not only text, but also graphics, drawings and photo-reproductions; it is also **conveniently accessible** in that once you have the book, you need no other artifacts in order to read it (except perhaps a pair of spectacles!), and its energy demands are minimal.

As pointed out in Section 1.6, print-based media still remain economically the most viable option for the support of learners and teachers (Mikk, 2000). Photocopies are much more expensive than printed material and the photocopies used are often illegally made from existing textbooks (Mhlongo, 2006), although most teachers are not even aware that by photocopying material for students, they are breaking the law and could face fines or years in jail (Ncokwane and Prabhala, 2006).

It is very difficult to define the concept “textbook,” because of the variation in literary or pedagogic styles, the way they are used and their philosophical, cultural and historical frameworks (Khutorskoi, 2006; Issitt, 2004). This is confirmed by Herlihy (1992b, v) in his remark that

> . . . as texts are examined, it is immediately apparent that they are not immutable and impersonal documents. What is a text and what is its place in a school are very complex questions.

The available definitions of the concept “textbooks” vary from books that are specially written and published for educational purposes, to any book used in the classroom (Mikk, 2000; Sitte, 1999; Johnsen, 1993), but most researchers refer to textbooks as those books that are specially written for use in didactic situations. They recognise the textbook acts as both “Mittel und Mittler” (content and facilitator) (Sitte, 1999) that contains the information and activities that are necessary to attain the desired outcomes (Khutorskoi, 2006).

In the role of information source, the knowledge represented in the textbooks used by one particular generation represents the knowledge that that society wants to transmit to its children to prepare them for life as worthwhile members of that society (Pingel, 1999, 5). Textbooks also organise the scientific knowledge and science as a discipline (Olesko, 2006). Therefore, textbooks from previous generations are sources of information on changes in science teaching and the bigger issue of the prevailing scientific paradigm described by Kuhn, referred to in Section 2.2.1 (Clericuzio, 2006; Guisasola et al., 2005).

Although some researchers see text as a mere vehicle that delivers this content without influencing learning (Clark, 2001), most researchers recognise the facilitator function of the textbook (Ogan-Bekiroglu, 2007). Since learning is an active constructive process that involves interaction with the textbook, the characteristics of a specific medium must influence the learning process (Kozma, 2001; Johnsen, 1993). The textbook is only a source of potential learning. What the students actually learn from textbooks are mediated by the school context (teacher, peers, instruction, assignments) (Mesa, 2004).

In their role as facilitators, textbooks can be used with a wide range of activities, ranging from individual to group activities and from lectures to inquiries (House, 2000); from introductory activities to application practice (Franssen, 1989a). The textbook's role as facilitator can extend the learning opportunities from the classroom where the teacher acts as facilitator, to the home where quality printed material can facilitate learning and guide learners through appropriate learning activities (Ogan-Bekiroglu, 2007) and even direct them to resources in their own environments. Textbooks provide learners with access to the entire curriculum in an integrated form to which they can refer at any time (Taylor, 2008a). According to Reddy (2005) textbooks are especially valuable to poorer communities where the school and textbooks are the only resources that most learners are able to access.

A textbook can be a book or be divided into a set of study aids, consisting of a textbook, workbook and teacher's handbook (Mikk, 2000; Sitte, 1999; Pingel, 1999). Each of these fulfil a specific purpose. Workbooks are designed to contain exercises and assessment questions and may even provide learners with space to answer them. When learners give written answers to questions in a single traditional textbook they have to answer the questions in a separate book. With the questions and answers in different books it can be very tedious to refer to both books when the learner is doing revision or the teacher is assessing the learner's level of competence. Some teachers prefer to have separate workbooks. This is only a practical arrangement and the questions must be developed and considered as integral part of the textbook (Nazarova and Gospodarik, 2006). Unfortunately workbooks are expensive. It is interesting to note that the Education Department of Namibia, for example, considers workbooks as too expensive for any subject other than mathematics and languages. They will only consider selecting workbooks that are designed in such a way that the learners do not write in the books and are therefore re-usable (National Institute for Educational Development of the Namibian Ministry of Education, 2005). In South Africa the role of the workbook, especially in basic education, is confirmed by the Department of Education's allocation of R 522 million in October 2009 for the development and printing of workbooks in all official languages for basic education (Pretorius, 2009).

As learning methods or textbook writing conventions improve more efficient textbooks can be composed (Mikk, 2000). In South Africa the last change in the science sillabus, prior to the implementation of OBE was in 1985. Many teachers in South Africa have been using the same textbook for about 12 years. In this time research on learning and teaching has provided new knowledge on conceptual change and the strategies that are effective in facilitating the process. Researchers in the fields of education and science are working together to design more effective teaching methods and instructional materials (Cocking et al., 2000).

The textbook was the focus of attention in this discussion. The next two sections are committed to an investigation of the interaction of the learners and the teachers with the textbook.

2.3.1 Learning from textbooks

Learning from textbooks adds another dimension to the complexity of the learning process – the dimension of *text and information-processing*. Science textbooks must represent science in such a way that it supports learners in mastering the discipline. But learners can only learn from textbooks that they can read and understand and textbooks are notoriously difficult to understand (Hsu and Yang, 2007). Consequently, the comprehension of texts deserves further discussion.

Different models of text comprehension can be used to explain how people interact with the text (Stahl, Jacobson, Davis and Davis, 2006; Walsh, 2006; Best, Rowe, Ozuru and McNamara, 2005; Norris and Phillips, n.d.; Boscolo and Mason, 2003; Otero and Campanario, 1990). The *Comprehension Integration (CI) model* of text comprehension was developed by Kintsch (Kintsch and Kintsch, 2005; Best et al., 2005; Boscolo and Mason, 2003; Broer, Aarnoutse, Kieviet and Leeuwe, 2002; Iding, 2000; McNamara, Kintsch, Songer and Kintsch, 1996). Three levels are defined in the comprehension process. The first level is the *decoding process*. Learners must convert the printed words to meaningful sentences in their minds. If the words are too difficult (semantic difficulty) or the reader has trouble putting the words together to form the sentence (syntactic difficulty) the first level of processing will be unsuccessful (Fry, 2002). The second level of text processing is called the *textbase*. The learner must integrate the different sentences to create a coherent text-level presentation (Best et al., 2005; McNamara et al., 1996). The text does not normally provide all the information relevant to the subject of the text (Boscolo and Mason, 2003). The reader has to infer connections and relationships not stated explicitly by the text (Walsh, 2006; Best et al., 2005; McNamara et al., 1996). Research has shown that coherent texts can facilitate this level of processing successfully. Coherent texts provide, where possible, the necessary connections and therefore, require little interference activity from the readers. (Best et al., 2005; Boscolo and Mason, 2003). The third level of the Comprehension Integration model is a processing of the textbase under the influence of the prior knowledge of the reader to construct a *situation model* (Kintsch and Kintsch, 2005). This is a mental model of the situation described by the text. Text content and organization can support learning and it is even possible for the text to moderate the lack of sufficient or accurate prior knowledge (Kendeou and Van den Broek, 2005; Mikkilä-Erdmann, 2002).

While it is obvious that textbooks must not be too difficult for learners to read, McNamara et al. (1996) warn against texts that are easy to understand. They contend that textbooks that are too easy could reduce the amount of active processing. Although this may increase direct recall of the text, it may lead to less effective learning (Boscolo and Mason, 2003; McNamara et al., 1996), since information that is actively generated is beter remembered and better put to use in novel situations (Clark and Salomon, 2001).

When considering the comprehension of texts with illustrations, both Kintsch's CI model of text comprehension and *Paivio's Dual Coding Theory* can provide a theoretical backdrop (Vekiri, 2002; Iding, 2000). Dual coding theory proposes that two distinct and independent but interconnected cognitive systems exist for the processing of verbal and non-verbal information. The dual coding theory can explain why Mayer (2003) found that students learn more deeply from words and pictures than from words alone (under the right conditions). It is obvious that the illustrations in textbooks can contribute to the learning process in science education and, therefore, influence the quality of the textbook. The visual design or layout of a textbook can give salience to some elements in the text: the elements can be marked as more important and more worthy of attention than others. This is attained through the relative size, sharpness of focus, relative positioning, tonal contract and colour contrasts (Hsu and Yang, 2007).

Jenkins proposed a *tetrahedral model of learning* that extends the interaction with text to an interaction of all the components of the learning environment (text or learning material included) (Nelson, 1996; Carney and Levin, 2002). The model involves a four-sided organizational framework representing the four underlying factors which affect learning. Each vertex of the tetrahedron represents a cluster of variables of a given type: characteristics of the learner, nature of the learning materials, learning activities and critical measures or tasks. Each edge represents a two-way interaction that influences learning. Each plane represents a three-way interaction and the whole figure represents the multi-dimensional process occurring from the interaction of all the variables. What constitutes quality in textbooks will be influenced by the textbooks' interaction with each of the other variables.

Penney et al. (2003, 419) stated that the dynamic interplay among variables in Jenkins' tetrahedral model can "change significantly as a function of the scientific

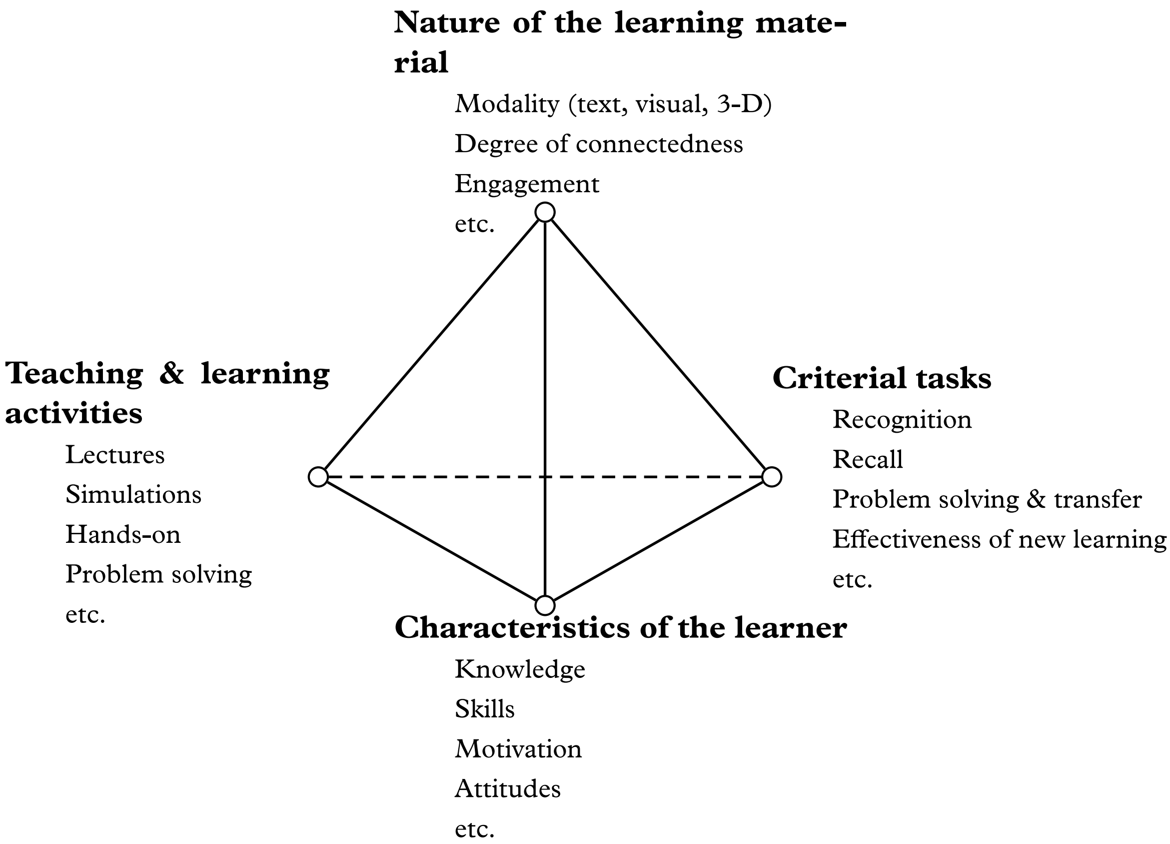


Figure 2.2: The Jenkins tetrahedral model (Bransford et al., 2004, 212)

domain of the study." The model was extended by Alexander and Kulikuwich (1994) to incorporate the context as overarching variable that situates learning and influences the interaction. For example, science texts are infamous for the mathematical and scientific symbols and diagrams that make the comprehension of the text even more daunting (Alexander and Kulikuwich, 1994).

It is my contention that the teacher should not merely be considered as part of the context in science classrooms. The information that teachers give to their learners about the material they use can significantly influence the process of learning from texts (Alexander, 1998). As facilitators of learning the teachers monitor the interactions described by the tetrahedral model. If they judge it to be necessary they can modify the learning material, learning activity or even the tasks during the learning process. The learner and his or her unique characteristics stand in the centre of the learning situation and should be included at the centre of the tetrahedral model, with the teacher as one of the vertices and the science classroom context as overarching variable.

Mayer (1999) proposed a theory of knowledge construction called the *SOI (Select, Organise and Integrate) model*. The SOI model of leaning fosters three cognitive processes in knowledge construction: selecting, organizing and integrating information. A visual representation of the SOI model is given in Figure 2.3.

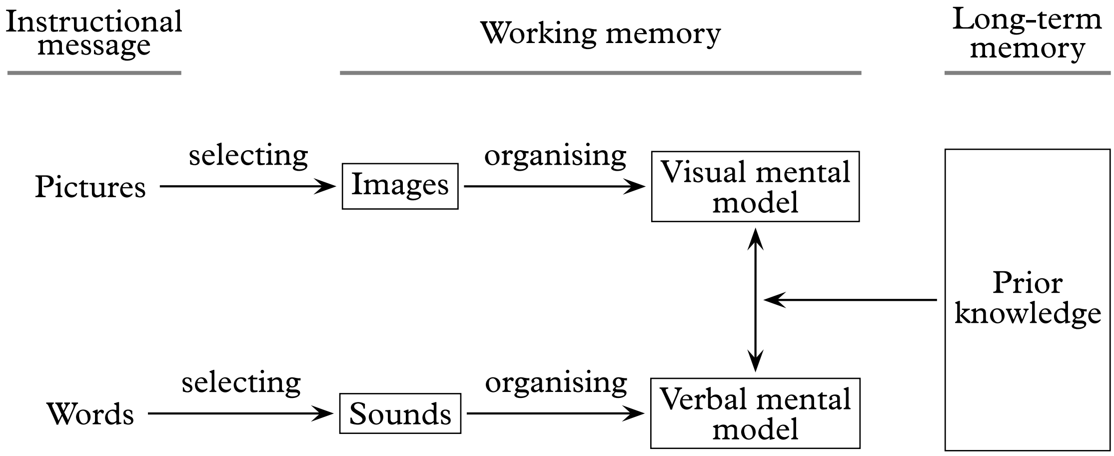


Figure 2.3: SOI model of Constructivist learning from words and pictures (Mayer, 1999, 149)

Text structures do not *per se* influence the results of comprehension and learning processes, but only in so far as certain cognitive activities of learners are initiated by them (Van Hout-Wolters and Schnotz, 1992). The process of integration of presented information and prior knowledge can be fostered by using textbook strategies like mind-maps, worked-out examples and elaborative questions.

One of the benefits of learning from text is that it prepares learners for lifelong learning. Adult learning from magazines, encyclopædia and other media constitutes an important component of informal lifelong learning. In this area text-based learning has increased in importance in our society with the diversifying of information channels (i.e. Internet and ICT) which makes it increasingly difficult to filter relevant from redundant information (Broer et al., 2002).

From the learners' interaction with the textbook we now turn our attention to the teachers' interaction with the textbook.

2.3.2 Teaching with textbooks

Science education has always been assosiated with the use of textbooks written especially for students (Klassen, 2006). Teachers have to face many challenges in their work (Anderson and Helms, 2001) and deserve all of the support they can get (Anderson, 1993, 157). This is especially true in science education (Muwanga-Zake, n.d.). Although all subjects require teacher input, Muwanga-Zake (n.d.) contends that teaching science requires more input than any other subject. He explains that science teachers have to prepare for practical work and care for the laboratory and equipment, but usually they have to teach the same number of classes as teachers in other subjects. Textbooks can help to reduce their burden.

Henson (2004) summarises three ways that teachers use textbooks: some try to avoid using it at all; the second group centres their teaching on it and supports it with other books, journals and newspapers and the third group of teachers design their own curriculum and just use the textbook along with other media as supplementary material.

Unfortunately, some unsuccessful teachers tend to overuse textbooks or use it inappropriately (Klassen, 2006; Ewing, 2006). Some even use textbooks to prevent chaos in the classroom by keeping the students occupied (Brändström, 2005). Unsuccessful teachers often use textbooks in a way that only results in a passive learning style. This overuse or inappropriate use of textbooks has created a "deep-seated anti-textbook ethos" in many educational communities (Issitt, 2004, 683). In reaction to such inappropriate use and overuse of textbooks, some members of the science education community and many teacher training programmes, like the introductory OBE training in South Africa, have aimed to make teachers independent of textbooks (Taylor, 2008a; Ball and Feiman-Nemser, 1988). However, when used wisely, the available textbooks have the potential to serve as a tool that enables the teacher to "do her best work with students" (Russell, 1997, 247). Furthermore, most teachers can not write their own textbooks or curriculum material (Department of Education of South Africa, 2001a) because (in the words of Russell (1997, 247)) it

> . . . is a job that requires people and resources; it requires a skilled team of . . . educators spending many thousands of hours writing, thinking, working in classrooms, and listening to students and teachers. We do not sell teachers short by recognizing that they cannot do this job.

Studies worldwide show that textbooks are still widely used (Lemmer et al., 2008; Arriassecq and Greca, 2007; Stern and Roseman, 2001; Sitte, 1999) and is considered as essential to effective teaching (Taylor, 2008a; Klassen, 2006). Table 2.1 lists a few examples of studies that measured the use of textbooks.

In South Africa the introduction of OBE created the impression that teachers have to be designers of learning programmes and their own learning support material (Department of Education of South Africa, 2001a). This aversion to textbooks has been identified by Taylor (2008a) and the Task Team for the Review of the Implementation of the National Curriculum Statement (2009) as one of the reasons why the education transformation in South Africa was unsuccessful. Taylor (2008a, 3) even contends that

Country	Research on textbook use
USA	▷ 96% of grade 9–12 science classes use published textbooks ▷ 59% of a national sampling of science teachers indicated that textbooks had a major influence on their teaching *(NSTA, 2003)*
France	Teachers use textbooks almost all the time (Pepin and Haggerty, 2003)
Germany	▷ 70% of teachers used mostly textbooks ▷ 20% of teachers used textbooks often ▷ 8% of teachers seldomly used textbook ▷ 2% of teachers never used textbooks *(Sitte, 1999)*
Austria	▷ Textbooks are the most used teaching aid. Textbooks were used in 87,4% of the cases where teaching aids were used *(Sitte, 1999)*.
Spain	▷ 92% of teachers use textbooks as basic reference for planning *(Huber and Moore, 2001)*

Table 2.1: Examples of studies that measured the use of textbooks

> . . . the disparaging attitude of most South Africans towards books and the inadequate quantities of reading and writing undertaken in the majority of schools, is nothing short of educational suicide on a national scale.

The view of the role of the teacher has now radically changed to that of facilitators spending their time on actively teaching using textbooks designed by experts in their field of study (Motshekga, 2009; Department of Education of South Africa, 2009). The South African Department of Education now sees the role of textbooks in the improvement of education (from 2010) as effective tools to ensure consistency, coverage, appropriate pacing and better quality in terms of instruction and content (Motshekga, 2009).

Different countries have different approaches to textbook provision. In France it is regarded as vital to provide students with textbooks, while in Germany pupils are expected to buy their own textbooks. In the Netherlands the freedom of choice of textbooks and learning material is guaranteed by the constitution, while twenty American states choose or "adopt" textbooks for statewide use (Tyson, 1997). From 2010 the South African Department of Education aims to provide every learner with a textbook in every subject area (Department of Education of South

Africa, 2009).

The way teachers use textbooks influences the way learners interact with their textbooks and even the image of science epistemology developed by learners (Ulerick, n.d.; Ninnes, 2002; Alexander and Kulikuwich, 1994; Eltinge and Roberts, 1993). Alexander and Kulikuwich (1994) found that students had a very good idea what content in the textbooks their teachers valued and this proved to be problematic when teachers follow the textbook rigidly and often refer to the textbook to check facts. It fosters the idea that science is a set of facts that is contained in the textbook.

2.3.2.1 Authority of textbook

Textbook authority or control is a complex concept and can have both positive and negative implications (Johnsen, 1993). What authority the textbook exerts in a classroom depends on how the learners and teachers view textbooks and respond to them (Kesidou and Roseman, 2002). Apple and Christian-Smith (1991) identify three ways in which people can potentially respond to texts:

- dominated (accept the message at face value),
- negotiated (reader may dispute a particular claim, but accepts the overall interpretation of a text) and
- oppositional (reject the dominant tendencies and interpretations in texts).

Most inexperienced and under-qualified teachers are dominated by textbooks in science education. They consider the textbook as the "correct" and sometimes even as the only source of knowledge and follow it rigidly (Ogan-Bekiroglu, 2007; Arriassecq and Greca, 2007; Tarr, Chávez, Reys and Reys, 2006; Pepin and Haggerty, 2003; Kesidou and Roseman, 2002; Wong, 1991; Otero and Campanario, 1990). An "unsuccessful" textbook in the hands of such a teacher can have a negative influence on learning (Ewing, 2006; Johnsen, 1993). According to Chester Finn (Whitman, 2004, I):

> To rely for one's course content and lesson plans on inferior instructional materials is like boxing with an arm tied behind one's back: success is apt to prove elusive.

In contrast, "successful" learning material can play a valuable role in inexperienced and under-qualified teachers' professional development (Newton and Newton, 2006; Litz, 2001; Iszak and Sherin, 2003; Huber and Moore, 2001; Singer and Tuomi, 2003). A good textbook has the potential to offer substantial and significant support to teachers. In the case of un(der)-qualified teachers, textbooks and exemplary materials are often the only sources of guidance and support readily available (Ogan-Bekiroglu, 2007; Newton and Newton, 2006; McKenney, 2001).

The application of educational reform policies can be impeded by teacher beliefs and knowledge (Mellado, Bermejo, Blanco and Ruiz, 2007; Aldridge et al., 2006; Lang, 2001; Anderson and Helms, 2001; Davis, 2003a) and research has shown that teachers are resistant to change in ideology or pedagogy (Brandt, 2005). One of the acknowledged strategies for attempting to change teacher beliefs and influence classroom instruction is the design and spread of curriculum material (Tarr et al., 2006; Fishman and Davis, 2006; Chambliss and Calfee, 1998). In South Africa this method will be used to improve education in the GET band from January 2010 (Pretorius, 2009). Changing textbooks is considered one of the most economical and efficient ways to improve the instruction in classrooms (Lubben et al., 2003; Chambliss and Calfee, 1998). This strategy assumes that the textbook does have at least a degree of authority or control in teaching. The assumption is justified by research that confirmed that well-designed curriculum material can successfully support (and influence) teachers when planning and implementing unfamiliar teaching strategies (Davis, 2009; Taylor, 2008b; Iszak and Sherin, 2003; Davis, 2003a; Watts and Simon, 1999; Smith et al., 1993). In this way both the education providers and the teachers benefit from textbooks during educational reform processes, because

> textbooks can support teachers through potentially disturbing and threatening change processes, demonstrate new and/or untried methodologies, introduce change gradually, and create scaffolding upon which teachers can build a more creative methodology of their own. (Litz, 2001, 6)

There can be a number of reasons why a teacher opts to follow a textbook rigidly. An unqualified or inexperienced teacher can use a textbook to survive, because it can provide guidance, support and security (Ansary and Babaii, 2002), while a well-qualified, experienced teacher can choose to make extensive use of a textbook because it is consistent with his or her own and researched-based views on education (Tarr et al., 2006; Ansary and Babaii, 2002). A busy teacher may use it to save time (Pepin and Haggerty, 2003). In this context the best textbook is one that "will be most useful, requiring the least amount of modification and change" (Stein et al., 2001, 18).

Whatever the teacher's reasons are for using a textbook, Parkinson (2002, 48) suggests that all teachers should have a negotiated response to the textbook. They should “take ownership” of the contents of the textbooks, judging its value in various parts of a lesson, deciding how they are going to use the activities in the textbooks, supplementing it or even discarding portions of the textbook according to his or her learners’ needs (Lubben et al., 2003; Pepin and Haggerty, 2003; Sherin, 2002; Tyson, 1997). In this way the textbook acts as an extension of the teacher and can be used by the teacher to extend the learning opportunities (Schramm, 1977).

2.3.2.2 Textbooks in OBE

In the process of curriculum change in South Africa some advisors introduced ideas that appeared to suggest that textbooks were obsolete in models where learners’ contexts and own experience were used as basis for education (Taylor, 2008a; Land, 2002). Teachers were encouraged to design their own learning support materials (Department of Education of South Africa, 2001a). Taylor (2008b, 24) contends that this aversion to textbooks is “one of the most damaging aspects of post-apartheid education.” It made enormous demands on teachers and left children with “. . . fragments of the curriculum, presented through standalone worksheets or isolated, short exercises written on the board” (Taylor, 2008b, 2).

The poor performance of children in South Africa in both international and local assessments forced the Minister of Education to appoint a task team in July 2009 to investigate the implementation of the National Curriculum Statements (Task Team for the Review of the Implementation of the National Curriculum Statement, 2009; Motshekga, 2009). In the light of their findings Minister Motshekga declaired OBE in South Africa “dead,” although it was not replaced by a new policy (Motshekga, 2009). From 2010 changes to the system will be made over a five year period to improve learner performance (Department of Education of South Africa, 2009). One of these changes will be that the importance of textbooks will be communicated to teachers and learners and that every learner will be supplied with a textbook in every subject (Motshekga, 2009; Department of Education of South Africa, 2009).

The textbook is the *written curriculum* that links the *intended curriculum* (articulated in the National Curriculum Statements) to the *enacted curriculum* or *implemented curriculum* (that is actually experienced in the classroom) (Amaral and Garrison, 2007; Tarr et al., 2006; Törnroos, 2005; Mesa, 2004). Consequently, textbooks and other learning materials are essential components of OBE, and this includes textbooks (Malcolm and Alant, 2004; Asmal, 1999; Potenza and Monyokolo, 1999). In their development of a theoretical framework for curriculum implementation, Rogan and Grayson ((2003)) even consider the availability of textbooks as part of the physical resources that influence a school's capacity to support innovation. This is also evident in work of Potenza and Monyokolo (1999). They claim that the success of OBE is determined by the presence of three key pillars (see Figure 2.4) or essential components of the curriculum, namely:

- curriculum development
- learning materials (which include textbooks) and
- teacher development.

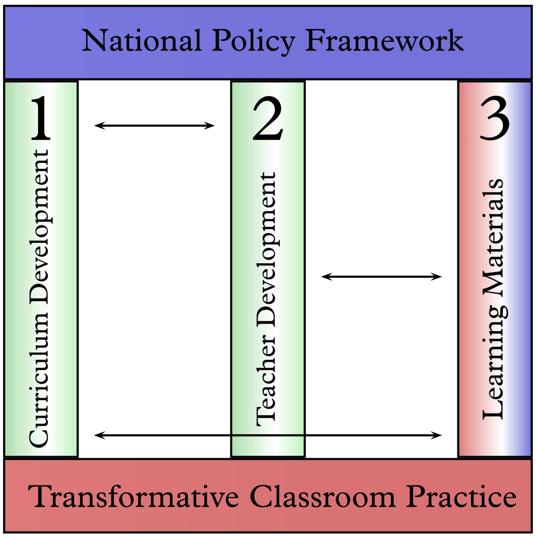


Figure 2.4: The three key pillars of the curriculum. (Potenze & Monyokolo, 1999:232)

The role of a single textbook to present concepts and content in an "organised, systematic and synthesised fashion" (Task Team for the Review of the Implementation of the National Curriculum Statement, 2009, 6) is important. Previous OBE strategies often used a multitude of texts to provide learners with the opportunity to learn to extract and compare information from different sources (Land, 2002). This strategy has a role to play in education, but it can not supply a systematic structure and ensure coverage of curriculum content.

Although the importance of textbooks is a major focus of improvements in education in South Africa from 2010 (Motshekga, 2009; Task Team for the Review of the Implementation of the National Curriculum Statement, 2009), government's budget allocations for Learning Support Material (LSM) confirmed that they have always considered it as important (Jacobs, 2001). Table 2.2 shows an increase for the 1999 to 2003 financial years, even if inflation is taken into consideration (Department of Education of South Africa, 2003b). The 2009 mid-term

Province	Year	Year	Year	Year	Year
	1998/1999 R million	1999/2000 R million	2000/2001 R million	2001/2002 R million	2002/2003 R million
Eastern Cape	42,1	120,0	155,5	180,0	272,3
Free State	26,9	72,1	75,7	77,8	84,1
Gauteng	52,3	95,4	153,4	176,0	240,0
KwaZulu-Natal	72,6	156,0	103,3	193,9	199,3
Limpopo	51,5	148,3	244,2	269,4	330,0
Mpumulanga	51,0	63,9	49,5	68,5	212,0
North West	45,9	65,2	53,0	53,6	50,0
Northern Cape	6,6	27,6	10,7	28,9	28,9
Western Cape	54,9	57,0	70,0	96,8	1490,0
TOTAL	392,6	794,6	920,2	1162,90	1565,6

Table 2.2: Learning Support Material (LSM) (textbooks and stationery) budget allocations (Department of Education of South Africa, 2003b)

budget allocation of more than R500 million for textbooks confirms the renewed belief in textbooks' contribution to effective education.

Teams of experts and teachers who are thoroughly familiar with the outcomes-based approach to education and a specific learning area will be able to design learning programmes with a wide variety of learning activities and the necessary print-based LSMs (Motshekga, 2009). With careful selection it would be possible for the teacher to choose appropriate textbooks for his or her learners in their specific Learning and Resource Context (LRC). Such textbooks could meet the majority of the curriculum statements and outcomes without placing too much of a burden on the teacher.

The *quality of a textbook* is a complex issue and it is not immediately apparent what a good textbook must look like. In Chapter 1 a good textbook was defined as one that has the potential to support the learner and teacher in attaining the desired science learning goals. A good textbook is, therefore, a textbook that incorporates *characteristics* that enable it to support the learners and teachers (Davis, 2003b). Consequently, the presence of these characteristics in a textbook will indicate its quality or it's potential to support the learner and teacher.

The complexity of the problem of identifying the characteristics that indicate textbook quality, demands a systematic approach to the problem. The overall purpose of supporting the learner and teacher can be broken down to a number of separately identifiable *functions* that contribute to fulfilling the overall purpose. Research on textbook characteristics is usually limited to one of these functions. This makes it easier to identify the characteristics that facilitate the separate functions, than trying to work with this complex issue as a whole. Therefore, in preparation for the identification of characteristic of good textbooks, the discussion in the next section centres around the functions that science education textbooks must fulfil in order to support the learners and teachers in the learning process.

2.3.3 Functions of textbooks

The purpose of science education textbooks is to support teachers and learners in the learning process (Litz, 2001; Garcia-Barros, Martinez-Losada, Vega and Mondelo, 2001; Franssen, 1989a). Kesidou and Roseman (2002, 523) assert that textbooks

> ... provide a coherent science program for students based on the best thinking available in the field, and material that supports teachers in making more thoughtful and informed decisions about their own students science learning.

The kind of support needed by teachers is different from the support required by learners. For example, teachers need to know how to present new concepts, while the learners, in turn, need to understand and apply these concepts. In this section the *support* required by the teachers and learners will be considered in more detail and broken down to a set of functions that textbooks must fulfil in order to support the teachers and learners.

Van Schalkwyk (1993) differentiates between the following general functions of media in teaching and learning:

- extends experience
- increases quality of teaching
- sharpens interest and observation
- nurtures skills
- enables individualisation and differentiation
- serves as diagnostic tool in the learning process
- fosters media literacy
- directs learning activities
- stimulates double unlocking (content unlocked for learner, learner unlocked or stimulated to participate and learn)

The functions textbooks can fulfil to support *teachers* are twofold (Reys and Reys, 2006; Russell, 1997):

- planning and teaching (day-to-day)
- professional development (long term)

In the context of the educational reform process textbooks must support teachers in the translation and execution of the curriculum (Taylor, 2008a). Van den Akker (McKenney, 2001) identified three aspects of that support:

- clearer understanding of how to translate curriculum ideas into classroom practice
- concrete foothold for execution of lessons that resemble the original intentions of the designers
- stimulation of reflection on one's own role with the eventual possibility of adjusting one's own attitude toward the innovation

Teachers' editions of textbooks can target the teachers' needs. It can provide explanations of curriculum requirements, as well as the content or subject knowledge and pedagogical knowledge (Singer and Tuomi, 2003), discussed in Section 2.3.2. STAMP 2000+ is an example of teacher training materials that facilitate in-service training and upgrading for teachers. It was designed by the Botswana College of Distance and Open Learning for participating sub-Sahara countries (Daniel and Menon, 2005).

Mikk (2000, 17–19) identifies the following functions of textbooks in its support of students in their learning:

- motivate students to learn
- represent information (transform and systemise)
- guide students to acquire knowledge
- guide students to acquire learning strategies
- aid self-assessment
- differentiate
- facilitate value education

From the preceding views the following list was compiled to summarise the functions of the textbook.

For the *students* textbooks must:

- motivate students to learn
- represent subject knowledge
    - transform, systemise and present
    - provide an appropriate view of the nature of science
    - provide visual representation of phenomena/apparatus
- guide student learning

    - identify prior knowledge
    - provide explanations and activities to facilitate knowledge acquisition and conceptual change
    - provide exercise and application opportunities
    - facilitate self-assessment
- guide students to acquire learning strategies
    - stimulate meta-cognition
    - scaffold learning strategies

For *teachers* textbooks must

1. aid teachers' planning

    - describe the relevant content or subject knowledge
    - provide pedagogical knowledge relevant to the content that is taught

2. aid professional development of teachers

    - develop their content or subject knowledge and nature of science (NoS) view
    - develop their pedagogical knowledge, beliefs and attitudes

In *general* it is also the function of textbooks to

1. co-ordinate with other educational aids

2. differentiate

3. facilitate value education

2.3.4 The textbook characteristics that contribute to fulfilling the functions of the textbook

The functions that textbooks must fulfil to support the learner and teacher was set out in the previous section. To fulfil each of the functions the textbook must exhibit certain characteristics. In the remainder of this section each function of the textbook will be considered separately, in order to ascertain the characteristics that must be present to enable the textbooks to fulfil that function. As mentioned in the previous section, the presence of these characteristics indicates that the textbook will probably be effective in supporting the learner and teacher in science education. Consequently, the presence of the characteristics acts as indicator of the textbooks' effectiveness and quality.

2.3.4.1 Characteristics that motivate students to learn

The motivation of students is one of the most important functions of textbooks (Mikk, 2000). Learners' motivation to learn was discussed in Section 2.2.4.2. A number of strategies and principles have been devised for the design of textbooks and other educational material to stimulate learner motivation to learn (Beck and McKeown, 2001; Keller and Burkman, 1993). Many of the strategies used in the past (and still used in some textbooks) have proved ineffective to motivate learners. For example, some visual strategies to improve motivation turned out to be counterproductive in the learning process. Daniels and Zemelman (2003) found that textbooks that try to compete with video visuals end in visual chaos. The success of the Harry Potter books prove that a well-written book can still hold children's attention, without existing visual stimuli (Ravitch, 2003).

Reading comprehension requires interaction with the text (discussed in Section 2.3.1). Therefore, researched-based strategies to motivate learners were designed on the premise that strategies that invite students to interact with text encourage more thoughtful reading and more connected learners, who want to learn more (Fordham, Wellman and Sandmann, 2002). These strategies should not involve adding interesting but unnecessary extra information. Dutch (2005, 35) contends that the

> . . . expectation that learning should be painless and entertaining clutters textbooks with unnecessary and distracting features that subsidise poor study habits, probably reduce the effectiveness of the text, drive up costs, and displace more worthwhile material.

The inclusion of unnecessary detail, jokes, et cetera related to the topic may actually distract some readers' attention from relevant content (Kesidou and Roseman, 2002). The term "seductive details" has been coined to describe these distracting details (Mayer, 2003; Boostrom, 2001; Shraw, 1998; Alexander, 1998; Harp and Mayer, 1997; Alexander and Kulikuwich, 1994). According to Budiansky (2001, 1) this strategy resulted in American science textbooks that have become

> . . . larger and flashier, chock full of colourful photographs, diagrams, 'activities', 'minilabs', sidebars about minorities in science, science in history and literature and art, and current issues such as the use of hormones in dairy cattle . . . The only thing the books utterly fail to do . . . is teach science.

In reaction to this unsuccessful strategy, many researchers have demanded textbooks without the paraphernalia that publishers claim will motivate learners. Instead of adding unnecessary information, the required content should rather be presented in such a way that learners are motivated to engage in the learning process. Crawford (2005) even contends that only intrinsic satisfaction, rooted in the learning activity, can sustain the effort required to learn any subject well.

Fortunately there are a number of examples of strategies that have proved to be successful in stimulating student motivation, active engagement in text processing, and learning science from text. Examples of such strategies are:

- portraying science and scientific inquiry as a human endeavour (Crawford, 2005)
- tying material (where possible) to everyday phenomena that children can experience (Dake, 2007; Budiansky, 2001).
- the use of variation and curiosity and stimulating need or relevance (Keller and Burkman, 1993);
- including characteristics that mark text as interpersonal communication, such as using verbs that represent more concrete actions (activity), including conversational tone (orality) and highlighting relationships and connections between the reader and the text (connectivity) and
- inserting a limited amount of tested literary devices into text (e.g. analogies).

2.3.4.2 Characteristics that ensure effective representation of subject knowledge

One of the first functions of textbooks was to make specialised knowledge available to learners (Villaverde, 2003). Today many other sources are available to provide this information, but gathering the relevant knowledge together in a single book is a practical approach that provides learners with "access to the whole curriculum in an integrated form" (Taylor, 2008a, 2). Textbooks have to represent information accurately to be effective (Roseman, Kesidou, Stern and Caldwell, 1999). Users usually take the accuracy of the content that is represented for granted, but research has shown that misleading or inaccurate statements occur in textbooks and result in incomplete understanding by learners (Dall'Alba, Walsh, Bowden, Martin, Masters, Ramsden and Stephanou, 1993).

Textbook authors have to consider what to represent; in what order to present it and how to represent it (Boostrom, 2001; Mikk, 2000). Writing a textbook provides the textbook author with the opportunity to select the content that is specified by the curriculum and transform it into a coherent tool that can facilitate the attainment of the learning outcomes (Mikk, 2000). Although some users may find the prearranged sequence and structure of textbooks annoying (Ansary and Babaii, 2002), the cohesion and structure of the text decreases the processing demand of the text on the reader (Alexander and Kulikuwich, 1994).

Textbook authors have to present the content within the framework of a specific view of science (as discussed in Section 2.2.1) (Johnsen, 1993; Apple and Christian-Smith, 1991). A textbook writer's view of science and science education determines the way in which the content is presented (Unsworth, 2001; Knain, 2001; Kearsey and Turner, 1999; De Berg and Greive, 1999). This may be in the form of information, an argument or an explanation (Chambliss, 2001). In view of current views of science accepted by most education departments, the textbook must present science not as a constant, unchanging body of knowledge, but as a dynamic, growing, and continuous quest for knowledge (Eltinge and Roberts, 1993). This is not true of science textbooks that tend to illustrate scientific knowledge, in the sense of showing, summarising and defining, rather than providing evidence and argument for the conclusions that are reached and presented (Norris and Phillips, n.d.). Clough (2006, 464) describes this when he argues that

> (s)cience textbooks, common cookbook laboratory activities, and most audiovisual materials downplay human influences in research, sanitise the processes that eventually result in knowledge, and portray science as a rhetoric of conclusions.

The way textbook authors present science is, therefore, of paramount importance when the textbook aims to facilitate understanding and other higher-order learning outcomes. Textbook content can be used to provide answers or stimulate students to explore questions and issues. Students can be fed information (Issitt, 2004) or be forced to make decisions or engage in other higher-order activities (Herlihy, 1992a).

2.3.4.3 Characteristics that guide student learning

There is a very important difference between a textbook and a reference book, like an encyclopedia (The Centre for Curriculum Materials in Science, 2006; Daniels and Zemelman, 2003). Although the knowledge in reference books is well presented, it is not adequate for a textbook that have to guide the students in constructing the concepts for themselves. This is why Hurd (Woodward and Elliott, 1990) called American biology textbooks "the most beautifully illustrated dictionaries." It failed to guide the learners to learn the content. Therefore, the most suitable presentation of the content in textbooks should be used, as determined by our current understanding of learning and of the cognitive processes involved with understanding textbooks (Klassen, 2006; Sulaiman and Dwyer, 2002; Sorrels and Britton, 1998; Krüger, 1983).

To enable constructive learning, textbooks must provide opportunities for activities that encourage cognitive engagement and create the conditions for conceptual change to occur (see Section 2.2.3). For example, learners with inadequate prior knowledge can be explicitly supported by providing access to information or "prior learning" activities that can help learners gain the knowledge that is lacking or change misconceptions. Instructional strategies for conceptual change in textbooks can include cognitive conflict strategies, like refutational text (Guzzetti, 2000) or prompts that guide learners to contrast common misconceptions with the scientific alternatives (Mikkilä-Erdmann, 2002; Roseman et al., 2001). Refutational text states common misconceptions and directly refutes it while providing the "scientifically acceptable idea" (Guzzetti, 2000).

Other strategies that can be employed in conceptual change teaching include the use of vivid examples (Budiansky, 2001) and provision of opportunities (questions or discussions) for the interpretation of results of activities in terms of the scientific concepts to be learned (Roseman et al., 2001). Textbooks can facilitate learner construction of new concepts, by illustrating it as intelligible, plausible and possibly fruitful, as required by Posner, Strike, Hewson, and Gertzog's conceptual change theory (Posner et al., 1982, 222). For example, the use of explanatory principles (plausibility, parsimony, generalisability, fruitfulness) in textbooks can contribute to conceptual change (De Berg and Greive, 1999).

The level of difficulty of textbook activities is very important. According to Anderson (1993) optimal development of knowledge requires activities of moderate difficulty: aimed at students' zone of proximal development (Anderson, 1993). This zone represents what a learner can accomplish with some assistance. It is the ideal level of task difficulty, conducting the learner to higher levels of understanding. This requires conceptual support, often called *scaffolding*, to help students to purposefully reason through a problem (Davis and Linn, 2000; Anderson, 1993). Textbooks can provide *scaffolding* or instructional support that help students learn but can be withdrawn when the learners have reached the outcomes (Hewson, 2004; Davis and Linn, 2000), like instructional "training wheels" (Wilen et al., 2004). For example, the text can provide questions that activate relevant prior knowledge and direct attention to relevant information.

Scientific writing is renowned for nominalisation and other characteristics that complicate text comprehension (Hsu and Yang, 2007) and textbooks should guide learners in their interpretation of these phenomena (Cicourel, 1985). Science texts, for example, are "bilingual" (mathematics and text) (Alexander and Kulikuwich, 1994) or rather multi-modal (verbal, mathematical and visual language combine to create new meanings) (Dimopoulos et al., 2005; Osborne, 2002; Ametller and Pinto, 2002). Textbooks should guide learners in the interpretation of the text as a whole. For example, new or difficult mathematical manipulations must be highlighted in the text or alternatively in an appendix.

It is not only the writing style that influences text comprehension. Text structures can influence the comprehension and learning processes by initiating certain cognitive activities of learners (Van Hout-Wolters and Schnotz, 1992). Consider the following examples.

- The *selection* of information by the learner can be facilitated by the use of text devices such as headings, italics, bullets, margin notes and white space (Sulaiman and Dwyer, 2002; Weiten, Deguara, Rehmke and Sewell, 1999; Mayer, 1999).
- Learners can be assisted in *organising* information by outlines, signaling headings, pointer words, structured illustrations, and coherent text structures (Sulaiman and Dwyer, 2002).
- The process of *integration* of presented information and prior knowledge can be fostered by using textbook strategies like advance organisers or mindmaps and worked-out examples (Dake, 2007).

Illustrations can also contribute to learning science, but the efficacy of illustrations for learning is a perennial question in many content areas and illustrations in science textbooks have been targeted by research with a range of perspectives (Stylianidou, Ormerod and Ogborn, 2002; Peacock and Weedon, 2002; Ametller and Pinto, 2002; Peña, 2001). In their analysis of USA science textbooks the AAAS found that lavishly illustrated textbooks rarely facilitate learning because the illustrations are "too abstract, needlessly complicated, or inadequately explained" (Roseman et al., 2001, 56).

Iding, Klemm, Crosby and Speitel (2002) proposed a taxonomy of illustrations based on the cognitive processes or activities that learners must engage in to understand the material. The illustrations are classified as knowledge acquisition, knowledge application or knowledge creation. In 2003 Mayer (2003) reviewed the research on the use of illustrations and developed research-based principles for the use of illustrations in book-based and computer-based environments.

Laboratory activities in textbooks can guide learners to reach learning outcomes, provided that it is designed to encourage cognitive involvement and independent thinking and inquiry (Amaral and Garrison, 2007; Hand and Keys, 1999). Section 2.2.4.3 discussed how this requires a move away from the traditional "cookbook-style" of laboratory activities, but even in this context the textbook can still perform a valuable service, since learners may need the structure and support of written "scaffolding" or "job performance aid," even with inquiry activities (Huber and Moore, 2001). Textbooks can play a valuable role in guiding learners to interpret and consolidate their learning experiences during or after laboratory activities (Appelbaum and Clark, 2001; Bancroft, 2002).

2.3.4.4 Characteristics that guide students to acquire learning strategies

With lifelong learning as a goal for education it is especially important to teach children not only content, but learning skills that will enable them to keep on learning, even when their formal education is completed. Besides guiding students to acquire knowledge, textbooks must, therefore, also guide them in developing learning strategies and skills (Mikk, 2000). Textbooks can guide learners to develop learning skills by including activities that help learners make summaries, organise their work, take notes and memorise information (Hudson, 1994). The most important learning strategy learners have to develop is the habit of reflection and self-assessment (Tamir, 1999). The following textbook characteristics can guide learners to develop the skill of meta-cognition (being aware and in charge of their own thoughts):

- highlighting learning goals
- providing opportunities for learners to identify their baseline assessment (discussed in Section 2.2.4.6)
- explaining the functionality of various learning activities
- providing questions to be answered prior, during and after passages
- providing prompts that scaffold self-reflection
- providing self-assessment opportunities

According to Hewitt (Stinner, 1992) traditional textbooks seduce learners to gauge their success by their ability to answer the questions found at the back of their textbooks. Students quickly learn that they only have to attend to words in boldface and memorise definitions to answer the end-of-chapter questions (Anderson, 1993).

Traditional textbooks are also criticised for the level of its questions. Most questions assess only knowledge of direct definitions and problems solved by straightforward substitution of given amounts into equations. These questions tend to be "low in cognitive level, inviting learners to follow a search-and-find learning strategy" (Ulerick, n.d.). If textbooks aim to encourage learners to gain understanding, the textbook assessment questions must also address understanding (Singer and Tuomi, 2003; Stern and Ahlgren, 2002; Dall'Alba et al., 1993).

Some further basic information gathering skills, like using an index, can enhance a learner's access to information in any book (Mann, Lu and Grzybowski, 1999). Textbooks should contain these devices and guide learners to develop the skills to utilise such devises. This can enhance their ability to collect information, which is stated as one of the critical outcomes in South Africa's OBE (Department of Education of South Africa, 2006).

2.3.4.5 Characteristics that facilitate coordination with other educational aids

Textbooks should refer to other educational aids. For example, textbooks can provide lists of books that cover the same content and even refer to films or story books in which the phenomena or principles that are discussed in the text are observed or applied. It may even refer to Internet sites where relevant information can be found, but it is necessary to bear in mind that Internet sites change rapidly and sites may cease to exist while the textbook mentioning it is still in use. In America a partnership between progressive U.S. textbook publishers and the National Science Teachers Association (NSTA) has led to SciLinks, a free service that offers teacher-approved internet resources tied to specific points in the textbook. It is accessed by logging on to the SciLinks site and entering the SciLinks number from the margin of the textbook (SciLinks, n.d.).

The most obvious educational aids mentioned in science education textbooks are laboratory equipment, but textbooks should also frequently refer to "equipment" available in the average home of the learners. Many textbooks include "kitchen-chemistry" or experiments that learners can do at home or in the classroom.

2.3.4.6 Characteristics that aid differentiation

According to Brandt (2005) differentiation in education is the creation of different learning situations for different students according to their needs (of matching teaching to learning needs). According to the Department of Education good textbooks "are written in a manner that allows adaptation of content for learners who experience barriers to learning" (Department of Education of South Africa, 2009, 6).

When differentiation is considered, one focus is on learners that learn the required content quickly and need *extra stimulation* to realise their full potential. The textbook can be an excellent aid in this regard by providing multiple exercises and application opportunities for learners of different levels of competency (Brändström, 2005). But attending to *prior knowledge* can also be an important differentiating strategy, because it addresses differences between learners. One of the most frequent complaints about textbooks is about the way in which prior topic knowledge is taken for granted (Ulerick, n.d.; Mikkilä-Erdmann, 2002). Learners' prior knowledge (domain, topic and alternative knowledge (Alexander and Kulikuwich, 1994)) must be identified and attended to. The prior knowledge of a specific learner can not be taken as point of departure in a textbook. The design of a textbook must provide support for learners with diverse prior knowledge systems. It must attend to all the alternative conceptions that occur frequently. Textbooks can provide scaffolding for learners with inadequate or inappropriate prior knowledge. In this regard research done by Sulaiman and Dwyer (2002) confirms that added cues in text enables low prior knowledge students to reduce the differences attributed to prior knowledge. Because it is difficult to attend to a variety of prior knowledge levels in one textbook, teachers in England, for example, use different textbooks for high, intermediate and low 'ability' learners (Brandt, 2005; Pepin and Haggerty, 2003).

2.3.4.7 Characteristics that facilitate value education

Textbooks can contribute to learners' acceptance of values. The most obvious values many stakeholders fight for in textbooks are social values. For example, the new South African Revised National Curriculum Statements are "overtly promoting the values of democracy, social justice, equity and equality, including non-racism and non-sexism" (McKinney, 2005, 1). The Oklahoma State Textbook Committee (n.d.b, 3) spelled out twelve criteria that must be considered when adopting materials. The first criterion refers to content, and the remainder are all devoted to values:

1. Align with recognised curriculum standards;
2. Are objective in content and impartial in interpretations, and which do not encourage or condone civil disorder, social strife or disregard for the law;

3. Do not degrade, and where appropriate, teach high moral standards including: (A) honesty (B) respect for parents, teachers, and those properly in authority; (C) the importance of the work ethic in achieving personal goals; (D) the existence of absolute values of right and wrong;

4. Emphasise the importance of the family as the core of American society and do not degrade traditional roles of men and women, boys and girls;

5. Include the principles of the free enterprise system;

6. Are designed to foster the intellectual development of the child by providing instruction in reading, writing and arithmetic, through centuries of academic endeavour, including an awareness of the religious and classical culture of the western world and its significance to the preservation of the liberties of the American people;

7. Present balanced and factual treatments to controversial, political and social movements without biased editorial judgements;

8. Do not promote illegal lifestyles or sexual behaviour, sadistic or degrading behaviour;

9. Do not include blatantly offensive language or illustrations;

10. Do not include violence for reasons of excitement, sensationalism or as an excuse for relevance. Violence, if it appears in textbook content, shall be treated in context of cause and consequence;

11. Treat the subject of historical origins of humankind in an objective and unbiased manner; and

12. Do not invade the privacy of the pupils or pupils' parents.

Besides moral values learners should learn the values associated with the nature of science and its relationships to technology, ociety and the environment. This was discussed in Section 2.2.1. Although direct instruction can present scientific values, the values will develop in a more enduring and transferable form if they are presented in all elements of science teaching (Council of State Science Supervisors, n.d.). Textbooks have to include activities that provide opportunities for learners to learn to identify and critically evaluate scientific knowledge claims and the impact of this knowledge on the quality of socio-economic, environmental and human development.

2.3.4.8 Characteristics that aid teacher planning and teaching

Textbooks can provide material that supports all teachers in making better, more thoughtful, more informed decisions about their students' learning (The Centre for Curriculum Materials in Science, 2006; Russell, 1997). All teachers can benefit from a map or indicators that give directions along the path they are travelling with their students (even if it just serves to confirm their own knowledge) and

> . . . a textbook or other instructional material can provide that road map. Individual teachers can, of course, take instructional detours and stop to enjoy the view of particular aspects of that curriculum along the way. But the curriculum contained in a textbook or other core instructional material provides teachers, administrators, and students a basic framework for instruction (Montgomery, 2006, 1).

Teachers are supported by textbooks that

- help them translate specific curriculum ideas into classroom practice (Davis, 2003b; Davis, 2003a; McKenney, 2001; Mozambique Ministry of Education, 2002; Watts and Simon, 1999; Chambliss and Calfee, 1998)
- state clear learning objectives (Watts and Simon, 1999)
- reflect appropriate pacing and weighing of content (Task Team for the Review of the Implementation of the National Curriculum Statement, 2009)
- provide key science ideas (The Centre for Curriculum Materials in Science, 2006; Malcolm and Alant, 2004; Singer and Tuomi, 2003; Watts and Simon, 1999), including the relevant facts (Mikk, 2000)
- structure and sequence subject matter (Task Team for the Review of the Implementation of the National Curriculum Statement, 2009; Reys and Reys, 2006; Pepin and Haggerty, 2003)
- identify relevant mathematical skills (Basson, 2002)
- alert teachers to students' prior conceptions about specific content (The Centre for Curriculum Materials in Science, 2006)
- suggest ways to activate students' prior knowledge (The Centre for Curriculum Materials in Science, 2006)
- provide examples of teaching strategies and activities (Taylor, 2008b; Reys and Reys, 2006; Henson, 2004; Pepin and Haggerty, 2003; Izsak and Sherin, 2003; Litz, 2001; Cook and Tulip, 1992; Dekker, 1983)
- can be used as the basis for interpretation and discussion (Litz, 2001; Mikk, 2000)
- scaffold student learning (The Centre for Curriculum Materials in Science, 2006)

- provide resources for continuous formative assessment to identify residual misconceptions and indicate areas of inadequate knowledge or skills (Lumby and Foskett, 2005; Chatterji, 2003; Davis, 2003b; Brooks, 2002; Skevington, 1994)
- provide recordable summative assessment measurements of the extent to which the individual learners have reached the outcomes (The Centre for Curriculum Materials in Science, 2006; Watts and Simon, 1999).
- offer logistical support by describing where to find supplies and how to utilise resources, especially if they offer kits, which provide a complete set of "ready-to-use" consumable and non-consumable supplies for the classroom (Davis, 2003b; Luft and Patterson, 2002)

2.3.4.9 Characteristics that aid professional development of teachers

The explosion in knowledge creation, technological development and understanding of cognitive processes discussed in Section 2.1.1 has some serious consequences for teachers. They must be able to comprehend and present their learners with new subject knowledge and implement new research results on learning in their teaching. This requires life-long professional development for the teacher.

The textbook or teachers' edition of a textbook can support teachers' professional development by providing

- content or subject knowledge
  The subject knowledge can be compactly represented in the teachers' edition to orientate the teacher (Singer and Tuomi, 2003). Discussions on frequently held misconceptions on the specific science content and effective strategies to bring about conceptual change in learners can alert teachers to their own misconceptions and change their conceptions (Davis, 2003a).
- pedagogical knowledge (Singer and Tuomi, 2003)
  When teachers use good textbooks, it can indirectly change teachers' knowledge and beliefs about how children learn and which teaching strategies are most effective. Malcolm and Alant (2004, 73) even refer to textbooks as "catalysts for changing teacher practice." Furthermore, teacher's editions that explain the results of research about how students learn and how the process of conceptual change can be supported, can play a valuable role in teachers' professional development (Huber and Moore, 2001; Singer and Tuomi, 2003; McKenney, 2001). It contributes to the teachers' pedagogical knowledge and, therefore, their professional development (Fishman and Davis, 2006). Textbooks must provide strategies for teachers to facilitate

meta-cognition in learners. These can include any activity that causes learners to reflect on what they have learned and even "sample classroom dialogues" (Davis, 2003b, 3). For example, answering high-level questions or preparing written or verbal presentations can compel learners to formulate and even evaluate their conceptions of the scientific content. A teacher who uses a textbook that provides these strategies for specific content, should gain generic pedagogical knowledge about the learning principles that are involved and should be able to apply it to other science content (Davis, 2003a).

2.3.5 Conclusion

The first part of this chapter was devoted to an exploration of education in general and science education in particular. This provided the backdrop for a discussion on textbooks in science education. The discussion cited research results about how learners learn from textbooks and described the models of text comprehension involved. The functions that textbooks must fulfil to support learners and teachers in science education were probed and finally attention was given to the characteristics that must be present in textbooks to enable them to fulfil the various functions and ensure effective support of the learner and teacher in order to be a good textbook.

Chapter 3 will focus on textbook quality and how the quality of textbooks can be assessed. Possible approaches to assessment investigations will be considered and existing methods and techniques used in quality assessment will be discussed.

Chapter 3

The assessment of the quality of science education textbooks

3.1 Introduction

In Chapter 1 I demonstrated that science education is important for all learners and that science education in South Africa is hampered by unqualified, under-qualified and inexperienced teachers. It was argued that good textbooks could play a valuable role in solving the problem, since it is a well-established fact that the utilisation of high quality textbooks can contribute to learner achievement. The quality of the available textbooks vary and teachers or education providers should choose the best textbooks available for teachers to use, or at least ensure that the textbooks chosen and used are adequate. Consequently, it is necessary to assess the quality of the textbooks during the process of textbook selection.

In Chapter 2 I illustrated what a good science education textbook should do and look like. I started by considering education in general and science education in particular. Subsequently, the role of the textbook in science education and the characteristics that contribute to textbook quality were investigated. Against this background of what a good textbook should do and what characteristics a good science education textbook should have, we can consider how to judge the quality of available textbooks.

The various methods that can be used to assess the quality of textbooks were briefly mentioned in Section 1.2.1. In this chapter textbook assessment is considered in detail. The available research results in this field have been investigated to identify a method or methods that is or are appropriate for evaluation of science textbooks.

3.2 Textbook research

Textbooks have often been branded the culprit for problems in society. Textbooks, especially history textbooks, have not only been indicated as the cause of many crises and wars, but also as the means to prevent them (Mikk, 2000; Pingel, 1999). This is still the opinion of many people. Sewall (2005) even stated that editors of history textbooks in America "give the nation's students a misshapen view of the global past and a false view of the global future" (Sewall, 2005, 500). The words of Commisso (2004) explain the rationale behind these opinions:

> When a child starts going to school, the book is a window to the world. Thoughts and images, ideas and stories that are conveyed by the textbook contribute gradually to shaping the thought and personality of a child, so that he/ she can change and become either an adult with prejudices, who discriminates and promotes violence, or an adult who is able to work for a better world, free of prejudice and inequality. His/her way of thinking is linked to the way in which the textbook is conceived, designed, and introduced to the child (Commisso, 2004, 11).

It is, therefore, understandable that investigations into the content of and treatment of delicate issues in textbooks have been responsible for much of the textbook research (Pingel, 1999; Johnsen, 1993) and is still the focus of many research projects (Dimopoulos et al., 2005). One example of a recent South African study that investigated how textbooks handle delicate issues is the HRSC study: *Textbooks for Diverse Learners: A critical analysis of learning materials used in South African schools* (McKinney, 2005). The study aimed to explore the extent to which the textbooks currently used in schools in South Africa "reflect and reinforce the post-apartheid vision of a non-racist, non-sexist, equitable society" (McKinney, 2005, 11). However, the research in the field of science curricula and textbook analysis is not limited to sensitive issues. The research is multifaceted and school science textbooks have been analysed, more often than not with regard to content, but also with a focus on linguistic and sociological aspects (Koliopoulos and Constantinou, 2005).

Textbook research has become an established field with many organisations that coordinate research and disseminate research findings through publications and international conferences (Pingel, 1999) and many articles on textbook research have been published. For example, the ERIC database reported 222 studies on school science textbooks in the period 1985 to 2002 (Dimopoulos et al., 2005). According to Weinbrenner (1992) research on textbooks vary from process-orientated to product-orientated, and includes the assessment of textbook influence. One of the organisations involved in textbook research is the International Association for Research on Textbooks and Educational Media (IARTEM), an independent, non-profit organisation. They organise bi-annual international research conferences and publish conference reports to

- promote research on textbooks and educational media,
- establish contacts between all parties interested in educational media and textbook issues, and
- strengthen the focus on educational media and textbook issues in teacher education and teacher training (IARTEM, n.d.).

Smaller groups like the Teaching Resources and Textbook Research Unit (TREAT) of the Faculty of Education at the University of Sydney works closely with IARTEM on textbook research. TREAT is a research unit in the School of Policy and Practice, Faculty of Education, at the University of Sydney (The University of Sydney, n.d.).

In the field of social studies UNESCO, in co-operation with the Georg Eckert Institut für Internationale Schulbuchforschung (Braunschweig, Germany), has been developing the *UNESCO International Textbook Research Network*. This network gathers and disseminates information on "new approaches, institutions and projects being carried out in various countries with regard to research into textbook development and revision of history, geography and social science textbooks" (Pingel, 1999, 3). It also coordinates the research and revision of social study textbooks on an international scale (Pingel, 1999).

A number of organisations interested in education or science education have made textbook quality the focus of special projects. One example is Project 2061 of the American Association for the Advancement of Science (AAAS). The project started in 1985 – the year Halley's Comet was last visible from earth. The next time this will happen will be in 2061. The name of the project was chosen as a reminder that today's education will shape the quality of children's lives as they come of age in the 21st century amid profound scientific and technological change. The project conducted a series of in-depth studies examining the quality of middle and high school science textbooks (Kulm et al., 1999). The project raised great concern about the quality of science textbooks in the USA (National Science Board, 2004; Jones, 2000; Budiansky, 2001). The National Academy of Sciences in America is another of the organisations that joined the action for better science textbooks in the USA. They appointed a committee to develop the capacity to select effective instructional materials. The committee developed a guide for the selection of instructional material for K-12 Science (Singer and Tuomi, 2003). A third example of a special project that focused on science textbooks is The Center for Curriculum Materials in Science (CCMS). It is a collaboration of Project 2061 of the AAAS, Michigan State University, Northwestern University and the University of Michigan. It focused on the analysis, design, and use of science curriculum materials funded through the National Science Foundation (Center for Curriculum Materials in Science (CCMS), n.d.).

To communicate research findings and to help teachers evaluate science textbooks numerous publications have been issued and a number of websites constructed. Examples of such websites are the AAAS site that reports the results of evaluations of American Science and Mathematics textbooks (ESchool News Online, n.d.; Roseman et al., 1999) and ScienceTextCentral.org, a website launched by the Environmental Literacy Council to help educators choose the best science textbooks for their classrooms (ScienceTextCentral, n.d.). The HRSC published a report to communicate the results of their research on the handling of race, gender, social class and disability in textbooks in South Africa (McKinney, 2005).

The Departments of Education of different countries have given attention to textbook quality. In the USA, for example, twenty different states "adopt" textbooks for statewide use (NSTA, 2003). The Education Boards of these states have developed their own sets of criteria for the evaluation of textbooks (Tyson, 1997). In 2001 the Texas Education Agency in Texas, one of the adoption states, even awarded a $80 000 contract to the Science Faculty at the Texas A&M University to check the accuracy of the proposed science textbooks (Jepson, 2002).

Another example of an education department that recognises the importance of textbook selection can be found in Hong Kong where an ad hoc committee of the Department of Education focuses on textbook quality. They formulated (and still update) a document on the principles for evaluating the quality of textbooks (Hong Kong Department of Education, 2002). Even in Luthuania, one of the last countries in Europe to set up a free textbook market, the Textbook Research and Information Centre for the Baltic Countries (VTIC) organised a conference in 1998 to discuss textbook analysis criteria (Bakonis, 1998).

In South Africa the quality of textbooks has also received attention from the National Department of Education. In 2005 McKinney (2005, 38) recommended the "creation of a unified national list rather than disparate provincial lists as is current practice" in order to prevent situations where a textbook selected for the list in one province is rejected by another. Currently the National Department of Education compiles a national catalogue of approved textbooks for the

FET band, from which schools, that receive textbooks from the Department, can choose (Task Team for the Review of the Implementation of the National Curriculum Statement, 2009; Motshekga, 2009; Mahlaba, 2006), although textbooks for the GET band are still approved on provincial level. The Western Cape, Gauteng and National Departments of Education have made their textbook analysis instruments available (Western Cape Department of Education, 2005; Mahlaba, 2006; Department of Education of South Africa, n.d.). These instruments will be discussed in Section 3.9.4.

The teaching of languages have always been closely associated with the use and selection of textbooks. Many articles in journals and also online journals discuss the work of researchers on the evaluation of textbooks for use in Teaching English First Language (TEFL) and Teaching English Second Language (TESL) (Miekley, 2005; Garinger, 2003; Ansary and Babaii, 2002).

A number of individuals or smaller organizations are also investing time and effort in working for better quality textbooks. One example is the The Textbook League – an independent California-based group (established in 1989) that publishes a bimonthly newsletter with textbook reviews (Chandler, 1999).

3.3 Defining textbook quality

Two complementary definitions for textbook quality were given in Section 1.2.1. The two definitions indicate respectively that if a textbook is good:

- the learners will *achieve the intended learning outcomes* (Bernier, 1996)
- the textbook contains the features and characteristics that make it *fit for the purpose* of supporting learners and teachers in the learning process (Sursock, 2001)

The two definitions represent what may appear to be two sides of the same coin: a textbook can obviously only enable learners to achieve the intended learning outcomes if it is fit to do so. However, these definitions provide two different focal points when considering the quality of a textbook. The first definition draws attention to the achieved learning outcomes or *attained curriculum* as indicators of the quality of the textbook, while the second points to the textbook or *potentially implemented curriculum* itself as a source of information on its quality or fitness for its purpose. The complementary nature of the two definitions is illustrated in the definition provided by Johnsen (1993, 221) :

> A (text)book will achieve optimal effectiveness if the material between its covers is written and adapted in such a way that allows most of the pupils, during the time available to them, with or without the guidance of a good teacher, to study the book and grasp the knowledge, understanding and skills specified in the curriculum, as measured by tests and examinations administered at different levels.

Evaluation can only be done by considering textbooks in relation to their purpose. According to both definitions it is imperative that the desired or intended learning outcomes or goals that the learners should reach, as well as the role of the textbook (the didactical functions thereof) in the process, should be specified in detail (Van de Grift, 1989; New Jersey Curriculum Framework, 1998). For example, if the curriculum stipulates recall of scientific facts, the first definition suggests that the number of facts the students are able to recall will indicate the textbook's quality, while the second definition requires the presence of features in the textbook that research has identified as features that enhance retention in order to qualify as a good textbook. Most current science education curricula, however, desire the development of scientific skills, understanding, and the ability to apply scientific knowledge, as discussed in Section 2.2.2. The quality of a textbook can therefore be observed in the skills, understanding and application of knowledge attained or outcomes reached by the learners who used the textbook or in the extent to which features are present in the book that will guide learners in acquiring these skills and understanding, and practicing the application of scientific knowledge.

3.4 Assessment, measurement and evaluation of textbooks

In the research on textbooks, and specifically research on textbook quality, the terms "assessment," "measurement" and "evaluation" occur frequently. Each of these terms is used to refer to an investigation with a slightly different purpose, although it is not always used with semantic precision. *Assessment* refers to the collection of data to describe or better understand an issue (Wikipedia, n.d.a). In the context of textbooks and their quality assessment, studies aim to gather data on aspects that influence textbook quality and to interpret such data in order to understand and describe it. In the context of learning and teaching the term "assessment" has become synonymous with "evaluation."

Evaluation is not merely a process of determining facts about things. Evaluation goes beyond describing (Rallis and Bolland, 2004); it generates knowledge that leads to a conclusion about the "worth, merit or significance" of the object under investigation (Scriven, 2003, 16). Evaluation can be formative and developmental or summative and judgemental or even both (MacDonald, 2006). Three key concepts merge in evaluation: systematic inquiry; judgement of merit, worth, value, or significance; and information for decision making. Furthermore, three basic activities are involved in evaluation: description, comparison, and prediction (Rallis and Bolland, 2004).

Measurement is the process of quantifying assessment data (Huitt et al., 2000). Measurement can convert qualitative data to quantitative data for easy comparison. For example, in their assessment of textbooks' concern for explanatory understanding, Newton, Newton, Blake and Brown (2002) divided the text into 'clauses' (text units) and classified according to the type of clause (according to a scheme). It was then possible to count the number of clauses that fit into each category. Another example can be found in the work of Adolfsson and Henriksson (1999). They coded the exercises and text of mathematics textbooks according to the categories of maths that were covered and calculated these as percentages.

Rallis and Bolland (2004, 4) contend that

> ... evaluation is characterised by disagreements as well as by consensus. What counts as data for evaluation? ... The very word evaluate is based on the word value and therefore suggests a judgment. But who makes the judgment – the evaluator or the person(s) who commissioned the evaluation?

Scriven (2003) identified eight models of evaluation. He classified these according to the role the evaluators play in the process:

- Quasi-evaluation model: External evaluators are sometimes called in to evaluate something to support decision-making, but the evaluation is limited to the client's preset criteria. In this study the development of the instrument is based on an investigation of the real evaluation process, but the application of the instrument is evaluation according to the Quasi-evaluation model. It is merely the gathering of data, not real evaluation. *A typical example of quasi-evaluation in textbook evaluation is a teacher on a panel that evaluates textbooks for approval by the education department and is restricted by the use of criteria prescribed by the education department.*
- Goal-achievement model: The overarching criterion is that the goals must be reached. This has also been referred to as the discrepancy model. The evaluation aims to identify any discrepancy between the goals and the actual achievement. No (expected or unexpected) side-effects are considered. *This*

is usually achieved through experimental studies.

- Outcomes-based model: Evaluation for the sake of increased accountability. The evaluation result must justify the product. *Publishers that aim to use the results of their evaluations as marketing propaganda or to gain departmental approval follow this model.*
- Consumer-orientated model: In the first three models the criteria were determined by the management. In the consumer-orientated model the criteria are formulated to determine if the consumer's needs are met. *For example, learners would prefer textbooks that pose no challenge and consumer-orientated evaluation may use this as criterion. However, research has shown that the ideal level of text difficulty is learners' zone of proximal development (discussed in Section 2.3.4.3).*
- Formative-only model: Not directed at a summative or overall evaluation, but only at improving the product. *This was the model of evaluation applied by Britton, Dusen, Gülgöz and Glynn (1989) when they investigated the effect of expert rewriting of text on the retention of text information.*
- Participatory model: All stakeholders are encouraged to become involved in the evaluation process. This includes collaborative or empowerment evaluation. *Teachers who are involved in the evaluation of textbooks follow the participatory model of evaluation that can empower them to better understand the strengths and weaknesses and use the textbooks more effectively.*
- Theory-driven model: Evaluation with the core function of generating explanations for the success or failure of the product.
- Power model: Evaluation that is committed to contribute to social engineering.

An evaluator's role can fit into more than one of the models described by Scriven (2003). In the next section I intend to justify the commitment of time and other resources to an investigation of textbook quality.

3.5 Reasons for evaluating textbook quality

The importance of textbook quality and the reasons why textbook evaluation is a worthwhile process has been addressed in Chapter 1. From the discussion it is clear that *textbook quality* is important because of its ability to

- increase learner achievement (Taylor, 2008b; Brandt, 2005; Reddy, 2005; Singer and Tuomi, 2003; Lubben et al., 2003; Kesidou and Roseman, 2002; Adolfsson and Henriksson, 1999);

- contribute to the professional development of teachers (Malcolm and Alant, 2004; Singer and Tuomi, 2003; Ansary and Babaii, 2002; Huber and Moore, 2001; Watts and Simon, 1999; Smith et al., 1993); and
- aid curriculum implementation (Task Team for the Review of the Implementation of the National Curriculum Statement, 2009; Davis, 2003a; Iszak and Sherin, 2003).

There are many reasons why textbook quality should be evaluated. These reasons fall within the three conceptual frameworks introduced by Chelimsky (MacDonald, 2006) to summarise the *purposes of evaluations*:

- evaluation for accountability (e.g. measuring its efficiency to provide data to assist teachers in their choice of textbooks) (Olivares-Cuhat, 2002; Sherin, 2002; Roseman et al., 2001; Litz, 2001; Tyson, 1997; Wong, 1991)
- evaluation for development (that leads to the publication of better textbooks and to enable teachers to identify the weaknesses in the textbooks they use in order to compensate for it) (Mozambique Ministry of Education, 2002; Roseman et al., 2001; Tyson, 1997; Chambliss, 1994; Britton et al., 1993)
- evaluation for knowledge (to contribute to the professional development of the teachers involved in the evaluation process) (Mozambique Ministry of Education, 2002; Litz, 2001)

The evaluation of textbooks should be done on a national scale (Task Team for the Review of the Implementation of the National Curriculum Statement, 2009). According to the Department of Education of the Republic of Mozambique (Mozambique Ministry of Education, 2002, 14) a *national evaluation system* will

- ensure better alignment between curriculum and textbooks
- ensure that the values and priorities the nation has identified are integrated into the education arena
- provide an effective mechanism to ensure equality of access and delivery at the school level
- assist with training and support of under-trained and under-qualified teachers
- ensure that state funds are spent on high quality, appropriate learning materials

This discussion on *why* textbook quality should be evaluated, is followed by a discussion on *how* it can be accomplished. The next section will be devoted to a discussion on the different approaches to evaluating quality and in Section 3.7.1 a classification system for all possible evaluation investigations will be proposed.

3.6 Approaches to assessment of textbook quality

The two definitions of quality, with their different focal points, mentioned in the previous section, form the basis for two different approaches to the evaluation of textbook quality. The quality can be assessed by

- determining whether the learning outcomes have been attained by learners using the textbook (first definition) or
- determining whether the textbook has the characteristics that makes it fit to guide learners to attain the outcomes (second definition).

The first definition and the corresponding approach establishes the programme goals and objectives. Data are collected about student performance and are then compared to the objectives to determine the effectiveness of a programme in helping learners to achieve the objectives (Rallis and Bolland, 2004).

Chambliss and Calfee (1998, 190) refer to a textbook of good quality as "a good learning tool." Consider the following analogous situation of another tool: A nutcracker is supposed to crack nuts. According to the first definition the quality of a nutcracker is therefore determined by its ability to crack a nut. However, there are usually more specific requirements. For example, a good nutcracker is not only able to crack the shell of a nut. The desired outcome is that the shell is cracked in such away that the nut is released without being broken or squashed. The most obvious approach to an investigation into the quality of a nutcracker will involve using the nutcracker to determine whether it can crack the shell of a nut, and also if it is cracked in such a way that the nut is released unbroken. It is possible for someone to observe the interaction between the user and the nutcracker while it is being used to crack a nut. This observation can provide information about how much effort is required to crack a nut. In the same way the textbook can be used to determine whether, how and how much the textbook contributes to reaching specific learning outcomes. This requires *experimental investigations* in the learning situation, where textbooks are used to gather information about their quality (Chambliss and Calfee, 1998).

Using, studying and experimenting with a few nutcrackers will provide the user with a good idea of the characteristics of a good nutcracker; or rather a nutcracker with the potential to crack effectively. For example, your research might show that nutcrackers made of metal perform better than those made of other materials. This provides the basis for another approach. When the nutcracker we evaluate can not be tested, for example when considering the purchase of a new one in a shop, it may then be inspected to determine if it possesses the desired characteristics. This can be done anywhere without requiring a nut. If the nutcracker does possess the ideal characteristics, it should be able to crack a nut effectively. This should however be verified (when possible) by using it (Britton et al., 1993). Similarly, research on the various aspects of learning with text and textbooks have indicated certain characteristics as contributors to learning (Iding et al., 2002; Newton et al., 2002; Glynn and Takahashi, 1998; Weiten et al., 1999). These characteristics were discussed in Section 2.3.4. A textbook can be inspected to determine if it has the desired characteristics. The "quality" of the textbook is regarded as the latent variable (Bernier, 1996) and the desired characteristics are the "indicators of the quality" or "distinguishable attributes of quality" whose presence are presumed to indicate levels of quality or "probable degrees of effectiveness" (Johnsen, 1993, 222). To determine a textbook's quality from this perspective requires *textbook analysis*. This technique, that originated after the end of World War I (Johnsen, 1993, 21), is the evaluating of data about textbook characteristics against a set of criteria to determine if it is fit to support the learner and teacher and enable the learner to reach the desired learning outcomes.

An additional, indirect, third approach is possible, using people as data source. This approach relies on other people's experience or opinions regarding the quality of a nutcracker. You can ask someone who is supposed to know about the quality of the nutcracker, like the shop assistant. When considering the quality of a textbook the respondents whose opinions are asked can be learners, teachers or experts on science or pedagogy. The respondents can give information on their opinion of the textbook or (if they have used it) on their experience of learning/teaching with the textbook. *Respondent opinions* are easy to acquire and frequently used. An indirect way in which the opinions of teachers are utilised in order to get an indication of textbook quality is the use of *popularity* as selection criterion. Publishers use sales figures (popularity) as a means of "proving" how good a specific textbook is. Popularity is seen as an indication of other teachers' positive opinion of the textbook: a book that sells well must be good because many teachers have chosen it (Ansary and Babaii, 2002).

3.7 Assessment procedures

3.7.1 Classification of assessment procedures

Different categorisations of textbook assessment approaches are given in the literature. The approaches to textbook evaluation discussed in the previous section corresponds to the categories of textbook assessment procedures suggested by Mikk (2000):

- analysis of textbooks
- experimental investigation (measuring outcomes)
- respondent opinions

As mentioned in Chapter 1 each of these methods has the potential to provide information on the quality of a book, but using two or more methods in conjunction can enable researchers to test the validity of results (Mikk, 2000). Huitt et al. (2000) classify textbook investigations according to the type of investigation: descriptive, correlational, or experimental.

Cunningworth and Ellis based their categories of types of material evaluation on the time of the evaluation in relation to the stages of its use (Litz, 2001):

- 'predictive' or 'pre-use'
- 'in-use'
- 'retrospective' or 'post-use'

A more comprehensive and explicit classification can be developed by considering the context in which the assessment is done and the type of data that is gathered. The data compiled in investigations (regardless whether it is gathered inside or outside the learning situation) can be either quantitative, qualitative or a mixture of qualitative and qualitative data. A suggested classification system for the different types of investigations involved in textbook assessment is depicted in Figure 3.1.

The remainder of this section will be dedicated to a quick overview of the types of investigations covered by the classification system. The merits and limitations of the different types of investigation will be discussed in the following section.

Experimental investigation in the learning situation are represented by the second column. Teachers can "test drive" new science materials (Davies, 2003). These 'in-use' investigations, alternatively called "try outs" or "learner verification" (Chambliss and Calfee, 1998, 176), can provide quantitative data, qualitative data or a mixture of both. Area E_1 represents investigations that gather

	Research in the learning situation		**Research outside the learning situation**	
	Experimental Investigation	**Respondent Opinion**		**Textbook Analysis**
Qualitative	E_l	RI_l	RO_l	A_l
Qualitative and qualitative	E_{ln}	RI_{ln}	RO_{ln}	A_{ln}
Quantitative	E_n	RI_n	RO_n	A_n
Stage of use	**In-use and retrospective**		**predictive**	

Figure 3.1: A suggested classification system for the different types of investigations involved in textbook assessment

only qualitative data. An example would be an investigation in which students are asked questions to determine how their understanding increased (if the learning outcome was reached) while using the textbook, while no attempt is made to quantify their understanding. Investigations in area E_{ln} produce both qualitative and quantitative data. The most common experimental investigations are represented by area E_n. Usually pre and post tests are utilised to quantitatively measure increase in skills, knowledge and understanding.

Besides measuring students' learning outcomes to get an indication of textbook quality, experimental investigations also provide the opportunity to study learners' interaction with text. One example is tests that measure text comprehension, instead of text acquisition (learning outcomes). In these tests the learners can refer to the texts during the test to find facts and formulate their answers. The tests give an indication of the quality of the text by measuring how well the text supports the learner in the learning process. It is important to realise that the learners who participate in these experiments must be representative of the learners for whom the textbook is intended.

Wilson (n.d.) mentions the following alternative measures of text comprehensibility that involve reader participation:

- reading speed procedures, which assume that a subject can read easy texts faster than difficult ones
- rating procedures, which require a subject to assess his or her own comprehension of a text on a scale of 1 (=low) to 5 (=high)

- recall procedures, which invite a subject to reproduce orally or in writing the content and structure of the whole or a part of a text
- cloze procedures, which omit every fifth word in a text and expect of a subject to fill in the missing words, amounting to one-fifth of the text
- question procedures, which use subjects' answers to questions about the content of a text to gauge their understanding of the text
- action procedures, which require a subject to read a text with instructions and then to carry out the prescribed action
- thinking-aloud procedures, which expect subjects to verbalise the process of decoding a text's meaning
- eye-movement procedures, which register the number, position and duration of eye fixations over print Although these procedures afford subtle and valuable insights into reader-text interaction, they are often complex, time-consuming and costly to implement and interpret.

The *cloze prosedure* mentioned by Wilson (n.d.) is widely used to determine how difficult the text is for students to comprehend, and therefore if it is fit for the purpose of supporting the learner. Sweet (2005) even calls the *cloze procedure* the cornerstone of reading assessment for ESL (English Second Language) teaching. The cloze prosedure involves deleting every *n*-th word in the text. As described by Wilson (n.d.), every 5th word is often omitted. Students are then asked to fill in the blanks. The higher the percentage of correct fillings, the easier it is to understand the text (Cain and Oakhill, 2006; Mikk, 2000; Johnson, n.d.). Mikk (2000) reported a number of studies that found good correlations between cloze tests and different measures of learning outcomes.

The third column represents investigations in which learners and teachers who have used the textbook give their comments on their experience of learning and teaching with the textbook. Teachers in America are often paid to pilot new textbooks and give publishers feedback on how well the lessons, exercises or experiments work (Hubisz, 2003; Tyson, 1997). Publishing companies like to use these comments in their marketing campaigns. This can be 'in-use' or 'retrospective' evaluations according to the time of use in relation to the time of the evaluation. The comments may be unstructured, or it can be structured by supplying a list of specific questions for the respondents to answer. RI_1 refers to investigations where no attempt is made to quantify the data.

The Nevada Department of Education (n.d.) uses the *predictive* opinion of evaluation committees in combination with the *reflective* opinions of teachers involved in classroom piloting of the textbook to assess textbook quality. This is a combination of RI_1 and RO_1 where participants were asked to answer "Yes, No or

Not Applicable” to seven overarching criteria (Nevada Department of Education, n.d.).

Area RI_{ln} represents investigations where the respondents in these investigations are asked to give a numerical rating of any of the textbook’s specific characteristics or its performance in specific functions, in addition to the written answers. For example, Nitsche (1992) questioned his students about their textbooks, using both open-ended questions and a scale from 1 to 9 to determine how helpful they experienced certain characteristics of the textbook.

If the respondents who have used the textbook are only asked to provide numerical ratings (on characteristics, functions or the textbook as a whole), the data will only be quantitative. These investigations will fall into area RI_n. There are many examples of instruments that were developed for this kind of evaluation (Abusharbain, n.d.). Examples of such investigations can be found in the ‘Kid-rating’ implemented by Muther and Conrad (1988) and the textbook evaluation form developed by Crystal Springs Books that uses a scale rating 1 (poor), 2 (fair), 3 (good) and 4 (excellent) (Crystal Springs Books, n.d.).

Assessment prosedures that consider the textbook in isolation, outside the learning situation, are represented by the forth and fifth columns of the diagram. These are ‘predictive’ or ‘pre-use’ evaluations. Area RO_l represents investigations where learners, teachers or experts who are not using and have not used the textbook are asked to give their opinions on the quality of a textbook, and the opinions are not quantified. Open-ended questions or checklists of desired characteristics can be used to help respondents formulate their opinions. The characteristics can be expressed as statements (California Department of Education, n.d.; Roseman et al., 2001; Chiappetta et al., 1993) or as questions (Garinger, 2003; Brown, 1998).

If respondents supply qualitative statements as well as ratings to textbook quality or aspects thereof the investigation will fall into area RI_{ln}, which refers to procedures that involve both qualitative and quantitative data. The National Science Foundation (NSF), for example, designed a review framework that requires written responses as well as an overall numerical rating on a five-point scale. An investigation can be designed to obtained both qualitative and quantitative data from respondents evaluating the textbook, or alternatively, respondents’ statements can be numerically encoded by the researcher. Encoding makes it possible to handle a lot of qualitative data and to compare huge volumes of complex unstructured data (Walliman, 2005). It is also possible to convert quantitative measurements or ratings to qualitative statements. For example, if the readability is measured with a readability formula, the result will be a numerical value, but the analyst can interpret it as stating that the "readability level is correct for the intended users." The qualitative statement is easier to understand than the numerical values that require knowledge of the relevant readability formula.

Some assessment procedures outside the learning situation produce only quantitative respondent opinions of textbook quality and will fall into area RO_n. One example of such an assessment can be found in the field of Teaching English Language. Brown (1998) developed a textbook evaluation system and a score sheet to help university conversation teachers select textbooks. It is based on teacher opinion and gathers only quantitative data.

Analysis of the textbook can give qualitative data (A_l), quantitative data (A_n) or a mixture of both (A_{ln}). A checklist of questions requiring only positive or negative answers can result in qualitative data, but if the number of positive answers are counted to give a summary of the evaluation data, this will result in a mixture of qualitative and quantitative data. The most common type of analysis, however, produces quantitative data that the analyst obtains by counting and measuring aspects of the textbook to give an indication of its quality.

3.7.2 Merits and limitations of assessment procedures

The analyst or evaluator must choose the best approach and methods to help understand the quality of the textbook according to the stated goals, objectives, context, and given available resources (Rallis and Bolland, 2004). The following merits and limitations of the procedures guide this decision.

3.7.2.1 Experimental investigations

Experimental investigations that measure the learning outcomes reached by students using the textbook is the most obvious approach to textbook evaluation (indicated by the first definition of quality as discussed in Section 3.5). Unfortunately this approach is not without its share of problems. The following aspects of experimental investigations can be problematic.

- Measuring techniques. Many different measures of learning outcomes have been utilised in educational research. These measures range from task performance and test scores to learners' perceptions of their learning. The measurement instruments used should give reliable and valid measures of the learning outcomes that have been reached (Van de Grift, 1989). If the measuring instruments influence the results it will give a misleading idea of the outcomes reached by the students (Mikk, 2000).

- Time and experimental design. Unfortunately designing and carrying out these experimental investigations can be a daunting task (Chambliss and Calfee, 1998) and it is often too time-consuming and difficult to implement (Britton et al., 1993). The textbook is only one of several factors that influence learning. To attain a measure of the textbook's contribution to the learning outcomes all other factors affecting the learning outcome must be taken into account in the experimental design (Montgomery, 2006; Reys and Reys, 2006; Clark and Salomon, 2001). The learners and teachers involved should be representative of all learners for whom the textbook is intended. A control group not using any textbook should also be used (Van de Grift, 1989). This control group should exist of learners with the same abilities (especially reading abilities), interests and background, and they should, preferably, be taught by the same teacher. It is also essential to consider alternative explanations for outcomes that are reached (Van de Grift, 1989). This ideal design is seldom realised and consequently Singer and Tuomi (2003, 38) of the National Academy of Science found that:

  > There is no substantial body of research that tries to evaluate the effectiveness of particular instructional materials as a separate variable in the total learning experience.

- Limited scale. Suitable large-scale experimental trials in which learning is measured are so difficult, expensive and time consuming that rigorous studies that relate textbook content to student achievement can rarely be implemented (National Science Board, 2004; Mikk, 2000; Britton et al., 1993; Johnsen, 1993). Consequently, experimental investigations are mostly limited to studies of specific topics or the effectiveness of specific teaching strategies in textbooks (Iding et al., 2002; Newton et al., 2002; Glynn and Takahashi, 1998).
- Ethical considerations. The research represents "severe interference with the teaching"(Johnsen, 1993, chapter IV) and there is also the possibility that the development of students can be hindered by participating in experiments with "unsuccessful" textbooks (Mikk, 2000).
- Uncontrolled novelty effects. Research subjects in the experiments tend to give increased effort when the textbook is new to them. This may yield an increase in achievement (Clark, 2001).
- Reflective versus predictive research results. Experimental investigations give valuable information for the revision of the textbooks and aid the design of future textbooks. However, when a teacher needs to select a textbook it

is impossible for the teacher to use every available textbook before deciding which one to prescribe for the learners. Since the financial implications make it impossible to buy new textbooks too often, teachers need to be able to select the best textbook available. Publishers need to assess the quality of textbooks in order to decide which textbooks to publish. In this regard predictive respondent opinion and textbook analysis are the only viable options.

Though experimental investigation is the most obvious approach to assessing textbook quality, it poses serious challenges to the researcher with regard to ensuring the validity of the results. Viadero (2007) reports that many experimental investigations yield no significant discernable positive effects for the use of specific textbooks when compared to other textbooks. She based this report on four rigorous randomised-control trials — experiments in which mathematics textbooks were randomly assigned to either a treatment or a control group.

With regard to textbook selection the biggest handicap of experimental investigations lies in the fact that it is unsuitable in a predictive capacity.

3.7.2.2 Respondent opinions

Studies that use respondent opinions as source of information on textbook quality are hampered by a number of problems. The following problems are associated with determining textbook quality through respondent opinions.

- Validity of opinions as indicators of quality: Everyone may have an opinion on the quality of a textbook, but it is not necessarily an informed opinion. The opinion of learners is usually only asked after they have used the textbook. Their comments are then based on their experience of learning with the textbook. When the textbook is considered in isolation the opinions of experts on science and education are valued most, because it can be assumed that their opinions are informed by their knowledge science and how children learn (Mikk, 2000). These experts can be experienced teachers or teacher educators.
- Difficulty in comparing different respondents' opinions: When respondents are asked to give their opinions on the quality of a textbook "you can get buried in information without a clear idea of how to condense it into a manageable form" (Goldsmith et al., 2000, 62). To make it easier to compare different respondents' opinions, investigators usually provide categories of even a list of characteristics that the respondents should comment on. To

make it even easier to compare opinions, respondents can be asked to rate the textbook with regard to every characteristic on a set list.

- Quick and superficial opinions: Although it is quicker to complete an assessment that only requires numerical rating, researchers have found that investigations that are limited to checklists that only require ratings on a numerical scale lead to quick, superficial evaluations (Singer and Tuomi, 2003). To prevent this from happening, participants are often asked to provide evidence for the qualitative ratings they give, because evaluators

  > ... will more likely provide critical well-thought-out judgements if they are asked to make a narrative response to evaluation questions or criteria, rather than make selections on a checklist (Singer and Tuomi, 2003).

- Formulating opinions: Respondents sometimes find it difficult to formulate their opinions. The assessment system based on teacher opinion that was developed by Brown (1998) (mentioned in the previous section) is unique in the way that Brown guides the teachers to give their true opinions. For example, he advises them to look at the textbook from the back (what he calls "Japanese-style") and to ignore the covers that are designed to stimulate interest. He also guides the teachers to evaluate the activities by asking them if there are any exercises or activities in the last three lessons that they would want to replicate in a textbook they would write. Their answers to these questions are then converted to a rating that is eventually summarised.
- The influence of context on opinions: Teachers with different teaching experiences, who teach learners from different backgrounds in classrooms with different resources, may have different opinions about the quality of a textbook, due to their contexts (Montgomery, 2006).

3.7.2.3 Textbook analysis

A number of problems can be encountered during textbooks analysis:

- Stakeholders who choose the criteria. Determining the criteria is the most problematic component of the analysis process (Van de Grift, 1989; Franssen, 1989b). According to Stein et al. (2001) textbook evaluations in the USA lack research-based criteria for evaluating instructional material. In the USA the textbook adoption system in many states have led to evaluation boards with stakeholders from every interest group (Whitman, 2004). This can result in "dumbed down textbook content in an attempt to render them inoffensive to every possible ethnic, religious, and political constituency" (Whitman, 2004). Ravitch (2003) contends that weeding out the controversial themes and clashes between good and evil from textbooks make textbooks dull and render them incapable of engaging learners' attention.

- Time versus thoroughness. Unfortunately textbook analysis is a tedious and time-consuming task (Davis, 2003b; Olivares-Cuhat, 2002; Fetsko, 1992). This creates a delicate balance between trying to consider the aspects that are relevant to textbook quality and trying to restrict the amount of time devoted to the evaluation. Most researchers limit their analysis of a textbook to two or more topics or chapters in the textbooks assuming that it will be representative of the whole (Pepin and Haggerty, 2003; Olivares-Cuhat, 2002). Many of the documented investigations of science textbooks are limited to analyses according to specific perspectives and single topics:
  - McKinney (2005) explored the extent to which textbooks currently used in South African schools reflect and reinforce the post-apartheid vision of a non-racist, non-sexist, equitable society.
  - Ninnes (2002) investigated how much of the achievements in space science of the major powers, particularly the USA, is included, compared to local achievements.
  - The handling of the siphon and the concept "pressure" in eight USA science textbooks were studied by De Berg and Greive (1999) to determine their use of *explanatory* principles, e.g. plausibility, parsimony, generalisability and fruitfulness.
  - Newton et al. (2002) examined the extent to which 76 primary science textbooks showed a concern for explanatory understanding. The text was divided into clauses (text units) and classified according to the type of clause (condition, consequence, explanation, purpose, prediction, aim, attention directing, irrelevant or not differentiated). In most cases the textbooks did not support learning by stating aims, making or asking for predictions or directing the reader's attention.
  - Eltinge and Roberts (1993) measure inquiry in science textbooks (linguistic content analysis).
  - Bazler and Simonis (1991) studied high school chemistry textbooks to determine whether they are gender fair and how this changed from 1970 to 1990. The photos, drawings and text pages in the textbooks were categorised as depicting males or females. The researchers also classified the textbooks by looking for depicted or mentioned male or female stereotype interests.
  - Allington (2002) examined the readability of textbooks.
- Subjectivity versus objectivity. Tension exists with regard to subjectivity and objectivity in textbook evaluation. Quantitative data is often the result of

an investigation that aims to provide more objective, easily verifiable results (Stein et al., 2001; Johnsen, 1993), but it requires thoughtful design to ensure that it is a valid indication of the textbook quality. It is necessary to be aware of the problems inherent in analysis. Research results can be misused and difficult to apply. For example: research shows that well chosen headings aid learning by indicating text structure, but just checking for the presence of headings or counting the number of headings will not give an indication of its contribution to learning. More time and effort is required to analyse the structure of the text (as indicated by the headings) to determine if it will contribute to learning.

- Relative importance of criteria. A problem with textbook analysis is that the proportions of questions characterising different aspects may cause some aspects of quality in textbooks to be unjustly emphasised (Mikk, 2000; Johnsen, 1993). The evaluation criteria should not only incorporate the features or attributes that contribute to quality in textbooks, but also address the issue of how important the features are to the quality, in relation to each other. This problem will be addressed in Section 3.9.3.5.
- Training of analysts. Apart from the need for sufficient time for the evaluation, adequate training for the evaluators is often mentioned as an important aspect of evaluations (Singer and Tuomi, 2003; Kesidou and Roseman, 2002; Tyson, 1997; Chambliss, 1994). It was already mentioned in Section 1.1 that teachers do not usually receive any training in the evaluation of textbooks during their teacher preparation programmes (Stein et al., 2001; Johnsen, 1993). A further problem is that teachers are often not familiar with research results about learning. Unfortunately even the training of analysts may prove problematic. Project 2061 of the AAAS (mentioned in Section 3.2) found that most science textbooks in the USA are of poor quality (Roseman et al., 2001; Roseman et al., 1999). These reports sparked a lot of interest from different stakeholders, but not everyone agreed with their results. Holliday (2003) pointed out that the American teachers' opinion of the available science textbooks did not correspond with the negative results of the AAAS analysis. He criticised the AAAS's methodology, saying that the definitions and examples or training used might have biased the analysts (Holliday, 2003).

3.7.3 Motivation for further focus on textbook analysis as evaluation method

There are several reasons why textbook analysis deserves further attention in this study. The main reasons are:

- Textbook analysis can predict effectivity: The main advantage of textbook analysis is that it enables us to predict the value and efficacy of textbooks before they are used in school (Mikk, 2000). This can inform the selection process and thus contribute to both science education and teacher development. This predictive value of textbook analysis is the determining factor in selecting textbook analysis as the focus for further investigation in this study.
- Textbook analysis is a systematic and transparent process: Although expert opinion can also aim to predict the value of the textbook to the teacher and learner the result is less transparent since the criteria used by the expert is usually not stated. Textbook analysis is chosen in this study because it also satisfies the need to develop a transparent, scientifically justifiable method for textbook quality assessment. According to Mikk (2000, 77)

  > (t)he difference between expert opinion and textbook analysis is that experts have no need to fix the rules clearly for justifying their opinions or they may have no rules brought to the level of consciousness.

3.8 Textbook evaluators

Who should act as evaluators of textbook quality? The most obvious candidates are the teachers who have used the book, but teacher training does not usually include training in the evaluation of textbooks (Stein et al., 2001; Johnsen, 1993) and teachers are often not familiar with research results about learning. Consequently, teachers that must act as evaluators must receive adequate training (Singer and Tuomi, 2003; Stein et al., 2001; Chambliss, 1994). The AAAS provided their Project 2061 curriculum material evaluators with seven days of training. Their training aimed to

- clarify the ideas that served as the basis of the analysis
- model the application of the instrument on examples
- provide opportunities to practice using the procedure (Kesidou and Roseman, 2002, 526).

Evaluators were also supplied with a notebook with examples, clarifications et cetera.

Textbook analysis requires evaluators that have the necessary background knowledge and are prepared to take responsibility for the analysis process (Stein et al., 2001). The Mozambique Operational Handbook for Textbook Evaluators and Managers addresses the selection of members of an evaluation by considering the following questions (Mozambique Ministry of Education, 2002, 22):

a. What makes a successful Subject Panel Member? The evaluator must

- be familiar with the policy documents.
- be familiar with the subject curriculum and related other subjects and international trends in the subject.
- be familiar with the learning environment in Mozambique, i.e. the classroom realities in both urban and rural areas.
- have in-depth experience and knowledge of the appropriate subject area.
- understand the role of textbooks in providing access to the curriculum for pupils and training and supporting teachers to implement the curriculum.
- be thoroughly familiar with the criteria for evaluating books and how they meet the needs for the curriculum in the subject.

b. What skills are required?

- Impartiality and fairness, high ethical standards and transparency.
- Time management.
- Critical evaluation skills: having an overview of the curriculum and subject and being able to assess qualitatively whether the materials are of a sufficiently high standard and whether they meet the needs and requirements of the curriculum for the subject.
- Group work skills.
- Own task management and quality assurance skills.
- Management and control and delivery.
- Speed-reading skills.
- Teaching experience preferable, including experience at the appropriate school level.

The extent of the independence of researchers that assess textbook quality can pose problems. UNESCO (Pingel, 1999, 16) points out that

> ... a free researcher ... is not restricted by ministerial guidelines and is entitled to criticise official curricula and approved textbooks frankly,

> but the final recommendations he makes are not binding in any way. On the other hand, acting on behalf of ministries, curriculum institutes or publishing houses, the researcher may have a more direct influence on the implementation of research findings but the aims, methods and subject of research are often defined by other institutions, thus limiting the freedom to openly criticise their products.

In some provinces in South Africa, like Gauteng, teachers can apply to act as members on evaluation panels that investigate textbook quality in order to compile a list of approved textbooks (Mahlaba, 2006). The evaluation is done according to the instrument supplied by the relevant education department. Although this limits their freedom to evaluate according to their own criteria, it provides a mechanism for a transparent and uniform evaluation system. Acting as evaluator can also be to the advantage of the evaluator. The process and even the use of an evaluation tool can serve as a training opportunity for the teachers. It forces them to consider the philosophy behind their teaching, the goals of their teaching and it can make them aware of aspects of textbook quality that they have not previously considered.

3.9 Textbook analysis instruments

During any evaluation the evaluator needs to gather data that can describe textbook quality, compare it to other textbooks or a benchmark, and predict if the textbook will support the teacher and learner in learning science. To do this they identify or develop data collection instruments or methods, identify data sources or samples, and develop plans for data analysis (Rallis and Bolland, 2004). A textbook analysis instrument can consist of a single method or it can incorporate a number of different methods or instruments to gather and analyse all the data needed.

There are a number of reasons why the use of a textbook analysis instrument is necessary. The first is to ensure *thoroughness* by ensuring that no important aspect is forgotten or ignored. The application of a textbook evaluation system can make textbook evaluation "a coherent, systematic and thoughtful activity" (Ansary and Babaii, 2002) and contribute to the *validity* of the evaluation results. The analysis of textbooks is the evaluating of textbook with regard to every relevant characteristic (Mikk, 2000). The multitude of aspects that must be considered during textbook evaluation makes it advisable to have an organised system or instrument for the evaluation of textbooks. As Wingate (2002) pointed out: many aspects of evaluations are "too important to leave to chance or intuition." When textbook analysis is done to inform the textbook selection process, the use of an instrument for the assessment can ensure that the selection committee has all the relevant information that can inform their decision (Goldsmith et al., 2000; Fetsko, 1992). A few vague opinions of teachers or other stakeholders can not provide the information on which an important decision could be based.

Furthermore, instruments for the assessment of textbooks enable the *easy application* of the criteria (Olivares-Cuhat, 2002; Franssen, 1989b) and *standardise* the process (Davies, 2003; Goldsmith et al., 2000), which in turn will contribute to the *reliability* of the evaluation results. It can change textbook analysis from a daunting task to a systematic activity. When an analysis instrument is available to the stakeholders it can "standardise the ways committee members review" textbooks (Goldsmith et al., 2000, 53).

Once an instrument is designed and approved, future evaluation panels can save the *time* needed to decide on the criteria for an evaluation (Davis, 2003b). The instrument can be constructed in such a way that future changes in the educational context can be accommodated with an adaptation in the weights of criteria or the addition of criteria not previously incorporated.

The use of an instrument for the assessment of textbooks can also provide *professional development* for teachers. It can be an efficient and user-friendly format for more experienced teachers or experts to share lessons learned from classroom experience or research with other and especially less experienced teachers (Wingate, 2002).

The last and most important reason why instruments should be used for the assessment of textbooks is to provide *accountability* (Mahlaba, 2006). Education providers endeavour to "ensure that the (evaluation) process is transparent, unbiased and fair to all parties involved" (National Institute for Educational Development of the Namibian Ministry of Education, 2005, 1). Standardised evaluation instruments ensure "fairness and consistency of application" (Mozambique Ministry of Education, 2002, 38) and can contribute to stakeholders' perception of the unbiased and fair character of the analysis and the selection results. The instrument makes it easier to compare different textbooks objectively (Goldsmith et al., 2000).

Different formats of instruments for the analysis of textbooks exist. The most obvious is a *checklist* that reminds evaluators what to look for (Wingate, 2002). One example is the 'schedule' used by Pepin and Haggerty (2003).

Instruments for the analysis of textbooks can be limited to content analysis, readability formulæ and/or checklists of desirable characteristics. But it can also be more extensive systems that involve ratings and weights. Assessment of textbook content and readability will be considered in the next section, prior to a discussion of analysis instruments that cover a wider spectrum of textbook characteristics.

3.9.1 Content assessment

It is obvious that the textbook used must cover the topics that are prescribed by the curriculum, but it can not be assumed that a textbook does cover all the prescribed topics. Analysis of the content is necessary, especially in the USA where different states may have different prescribed content, but publishers print only one textbook that is used in different states.

Penney et al. (2003) classify the methods used for content analysis as

- qualitative or non-frequency analysis
- frequency counts of topics or terms
- frequency counts of combinations of categories
- spatial analysis or percentage of overall text

Qualitative or non-frequency analysis ascertain the presence of a topic (or outcome) in the textbook. The simplest method to determine if the topics stipulated by the syllabus are present in a textbook is to look at the book's content (the "flip test") (Marshall, 1991, 150) or to to consider the table of contents (Chambliss, 1994). The result of the analysis can be represented by a *table* with rows representing the different textbooks and columns in which is indicated whether the topic is included in the textbook, according to the topics indicated by the textbook's table of content (Ade-Ridder, 1989). Alternatively rows can represent the different topics and the columns each represent a textbook. Table 3.2 provides an example of this representation (Henson, 2004, 65). Note that the first column refers to topics of chapters in a teacher training textbook and not to criteria for textbook evaluation. For example Book F and Book G are the only books that do not contain chapters on adolescence and learning.

Even if the topic is included in the textbook it can not be assumed that every single standard or outcome in the topic is covered. A more detailed content analysis of the textbooks can be made to check for the presence of every single standard or outcome, although this can be very time-consuming. The extract from the Alabama Content analysis instrument in Figure 3.3 shows how the presence of every outcome is considered and not only the presence of the topics.

Chapter Topics	Book A	Book B	Book C	Book D	Book E	Book F	Book G	Book H
Adolescence and learning	x	x	x	x	x			x
Planning	x	x	x	x	x	x	x	x
Classroom management	x	x	x	x	x	x	x	x
Evaluation	x	x	x	x	x	x	x	x
Teaching styles	x	x	x	x	x			x
Motivation	x	x						x
Multicultural or disadvantaged students		x		x	x			x
History and aims		x						x
Audiovisual aids								
Teaching special pupils		x		x				x
Communication			x					x

Figure 3.2: Comparison of the content of different teacher training textbooks Henson (2004, 65)

Physical Science

Textbook__________________ Edition________ Publisher____________

Author __________________ Date of Publication ________ Series __________

Content Standards (Bullets and Examples)	Correlation (Yes/No)
1. Recognize periodic trends of elements, including the number of valence electrons, atomic size, and reactivity. • Categorizing elements as metals, nonmetals, metalloids, and noble gases • Differentiating between families and periods • Using atomic number and mass number to identify isotopes	
2. Identify solutions in terms of components, solubility, concentration, and conductivity. • Comparing saturated, unsaturated, and supersaturated solutions • Comparing characteristics of electrolytes and nonelectrolytes • Describing factors that affect solubility and rate of solution, including nature of solute and solvent, temperature, agitation, surface area, and pressure on gases	

Figure 3.3: Extract from the Alabama content analysis that questions not only the presence of topic in the textbook, but how well the topic is covered (Alabama Department of Education, 2009, 1)

Penney et al. (2003) also distinguish between methods used for content analysis that require analyst involvement and computer techniques. Computers are used in the process to save time and/or cost. The use of *computer programs* to determine the degree of alignment between curriculum and textbook does provide cheaper analysis, but also more superficial correlation analysis. Computers can check if the programmed "keywords" (topics and terminology prescribed by the curriculum) are present in the textbook (Tyson, 1997). This can lead to invalid results if the keywords are not chosen with care, since chapter titles may describe the contexts in which the ideas are presented rather than naming the ideas themselves (Goldsmith et al., 2000). For example, a chapter on "chemical change" can be titled "Burning birthday candles."

Textbook publishers often undertake the task of creating a *correlation analysis table*. These documents cross-reference the required content with the content of the textbook. The pages of citations indicate where in the textbook each required outcome of the topic is addressed. Unfortunately it gives no indication of how well the topic is covered (Tyson, 1997). Since publishers provide the data as part of their marketing of the textbook it is unrealistic to expect them to spend time describing the depth of coverage or the pedagogical approach, if it is not demanded by the majority of textbook buyers.

Even if a qualitative or non-frequency analysis shows that every topic and every single outcome is present in the textbook it can give no indication how well it is covered. More information is needed to get an indication of how well the content is covered. *Frequency counts* determine the number of times the topic (or outcome) is mentioned. It is assumed that a textbook that refers to a topic repeatedly is more likely to cover the topic thoroughly. The logic in this assessment becomes even more suspect when publishers know that this method is used. By adding the topic word repeatedly they can ensure good analysis results.

Textbooks can also be analysed to determine in *combinations of categories* are present. For example, the computer analyst can determine if the combination of a specific topic and a specific skill (for example 'density' and 'investigation') is mentioned in the same paragraph anywhere in the textbook. *Spatial analysis* measures the number of lines or pages devoted to a particular issue. The percentage of the whole devoted to an indicator can then be calculated. This renders an absolute as well as a relative measure of the importance the textbook attaches to the matter in question (Ninnes, 2002; Johnsen, 1993). The percentage of the total pages devoted to specific topics was used to evaluate and compare the coverage of science, technology and society standards in two American science textbooks (University of Alabama, n.d.).

2. For each content strand listed below, check off (4) whether or not it is covered in the program. Then rank how well the materials address that content:

1	2	3	4	5
Not well-addressed		Somewhat or satisfactorily addressed		Very well-addressed

Content	*Covered?*	*Ranking*				
Algebra	❑	1	2	3	4	5
Functions	❑	1	2	3	4	5
Geometry from a synthetic perspective	❑	1	2	3	4	5
Geometry from an Algebraic perspective	❑	1	2	3	4	5
Trigonometry	❑	1	2	3	4	5
Statistics	❑	1	2	3	4	5
Probability	❑	1	2	3	4	5
Discrete mathematics	❑	1	2	3	4	5
Conceptual underpinnings of calculus	❑	1	2	3	4	5
Mathematical structure	❑	1	2	3	4	5

Figure 3.4: Extract from the textbook evaluation instrument used by Waltham Public Schools, Waltham, MA, and Education Development Center, Newton, MA (Goldsmith et al., 2000)

Most textbooks contain the “topics” specified by the curriculum. The difference lies in the approach used to deal with it (Boostrom, 2001). Textbook analysis can not be limited to the presence of topics. It must also address the issue of how well the textbook covers the content. Waltham Public Schools, for example, considered every topic in a mathematics textbook and gave a rating of 1 to 5 for how well the textbook covered that topic. Figure 3.4 shows an extract of their evaluation instrument that shows the grid that is used to record the ratings.

Many researchers warn that the methods used to judge the content of textbooks and their alignment with syllabi are often ‘dysfunctional’ (Tyson, 1997). It does not give any indication whether the coverage is good or even adequate (Kesidou and Roseman, 2002; Stein et al., 2001). In the evaluation checklists and instruments the consideration of content has developed from a general judgement on “content” to “alignment” with prescribed standards or outcomes (Tyson, 1997). The AAAS expressed the view that a thorough examination of a material’s treatment of a few carefully selected learning goals would be more revealing than a superficial look at content alignment to many learning goals (Kulm et al., 1999). For example, the AAAS has developed the following questions to aid assessment of the content (American Association for the Advancement of Science, n.d.a):

- Does the content address the substance of the specific benchmark(s) or only the benchmark's general topic?
- Does the content reflect the level of sophistication of the specific benchmark or are the activities more appropriate for targeting benchmarks at an earlier or later grade?
- Does the content address all parts of the learning goal?

The New Mexico State University College of Education reported an evaluation of the content of mathematics textbooks in which the AAAS content criteria mentioned above were used to evaluate both the presence of content and the depth of coverage. The results are represented in the chart in Figure 3.5.

The table of contents can provide more information than just the presence of a subject in a textbook. Chambliss (1994) evaluated the tables of content to get an indication of the structure of the textbook and the presentation of the content. Textbooks that are poorly structured and would confuse students are thus identified and not investigated any further.

3.9.2 Readability assessment

Readability is an obvious characteristic of a good quality textbook, since learners can only benefit from the use of textbooks that they can read and understand (Allington, 2002). Legibility ("Can you make out the word?") must not be confused with readability ("Can you understand what the author is saying?") (Schultz, 1997). Modern printing is of such consistently high quality that concerns about legibility focus only on the artwork, like complex and confusing diagrams. The readability of science textbooks are, however, important and readability measurement has been the focus of research since the beginning of the 20th century (Stahl, 2003). Textbooks can be tested for usability using experimental measures in the classroom with the intended users. In Section 3.7.1 a number of experimental measures of text comprehensibility were mentioned. In textbook analysis where the textbook is considered in isolation *readability formulæ* are applied as objective and quantitative approaches to assess the readability of a textbook (Fry, 2002; Armbuster, Osborn and Davidson, 1985).

The theoretical framework within which readability is considered is the research on reading and reading comprehension and subsequent development of models of text comprehension, such as Kintsch's Comprehension Integration (CI) model of text comprehension (Kintsch and Kintsch, 2005; Best et al., 2005; Boscolo and

Archived Textbook Evaluations Summary Chart

This chart provides comparisons on how the textbooks scored on content for detailed information charts on the *instructional quality* of each textbook, click on the textbook name. The instructional quality charts had too many good categories for a single comparison chart such as the one below. There is a nice visual rating chart similar to the one below for each of the texts.

Content Scale Most content Partial content Minimal content

	Benchmarks					
	Number Concepts	Number Skills	Geometry Concepts	Geometry Skills	Algebra Graph Concepts	Algebra Equation Concepts
Connected Mathematics	Partial	Most	Partial	Most	Most	Most
Heath Passport	Minimal	Most	Most	Most	Partial	Partial
Heath Mathematics Connections	Partial	Most	Minimal	Most	Partial	Partial
Math Advantage	Minimal	Most	Minimal	Partial	Partial	Minimal
Mathematics: Applications and Connections	Partial	Most	Partial	Most	Minimal	Minimal
Mathematics in Context	Most	Most	Partial	Most	Most	Most
Mathematics Plus	Most	Most	Partial	Most	Minimal	Partial
MathScape	Partial	Most	Partial	Most	Minimal	Most
Middle Grades Math Thematics	Partial	Most	Partial	Most	Partial	Most
Middle School Math	Most	Most	Partial	Most	Partial	Most
Middle Grades Math	Most	Most	Partial	Most	Partial	Most
Saxon Math 65, Math 76, and Math 87: An Incremental Development	Partial	Most	Minimal	Most	Minimal	Minimal
Transition Mathematics	Most	Most	Partial	Most	Partial	Most

Figure 3.5: Example of a Textbook Evaluations Summary Chart (College of Education: New Mexico State University, n.d.)

Mason, 2003; Broer et al., 2002; Clark, 2001; Iding, 2000) that was discussed in Section 2.3.1. The first level of processing is processing text structure (words, sentences and linguistic relationships between them) (Best et al., 2005; Iding, 2000; Broer et al., 2002). Any difficulty with words, (semantic difficulty) and sentences (syntactic difficulty) can impair the first level of processing (Fry, 2002; School Renaissance Institute, 2000). Research has confirmed the influence of vocabulary difficulty on text comprehension (Mikk, 2000). Stahl (2003) found a correlation coefficient of more than 0,90 between measures of vocabulary difficulty and reading comprehension.

The reading level of a textbook can have a great influence on a student's understanding and learning. Consider a student reading a book at his "instructional reading level" (95% accurate). He will skip or misread as many as 5 words in every hundred. That will be 10–25 words on every page of a high school science text. The words he misread will obviously be the unfamiliar or technical vocabulary (Allington, 2002). This is especially true in science textbooks that are usually more difficult to read than narrative texts (Hsu and Yang, 2007; Galili and Tseitlin, 2003; Freeman and Person, 1998). Physics redefines words in everyday use, provides them with a specific meaning, making them inflexible and far less obvious to the reader (Galili and Tseitlin, 2003). Physics textbooks

> ... contain no unnecessary information; each word is important and has its precise meaning. Thus, students can not read *most* of the words and get the information; they have to read *all* the words (Koch, 2001, 759).

Johnson (n.d., 9) mentions three further reasons why science textbooks should be below the reader's reading level, rather than above:

- the numerical parts of the text increase the processing level
- the student's motivation may not be the same as his or her motivation in texts in other subjects
- the textbooks in science education are often used without teacher support, for example with homework

Readability formulæ can target the textbook's syntactic difficulty, semantic difficulty or both. The *syntactic difficulty* or the complexity of the sentences can be measured in terms of the average number of words per sentence (number of words ÷ number of sentences) (Burns, 2006; Stahl, 2003; Fry, 2002; School Renaissance Institute, 2000). The *semantic difficulty* or difficulty of text vocabulary can be determined by

- average number of characters per word (Stahl, 2003)

- average number of syllables per word (number of syllables ÷ number of words)(Stahl, 2003; Johnson, n.d.)
- percentage "easy" words (number of words in the text that occur on a specific list of "easy" words ÷ total number of words × 100) (Stahl, 2003).

3.9.2.1 Traditional readability formulæ

Many different readability formulæ have been developed that incorporate the measurements discussed in the previous section. A number of the examples of readability formulæ are analysed in Table 3.1.

A more detailed discussion of the formulæ can be found in Appendix B. The different formulæ reveal how different researchers view the relative importance of the semantic and syntactical difficulty to the overall readability. In the Dale-Chall and the Flesch-Kinkaid formulæ the semantic difficulty is given much more weight than the syntactical difficulty, but the Gunning FOG readability test affords them equal weights.

3.9.2.2 Limitations of traditional formulæ

Traditional readability formulæ have definite limitations and problems associated with their interpretation (Armbuster et al., 1985; Chambliss, 1994). While the counting of words and sentences and applying readability formulæ give *reliable* results, the results are not necessarily a *valid* measure of the readability (Allington, 2002; Armbuster et al., 1985). This can be explained by the second and third levels of processing required for comprehension and integration of the text. The second level of Kintsch's Comprehension Integration (CI) model of text comprehension is called the *textbase* (Kintsch and Kintsch, 2005; Boscolo and Mason, 2003). On this level the reader must integrate individual sentences into a coherent text-level representation (Best et al., 2005).

The text does not provide all the information that the reader requires in order to understand it, the reader has to "fill in the information gaps" (Kintsch and Kintsch, 2005; Best et al., 2005; Boscolo and Mason, 2003; Walsh, 2006). The text can consist of independent clauses without enough coordinating conjunctions. If there are too many gaps, or the gaps are too large for the readers to bridge, comprehension of the text will be limited or impossible. Best et al. (2005) identify this as a common problem in science textbooks. According to Boscolo and Mason (2003) there are two main variables that contribute to the process

Sample size	Semantic difficulty	Syntactical difficulty	Reading level
Dale-Chall Readability Formula (School Renaissance Institute, 2000; Dale and Chall, 1948)			
1 passage with 100 words for every 50 pages in textbook	% familiar words (F) Dale-Chall list: 3000 words familiar to fourth graders	average words per sentence (L)	Raw Score = $0,0496L + 0,1579F + 3,6365$ Use Dale-Chall Table to convert it to grade level.
Gunning FOG Readability Test (Johnson, n.d.)			
3 passages of 100 words each	% words with three or more syllables (N)	average words per sentence (L)	$(L+N)\times 0,4$
Fry Readability Graph (Mikk, 2000)			
three 100 word samples	average number of syllables per 100-word sample	average number of sentences (y)	use x and y to read the reading age (in years) from the Fry graph
Flesch Reading Ease Test (Giles and Still, 2005)			
systematically selected 100-word samples	syllables per 100 words (*wl*)	average number of words per sentence (*sl*)	Reading Ease = $206,835 - 0,846wl - 1,015sl$ Use Flesch's scale to get the difficulty classification
Flesch-Kincaid Formula (Giles and Still, 2005; Johnson, n.d.) developed by USA Department of Defense and used in Microsoft Windows and Corel Wordperfect			
	average number of syllables per word (N)	average sentence length (L)	grade level = 0,39L + 11,8N -15.59
'SMOG' Formula of McLaughlin (Johnson, n.d.; Newton, 1990)			
select samples of 30 consecutive sentences	average number of words with 3 or more syllables (N) in 30 consecutive sentences	-	grade level = $\sqrt{N} + 3$, Reading Age = $\sqrt{N} + 8$ years
Lexile Framework developed by Metametrics (commercial) (Fry, 2002; School Renaissance Institute, 2000)			
samples of 150 words	percentage of the words that occur on the American Heritage Intermediate Corpus	average number of words per sentence	not available (closed commercial standard)
ATOS (Advantage-TASA Open Standard) developed by the School Renaissance Institute (Burns, 2006; Fry, 2002; School Renaissance Institute, 2000)			
the entire text is analysed by computer	average grade level of words and characters per word	words per sentence	

Table 3.1: Examples of readability formulæ

of generating connections by inferring or filling in the gaps. The first is text coherence and the second is the reader's contribution (prior knowledge and knowledge of text structure). *Text coherence* refers to the sentences and how parts of the text are connected in order to facilitate the forming of a coherent text representation (Boscolo and Mason, 2003). Although research on readability formulæ suggests that using less complex sentences can increase readability and learning (Britton et al., 1993), some techniques for shortening sentences can have disastrous effects. If sentences are shortened by dividing separate clauses and deleting connectives it may have negative effects since readers must infer the missing connectives (Kintsch and Kintsch, 2005; Armbuster et al., 1985). Coherent texts provide more of the connections needed to form the textbase and therefore require less interference activity from the readers (Best et al., 2005; Boscolo and Mason, 2003). Consequently, the coherence of the text influences its readability and ultimately its quality as learning support material (Newton et al., 2002; Beck and McKeown, 2001; McNamara et al., 1996). Measuring the coherence of text can therefore contribute to the assessment of the ability of the text to support the learning process especially in science textbooks that are typically "low-cohesion" (Best et al., 2005). At a local level cohesion is maintained by using explicit connections between clauses, rather than requiring the reader to infer the connection. On a global level "coherent texts are well organised and their structure is clearly signaled (e.g., by using section headings, topic introducers, order markers, and the like)" (Kintsch and Kintsch, 2005, 85).

Prior knowledge is the second variable that influences construction of the textbase. The experts involved in writing the text omit information that they assume their target learners possess as prior knowledge (Best et al., 2005). If the learner's prior knowledge is inaccurate it can interfere with the processing of the text. The text must anticipate and moderate readers' lack of sufficient and accurate prior knowledge (Kendeou and Van den Broek, 2005).

The third and final level of text processing is the situation process: a processing of the textbase under the influence of the prior knowledge of the reader (Iding, 2000; Broer et al., 2002). Examples of text characteristics that contribute to this level of processing are the genre of text (Kearsey and Turner, 1999), and the use of explanatory principles (plausibility, parsimony, generalisability, fruitfulness). Recent textbook evaluations tend to include many factors that aim to determine *learnability* and *text effectiveness* rather than *readability* (Boscolo and Mason, 2003; Mikk, 2000). Readability assessment have

> . . . moved from the traditional, largely age-related requirements for the simplification of vocabulary and syntax toward an increasingly psycho-

and sociolinguistic recognition of the scope of the problem of *communication* in textbook texts (Johnsen, 1993, 215).

3.9.2.3 Readability of the textbook and reading ability of the learner

The challenge in assessing readability is not only to determine how difficult a textbook is to read, but to match the learner's reading ability with the textbook reading difficulty. According to Freeman and Person (1998, 12) "textbooks have simultaneously been criticised for being both too easy en too difficult." Texts that are easy for a reader to process reduce the amount of active processing. Although this may increase direct recall of the text, it may lead to less effective learning (Kintsch and Kintsch, 2005; Boscolo and Mason, 2003; McNamara et al., 1996). Textbooks should be chosen according to learners' zone of proximal development (ZPD), within the range of challenge in which maximum growth can occur (School Renaissance Institute, 2000). It must be challenging enough to engage the reader, but not so difficult that readers find it frustrating (Clark, 2001; School Renaissance Institute, 2000; Mikk, 2000; Chall and Conard, 1991). The level of the text must be matched to the level of the reader's ability (Burns, 2006; Fry, 2002). It is necessary to keep in mind the fact that reader characteristics like motivation, interest, purpose and perseverance in the reading test situation may differ from the same reader's characteristics when reading science texts (Guthrie and Wigfield, 2005; Armbuster et al., 1985). In this research we are only concerned with assessing textbook quality. Although reader characteristics are not measured during textbook analysis, the writers and analysts should anticipate the audience (Giles and Still, 2005).

3.9.3 Complex analysis systems

Some analysis instruments consist of a basic list of desired characteristics that can be checked off, but more elaborate instruments have been developed that include a number of elements. Ansary and Babaii (2002) suggest that the following elements should form part of any system for textbook evaluation:

- a predetermined preferred set of characteristics of textbook
- a system within which one may ensure objective, quantified assessment
- a rating method that can provide the possibility for a comparative analysis
- a simple procedure for recording and reporting the evaluator's opinion
- a mechanism by which the universal scheme may be adapted and/or weighted to suit the particular requirements of any teaching situation

- a rating trajectory that makes possible a quick and easy display of the judgements on each and every criterion,
- a graphic representation to provide a visual comparison between the evaluator's preferred choices as an archetype and their actual realization in a particular textbook

The remainder of this subsection will be dedicated to descriptions and explanations of the various elements of analysis instruments. Relevant examples will be mentioned. Procedures for the development or designing of the elements and whole instruments will be discussed in the next section.

3.9.3.1 Criteria or desirable characteristics

Evaluation often depends on context and even on preferences (Scriven, 2003, 29) and therefore there is no general consensus on evaluation criteria for textbooks or teaching media (Ansary and Babaii, 2002; Johnsen, 1993). For example, teachers use textbooks in different ways and therefore their priorities in textbook characteristics will vary and no already-available checklist of criteria can be used to judge teaching material as long as the specific requirements in a specific teaching situation are not identified (Ansary and Babaii, 2002). At best a group of teachers in similar contexts (e.g. same country, educational system, and subject) can use the same evaluation criteria. Consequently, textbook evaluation does not start with an investigation of the textbooks. According to Montgomery (2006) one of the biggest mistakes analysts make in conducting an evaluation is to start the evaluation before they have adequately discussed their instructional priorities and formulated the evaluation criteria. The determination of the criteria is exactly the core problem in the evaluation of textbooks mentioned in Section 3.7.2.3 (Mikk, 2000; Van de Grift, 1989; Franssen, 1989b). Montgomery (2006) recommends that the following questions be used to guide the discussion on instructional priorities:

- How will the textbook be used?
- What are the curricular goals we hope this material will help us meet?
- How important are ancillary materials, or should we focus only on the core materials?
- What sort of instructional approach would best fit with our curricular goals?
- Will the book be the primary source of content for the course, or will content come from elsewhere?
- Does the district already have common assessments that we hope this book will help our students to meet, or do we expect that the textbook will provide

assessments for us?

There are many aspects of a textbook that can influence its success in supporting the teacher and learner, and researchers have identified many of the factors that can add or detract from the learning process (Johnsen, 1993). These factors can be used to formulate criteria for textbook analysis. To consider *all* the relevant factors it is necessary to create a list of criteria or desirable characteristics, prior to the textbook analysis. This list can then be consulted when the textbooks are evaluated (Wingate, 2002). It is important to reach consensus on the evaluation criteria before the actual analysis of the textbook to ensure an objective approach to the task.

Some criteria may be so vital that any textbook that does not meet that criteria can not even be considered as a candidate for selection. Stein et al. (2001) refer to these criteria as *screening criteria*. When numerous textbooks are considered the use of screening criteria can help reduce the number of textbooks that have to be thoroughly analysed. The following have been used as screening criteria:

- Content is usually the screening criterion. The Mozambique Ministry of Education only considers textbooks for approval if it covers at least 70% of the syllabus.
- Durability standards can be the screening criterion. For example, the National Institute for Educational Development of the Namibian Ministry of Education (2005, 2) sets minimum standards for durability as follows:

  Paper:
  All textbooks should be printed on Bond 70 or 80 gm paper. Super 60 (60 gm) should not be used.
  Binding:
  Textbooks/workbooks up to 120 pages – saddle-stitched in the centre of the pages.
  More than 120 pages – only thread-sewn.
  No books with *perfect binding* (pages glued) should be submitted.
- The publication date is sometimes considered as a screening criterion (Ade-Ridder, 1989). The cost of new textbooks forces some education departments to limit the acquisition of new textbooks and use textbooks for a number of years. The National Institute for Educational Development of the Namibian Ministry of Education (2005), for example, prescribes that textbooks should be used at least five years before a change to a new textbook can be considered. The use of publication date as screening criterion for science textbooks is justified by the fact that scientific knowledge growth is swift, textbooks can quickly become outdated and the fact that research

on learning continuously provides new information on learning that must be considered in the teaching approach in the textbook. On the other hand, Hubisz (2003, 6) argues that

> . . . a middle school science course should not be about the latest scientific discoveries or complex subjects like black holes, the quantum theory, or the theory of relativity. The latest scientific findings cannot be meaningful to students with little or no background. In order to appear current, all of the books . . . devote valuable space to these issues.

Some criteria can act both as screening and general criteria. For example, coverage of the content topics can act as screening criterion, while the depth of coverage can be considered a general criterion. If all topics specified by the syllabus are not present, the textbook is removed from the list of textbooks that are considered for selection. The textbooks that cover all the topics are then analysed to assess how well they cover the content.

The number of criteria can range from a few to as many as 180 (Chambliss, 1994) and they are usually grouped in categories and even in subcategories to make it easier to handle (Ogan-Bekiroglu, 2007). For example, the instrument used by the National Department of Education in South Africa has 40 criteria in 5 categories and the instrument used by the Gauteng Department of Education has 51 criteria in 12 categories (see instruments in Appendix B). When irrelevant criteria are included in the list it will escalate the time-consuming nature textbook of analysis (as discussed in Section 3.7.2.3). Therefore, it would be prudent to limit the criteria to important aspects that have a marked influence on the quality of the textbook, but at the same time all important aspects must be included. The criteria should include both internal criteria (which are subject related) and external criteria (which provides a broader view of the book) (Ansary and Babaii, 2002) and should be "objective" and "verifiable" (Stein et al., 2001, 16). Even the weight of the textbook can be a criterion in determining the textbook's usability (Dake, 2007).

A clearer picture of a set of criteria and the categories and subcategories can be obtained when the hierarchical structure is presented as a tree. For example, the criteria of the MEAS (Media Assessment) group (Astleitner, n.d.) for the evaluation of educational media in general can be represented in the diagram in Figure 3.6. The MEAS group did not publish any explanation for the inconsistent use of question marks with the criteria. The MEAS group assessed the quality of both printed and electronic media on a commercial basis under the leadership of Hermann Austleitener of Salzburg.

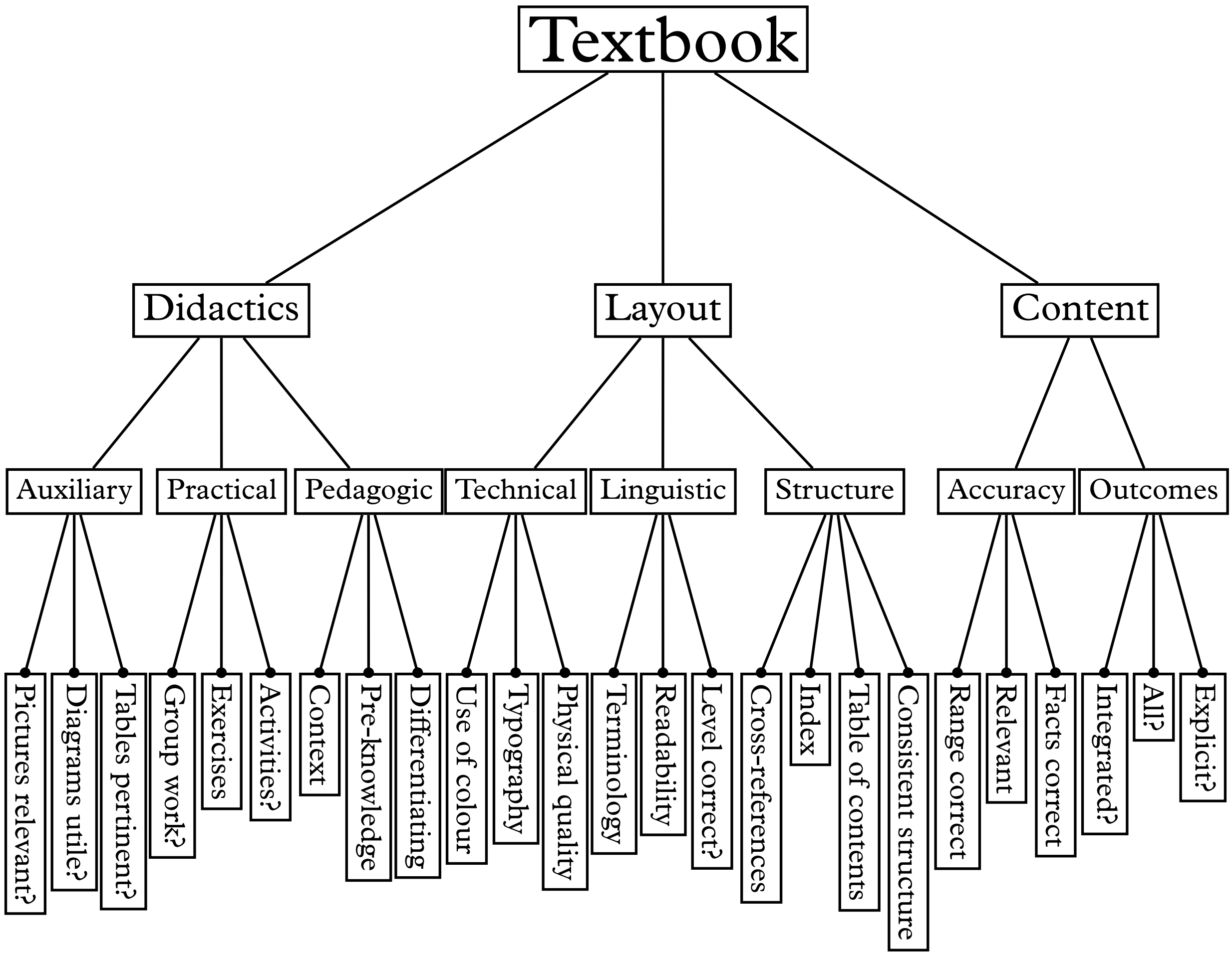
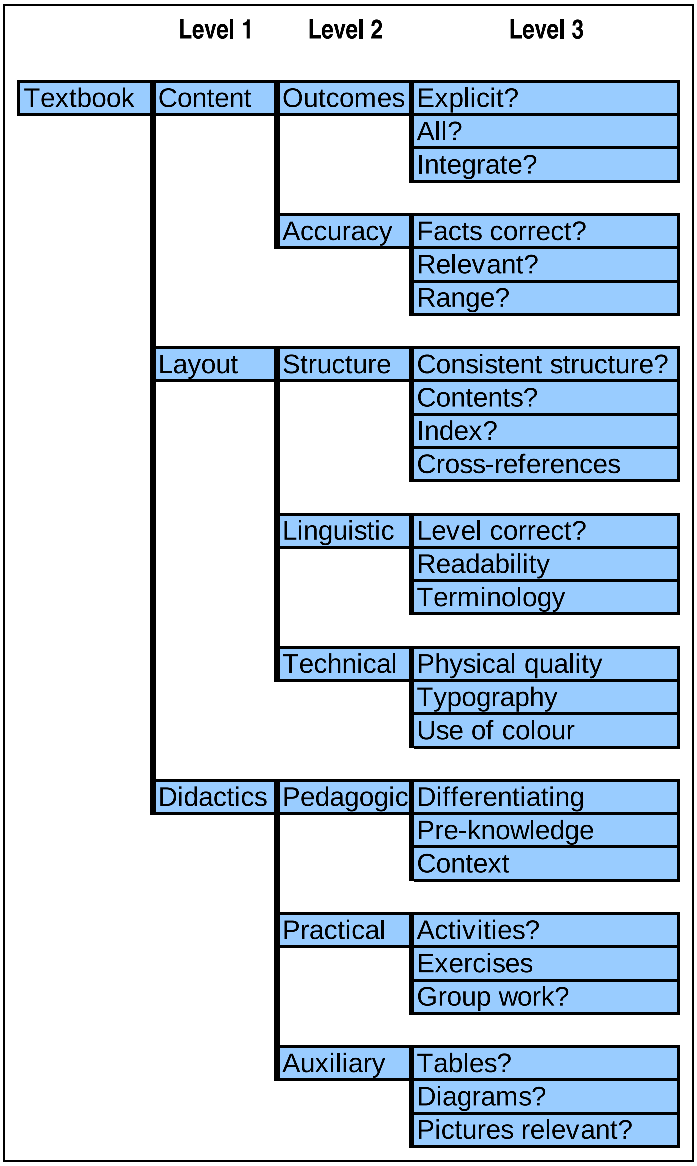


Figure 3.6: The criteria used by the MEAS group (Astleitner, n.d.) represented as a hierarchy with levels represented horizontally

The hierarchy of criteria used by the MEAS group is transformed and represented with horizontal levels in Figure 3.6. If the evaluation instrument contains more criteria on the lower level this representation becomes difficult to fit on a page. It is easier to represent the levels from left to right, as in Figure 3.7.

Some of the existing instruments for the evaluation of science textbooks will be discussed in Section 3.9.4 and their criteria end categories will be compared. Valid data about the quality of the textbook can only be obtained by the use of criteria based on thorough research on the characteristics of the textbook that contribute to learner achievement of the science education goals. In this regard many researchers deplore the lack of available research-based criteria for evaluating instructional materials (Mozambique Ministry of Education, 2002; Stein et al., 2001; Goldsmith et al., 2000; Franssen, 1989b). Various approaches followed by researchers in addressing the problem of criteria formulation will be discussed in Section 3.10.1.



(The inconsistent use of question marks is maintained as published by the MEAS group)

Figure 3.7: The criteria and categories used by the MEAS (Astleitner, n.d.) to evaluate textbooks represented as a hierarchy with vertical representation of levels

Criterion 1. ***The program's learning goals are challenging, clear, and appropriate for the intended student population.***

Indicator a. The program's learning goals are explicit and clearly stated.

Indicator b. The program's learning goals are consistent with research on teaching and learning or with identified successful practices.

Indicator c. The program's learning goals foster the development of skills, knowledge, and understandings.

Indicator d. The program's learning goals can include important concepts within the subject area.

Indicator e. The program's learning goals can be met with appropriate hard work and persistence.

Figure 3.8: Extract from the evaluation criteria of the U.S. Department of Education Expert Panel on Mathematics and Science Education, with the relevant indicators

3.9.3.2 Indicators of the presence of characteristics

For every criterion to be considered evaluators need to know exactly what they should look for in the textbook. *Indicators* can be provided that "specify what constitutes evidence for meeting each criterion" (Kesidou and Roseman, 2003, 536). Specifying the indicators gives a more objective approach to the analysis and contributes to the reliability, because it enables different analysts to look for the same indicators in the textbook to determine if it exhibits a specific characteristic (Stein et al., 2001).

Both the instruments used by the AAAS and the U.S. Department of Education Expert Panel on Mathematics and Science Education address this issue in detail. They prescribe how the presence of a specific characteristic can be determined or measured by providing indicators that can be easily identified in textbooks. Figure 3.8 displays an example from the analysis instrument of the Expert Panel and an example from the AAAS instrument is presented in Figure 3.11.

3.9.3.3 Rating scheme for criteria

Some of the checklists have criteria formulated as questions that simply require a "Yes" or "No" answer. In this case the percentage of positive answers gives an overall picture of the textbook quality. The evaluation instrument of Figure 3.9 is one such example.

CRITERIA	**RATING**		
I. ADDRESSING THE GOALS OF ELEMENTARY SCIENCE TEACHING AND LEARNING			
Does the material focus on concrete experiences by the children with science phenomena? Reason:	Yes	No	NA
Does the material enable children to investigate important science concept(s) in depth over an extended period of time (core materials only)? Reason:	Yes	No	NA
Does the material contribute to the development of scientific reasoning and problem-solving skills? Reason:	Yes	No	NA
Does the material stimulate student interest and relate to their daily lives? Reason:	Yes	No	NA
Does the material allow for or encourage the development of scientific attitudes and habits of mind, such as curiosity, respect for evidence, flexibility, and sensitivity to living things? Reason:	Yes	No	NA
Are assessment strategies aligned with the goals for instruction? Reason:	Yes	No	NA
Will the suggested assessment strategies provide an effective means of assessing Yes No NA student learning? Reason:	Yes	No	NA

Figure 3.9: Example of an evaluation instrument that requires “Yes” or “No” answers (National Science Resources Center of the National Academy of Sciences, 1998, 412)

In a more complex assessment consideration is not limited to an investigation whether the characteristic is present or absent in a textbook. The textbook is also studied to determine how well the characteristic is exhibited in the textbook. This is the element of textbook analysis or evaluation where measurement and comparison enter into the process. Both qualitative and quantitative information can be involved. How many activities are provided per topic or what percentage of the activities is learner-centered or inquiry based is an example of a quantitative measurement. The reading level that is calculated for a textbook is another example of a quantitative measurement.

A rating scheme makes it possible to quantify the quality of different aspects of the textbook and produce comparable numeric values. A typical rating procedure is the use of a five-point Likert-type scale ranging from zero to four. The score awarded to the textbook reflects the level of agreement with the criterion or requirement stated in the list (Olivares-Cuhat, 2002). The Oklahoma evaluation instrument, represented in Figure 3.15, is an example of a five-point Likert-type rating *scale*. Alternatively, terms like "outstanding," "good," et cetera are used. Consider the extract of the Alabama instrument in Figure 3.13 and the New Bruinswick instrument in Figure 3.14. Unfortunately this type of rating scores does not reflect the thinking of the analyst and the reasoning used to decide on the rating for the textbook (Maree et al., 2006). The Namibian Department of Education addressed this problem by providing space on their scoring sheets (see Figure 3.10) where analysts can record their justification for the rating (National Institute for Educational Development of the Namibian Ministry of Education, 2005). Figure 3.10 shows some of the criteria, as well as the indicators that are provided. Unfortunately this does not solve the problem of how to decide which rating on the scale is appropriate.

A more complex textbook evaluation procedure guides the analysts in their awarding of rates. In the previous section it was described how the AAAS *rating schemes* provide indicators whose presence in the textbook enables the textbook to satisfy the criterion. The presence of more of these indicators suggests that the textbook is better at satisfying the criterion. The AAAS used the terms "excellent," "satisfactory" or "poor" in their rating scheme. A rating of "excellent" is the benchmark of what a perfect textbook should look like. Consider, as an example, the AAAS rating scheme for one of their criteria in Figure 3.11.

The indicators can be incorporated in a *rubric* to assist the rating of textbook quality for every characteristic (Chambliss, 1994). A scoring rubric is a qualitative and descriptive scoring guideline that describes the characteristics of different levels of performance (Kan, 2007; Singer and Tuomi, 2003). Scoring rubric scales are

2.	CONTENT (Please note 2.1 to 2.5 are counting out of 10)	
2.1	How consistent is the approach used in the book with the syllabus? : • Does the book integrate relevant learner activities and teacher demonstrations? ………………………………………………………………………………………… … ………………………………………………………………………………………… ………………………………………………………………………………………… ………………………………………………………………………………………… ……………………………………………………………………(10 Fully to 1 Scarcely)	
2.2	To what extent are relevant knowledge objectives catered for in the book? • Does the book cover competencies (syllabus topics and objectives)? • Does the book follow a policy accepted pattern (LCA e.g. social constructivist approach/co-operative learning)? ……………………………………………………………………………………… ………………………………………………………………………………………… ………………………………………………………………………………………… ………………………………………………………………………………………… ……………………………………………………………………(10 Fully to 1 Scarcely)	

Figure 3.10: Extract from the Namibian Instrument for Textbook Review (National Institute for Educational Development of the Namibian Ministry of Education, 2005, 14) that shows indicators and a basic rating scheme

usually organised in a table that consists of vertical rows and horizontal columns. Every row represents a criterion where the first cell identifies the criterion and the other cells contain detailed descriptions of indicators that should be present in order to receive the various ratings or scores (Kan, 2007; Council of State Science Supervisors, n.d.). The Council of State Science Supervisors in the USA developed such a rubric for evaluating materials' ability to stimulate classroom inquiry (Council of State Science Supervisors, n.d.). Figure 3.12 shows an extract from a rubric used by Virginia to evaluate history and social science textbooks.

3.9.3.4 Procedure for recording and reporting the rates

Some textbook analysis instruments consist of a list of aspects that must be considered or questions that must be answered during an analysis, but most instruments indicate how analysts should record their individual judgements (National Institute for Educational Development of the Namibian Ministry of Education, 2005). This can be done in long written reports, but most textbook analysis instruments are accompanied by a specific score sheet where the ratings are recorded. The ratings can be entered as numbers; it can be circled when the possible values are printed on the sheet (Litz, 2001) or a cross can be made in a column representing the rating (Crystal Springs Books, n.d.). Figure 3.13 shows an example of such a score sheet.

Some score sheets are designed to accommodate the results only one book. In such cases analysts will have to consider all the different score sheets in order to

In the AAAS category: *Attending to prerequisite knowledge and skills*,
one of the criteria is

Does the material specify prerequisite knowledge/skills that are necessary to the learning of the key ideas?

The AAAS provides the following indicators of meeting this criterion:

1. The material alerts the teacher to specific prerequisite ideas or skills (versus stating only prerequisite topics or terms).
2. The material alerts teachers to the specific ideas for which the prerequisites are needed.
3. The material alerts students to prerequisite ideas or experiences that are being assumed.
4. The material adequately addresses (provides instructional support for) prerequisites in the same unit or in earlier units (in the same or other grades). (The material should not be held accountable for addressing prerequisites from an earlier grade range. However, if a material does address such prerequisites they should count as evidence for this indicator.)
5. The material makes adequate connections (provides instructional support for connections) between ideas treated in a particular unit and their prerequisites (even if the prerequisites are addressed elsewhere).

Rating Scheme:

- Excellent: The material meets indicators 1, 2, 3 or 4, and 5 for all or most prerequisites.
- Satisfactory: The material meets indicators 1, 2, 5, and either 3 or 4 for some prerequisites.
- Fair: The material meets indicators 5 and either 3 or 4 for some prerequisites.
- Poor: The material meets no more than one indicator.

Figure 3.11: Extract from the AAAS instrument that includes indicators and a rating scheme

History and Social Science
Textbook and Instructional Materials Review Rubric
Section II - Design and Organization of Instructional Materials

Adequate **A**	**Limited** **L**	**No Evidence** **O** (Note: Provide examples to support this rating.)
Criterion 1 - Materials support the goals of the History and Social Science Standards of Learning as outlined in the introduction to the 2001 History and Social Science Standards of Learning.		
– Objectives and materials are aligned with the standards. – Sufficient coverage of skills and concepts is evident.	– Limited connections between the standards and the materials are noted. – Limited support of skills and concepts is evident.	– There is no correlation between the materials and the standards.
Criterion 2 - Instructional materials reflect a coherent and logical sequence of instruction.		
– Materials provide for an orderly, well-planned, logical progression of facts and concepts with attention to the goals of the standards.	– Sequence of content and materials is inconsistent.	– No sequence of content and materials was identified.
Criterion 3 - Reading level is appropriate to the age and/or grade level of students.		
– Reading level is generally appropriate for the intended student population. – Reading level remains consistent throughout the text.	– Reading level is often not on grade level. – Reading level varies throughout the text.	– Reading level is not appropriate.
Criterion 4 - Organizational properties such as the table of contents, indices, and glossaries are well designed and useful.		
– Organizational properties of the materials assist in understanding and processing the content. – Organizational properties are generally easy to use.	– Organizational format is inconsistent or not well-defined. – Organization of properties are unclear or confusing.	– Organizational properties are insufficient or not present.

Figure 3.12: Extract from the rubric for History and Social Science Textbook and Instructional Materials Review used by Virginia

TEXTBOOK APPRAISAL FORM

NAME OF BOOK ______________________________ **COPYRIGHT** ________

PUBLISHER ______________________________ **GRADE** __________

CODE: 1 - Excellent, 2 - Good, 3 - Fair, 4 - Poor, 5 - Not Acceptable

I. PHYSICAL FEATURES

A. Construction - Quality of Paper and Binding	1	2	3	4	5
B. Attractiveness	1	2	3	4	5
C. Size and Shape	1	2	3	4	5
D. Table of Contents and Index	1	2	3	4	5

COMMENTS: ____

Figure 3.13: Extract from the score sheet used by Alabama for recording ratings during textbook analysis

compare the different textbooks. When the score sheet makes provision for the ratings of a number of textbooks, it establishes the possibility of direct comparison between the various textbooks, with regard to every criterion. The New Bruinswick Public Schools Textbook Evaluation Form in Figure 3.14 is an example of a score sheet that caters for a number of textbooks on a single sheet.

A score sheet for recording and reporting rates is valuable to all the stakeholders, for a number of reasons:

- All stakeholder benefit from the use of score sheets because the procedure facilitates *adherence to the evaluation system* and, therefore, contributes to thorough analysis of the textbook. Just like using a textbook in class does not ensure that the teacher use the textbook as intended, textbook analysis instruments are not always used as intended. With a time-consuming task like textbook analysis, the temptation is always there to do a superficial scanning of the book, instead of considering every question or aspect individually. When the analysts are required to put their judgements on individual questions or aspects in writing, they are coerced into considering these individually. Score sheets that provide space for analysts to cite page numbers and other references as explicit evidence to support their judgement adds further "rigor and reliability" to the decisions (American Association for the Advancement of Science, n.d.a). Citing actual page numbers, as evidence for a judgement, also helps "reduce evaluator bias" (Stein et al., 2001, 16).

RATING PROCEDURES	SERIES IDENTIFICATION
After each point has been carefully studied record your judgment in the space provided using the following scale: 3= Outstanding 2= Good 1= Fair 0= Unsatisfactory	A. B. C. D.

PHYSICAL CHARACTERISTICS	Series A	Series B	Series C	Series D	Series E
Size and type, style of print, and the binding provide for easy reading and use.					
Illustrations are related to the maturity levels of the readers.					
DEVELOPMENT Authorship included competent writers in the filed.					
AFFIRMATIVE ACTION The material avoids stereotypic treatment of women and racial, religious and cultural minorities.					
The material includes broad presentation of ethnic groups in a variety of roles both in illustration and in content.					
The material portrays senior citizens as active participants in the roles shown in photographs.					
TEACHER'S EDITION & ANCILLARIES Appropriate learning activities are suggested which includes use of technology.					

Figure 3.14: Extract from The New Bruinswick Public Schools Textbook Evaluation Form that can be used to record and compare the ratings of five textbooks (New Bruinswick Public Schools, n.d.)

- All stakeholders can benefit from score sheets that allow them to view the individual ratings. It can act as *verifiable justification* for the final analysis results (Stein et al., 2001). It is even possible to show all ratings given by individual analysts, as well as the final aggregated ratings. For example, the ratings of five analysts are displayed in the Textbook Evaluation Tool, a self-contained desktop application, developed by commercial, digital designers from Imuse. (Imuse: digital design company, n.d.) Unfortunately, detailed reports of textbook evaluations are often not available. For example, the NIED of Namibia (National Institute for Educational Development of the Namibian Ministry of Education, 2005) considers completed evaluation instrument sheets as confidential documents, that are filed for at least two years, to be referred to in cases of enquiry.
- The analysts benefit from the use of score sheets because it saves them the *time* a long written report would have taken. Some score sheets even include the criteria next to the position where the rating should be recorded. Examples are the Textbook Evaluation Form used by Crystal Springs Books (Crystal Springs Books, n.d.) and the Standards-Based Textbook Evaluation Guide developed by the Indiana University, for the evaluation of Foreign Language Textbooks (Cisar, n.d.). The analyst does not even have to refer to a separate list when analysing the book. An analyst who has used the score sheet for a number of analyses can become accustomed to the score sheet, which will further decrease the time per analysis.
- The analysts also benefit because the use of score sheets can help "*maintain a professional atmosphere* during what can sometimes prove to be a stressful process" (Stein et al., 2001, 16).
- Some score sheets can be directly used to *represent the results* of the analysis. Score sheets can provide so much more information than just final aggregated results that rank textbooks according to overall quality. It could help teachers to identify weaknesses in the approved or chosen textbooks and provide publishers with information on how to improve the quality of the textbooks they publish.

3.9.3.5 Relative importance of criteria or weights

When every question or criterion receives the same marks or attention, the proportions of questions or criteria in a textbook analysis instrument that focuses on different aspects are important. It may cause "inconsistency in criteria" if it is not considered thoroughly (Johnsen, 1993). Some aspects of quality in textbooks can inadvertently be emphasised at the expense of others (Mikk, 2000; Johnsen,

Oklahoma State Textbook Committee
Evaluation Form for Mathematics K-12
Please circle one for each of the following criteria: 1=strongly disagree 2=disagree 3=no opinion 4=agree 5=strongly agree

Committee member ____ Advisor____

Publisher____ Title of Bid Item ___________________________Grade ___________

Type of Bid Item: Comprehensive, Ancillary or Supplemental (Circle One)

A. Criteria for Instructional Materials

1. Content in the instructional materials align well with Oklahoma's Core Curriculum, *Priority Academic Student Skills (PASS).* **1 2 3 4 5**
2. Materials are recently copyrighted and reflect current research for the best practices in mathematics. **1 2 3 4 5**
3. Materials are interesting, engaging, and effective for all students (e.g., ethnic, rural, urban, disabilities, English language learners). **1 2 3 4 5**
4. Concepts are logical and accurate. **1 2 3 4 5**
5. Instructional materials provide a logical progression for developing conceptual understanding of mathematics. **1 2 3 4 5**
6. Materials integrate the use of technology for teaching and learning mathematics. **1 2 3 4 5**
7. Instructional materials incorporate a variety of strategies and forms of assessment. **1 2 3 4 5**
8. Content integrates other subject areas and teacher's materials provide specific ideas for integration. **1 2 3 4 5**
9. Materials provide opportunities for students to explore and investigate mathematical content in an in-depth way. **1 2 3 4 5**
10. Materials provide opportunities for students to apply their understanding of the concepts. **1 2 3 4 5**
11. Materials give as much attention to problem solving, reasoning, communicating and application as is given to simply recalling information and performing computation. **1 2 3 4 5**

C. Readability, Graphics, Illustrations and Format

1. The readability of the materials is appropriate for the age, grade and maturity level of the students. **1 2 3 4 5**
2. Index, glossary and table of contents are complete and easy to use. **1 2 3 4 5**
3. Mathematics vocabulary is introduced or reinforced throughout the text. **1 2 3 4 5**
4. Graphics and illustrations help students understand materials and promote thinking and problem solving. **1 2 3 4 5**
5. Illustrations, charts, maps and graphs are conveniently located, and are clear and meaningful. **1 2 3 4 5**
6. Photographs and pictures help clarify the text. **1 2 3 4 5**
7. Binding is durable and soil resistant. Paper is good quality. Print is appropriate size, color and clarity to enhance readability. **1 2 3 4 5**

Section C total [ ]

D. Electronic-Based Materials (if applicable)

1. The means of response (i.e., mouse, single key-stroke, whole words or sentences, voice input) are grade level appropriate. **1 2 3 4 5**
2. Sound, when present, can be controlled. Graphics, color and special effects are motivational for users. **1 2 3 4 5**
3. Adequate spacing exists between lines of text. Letter size is

Figure 3.15: Extract from Oklahoma State Textbook Committee's score sheet that includes the wording of the criteria next to the position where the rating must be recorded (Oklahoma State Textbook Committee, n.d.a)

1993). For example, if most of the questions centre around the layout of the book, it will convey the idea that layout is more important than pedagogical approach. The evaluation criteria should not only incorporate the features or attributes that contribute to quality in textbooks, but also address the issue of how important the features are to the quality, in relation to each other (Pingel, 1999). To do this Montgomery (2006, 5) advises that

> criteria should be weighted in such a way so as to emphasise those aspects of academic content or instructional design that the district deems most important.

Other researchers like Bernier (Bernier, 1996) and Tucker (Ansary and Babaii, 2002) also advise and justify the use of weights and some of the available analysis instruments incorporate weights. The evaluation instrument for the Foundation and Intermediate Phase of the Gauteng Department of Education (Mahlaba, 2006) and the Namibian evaluation instrument (National Institute for Educational Development of the Namibian Ministry of Education, 2005) are examples (see Appendix B). Although researchers differ on the relative importance of criteria, most of them agree that textbooks are to be judged primarily in terms of how well they are likely to support student learning of important learning goals (Stein et al., 2001; Kesidou and Roseman, 2003).

Figure 3.16 shows a representation of the hierarchical structure of criteria as well as the weights used by Nogova and Huttova (2006) in the instrument they developed for the long-term project entitled "Textbook Policy in Slovakia". Since the total of the rating is equal to 100, the numbers after the categories and criteria are both the ratings and the percentages that the specific criterion or category contributes to the overall quality. This representation shows clearly that they consider content selection (18%) three times as important as social correctness (6%).

3.9.3.6 Representation of evaluation results

It can be problematic to determine how much information to include in representations of the textbook analysis results and how to represent these. The results of textbook assessments must be represented in a format that supplies all the information a teacher may need to select a book, identify a textbook's weaknesses or justify his or her choice of textbook. The contexts of teachers, learners and resources vary from class to class, therefore different teachers may value different aspects of textbooks. For example, an inexperienced teacher may value the proposed learning activities more than an experienced teacher and a teacher in an area with limited laboratory equipment would prefer a textbook with activities

	Level_1		Level_2	
Textbook	Compliance with PPD	17	Compliance with curriculumn	7
			Logical structure	3
			Compliance with course standerd	7
	Personal Development	18	Development of basic skills	7
			Student integration into community	5
			Usefulness in everyday life	6
	Content selection	18	Accuracy of information	8
			Conciseness of content	7
			Balance between elementary and additional content	3
	Methodological approach	24	Stylistic characteristics of text	7
			Content accessing methods	5
			Presentation of content	7
			Formulation and system of questions and tasks	5
	Graphic layout	17	Graphic representation of text	6
			Clear graphical presentation of content	4
			Matching pictorial material and text	3
			Quality of illustrations	4
	Social correctness	6	Respecting social correctness	6

Figure 3.16: Representation of the structure and weights (expressed as percentages) of the instrument developed by Nogova and Huttova (2006, 336–337)

with alternative learning activities that do not require laboratory equipment. In this regard Ade-Ridder (1989) even suggests that no attempt should be made to indicate which textbooks are overall superior to others. The results of textbook analysis should rather be represented in a way that enables teachers to determine the best fit for their needs. Providing the results for every category or even for every criterion provides the teachers with the optimum information on which to base their selection decisions. A more satisfying solution to the problem is to provide evaluation systems with mechanisms by which the weights can be adapted to suit the particular requirements of any teaching situation. The ratings can be done by experts and every teacher can adjust the weights to determine the optimal textbook for his or her class. This is in accord with the advice given by MacDonald (2006, 11) about evaluation data in general:

> Making our evidence or data more readily available to others may also provide opportunities for different conclusions to be reached or other questions to be addressed, thus making it of use to others in possibly very different contexts.

The specific pages, sentences, activities, et cetera in a textbook that prove that the textbook satisfies a criterion, are called sightings. The sightings that justify the ratings can be published with the analysis results. For example, the page number and wording or activity, which is the sighting on which the rating is based, is published with the results of analysis of the content of mathematics textbooks according to the AAAS criteria in Figure 3.17.

Evaluation results must be represented in such a way that different textbooks can quickly and easily be compared. Since long discussions will discourage teachers (with already full programs) from reading it and profiting from the assessment results, tabled results and visual representations that give quick access to the information are favoured (Ansary and Babaii, 2002). *Tabulating* results makes it easier to compare the books. As discussed in the previous paragraph it is wise to tabulate not only the final results, but the results of every category. One example where this is done is the Standards-based Textbook Evaluation Guide developed by Indiana University (Cisar, n.d.). Henson (2004, 65) encourages teachers to use a table or "comparison chart" when comparing more than three books. The AAAS, for example, presented the results of the analysis of science textbooks in a table, using a column for each textbook and a row for each criterion. Figure 3.18 shows how they used symbols to represent the ratings (poor, fair, satisfactory, very good and excellent).

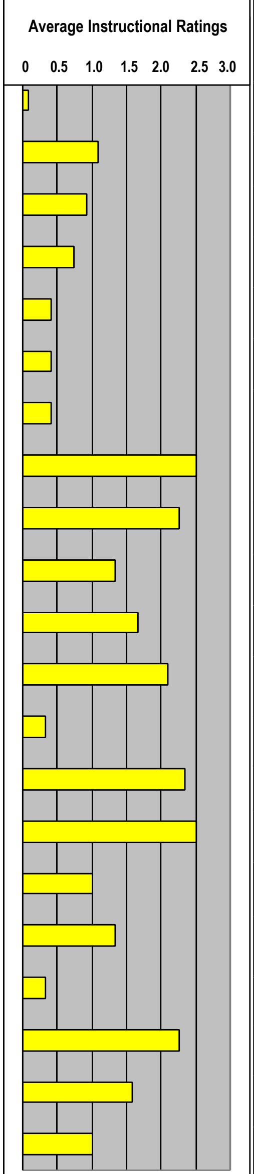
Graphic representations can depict the quality of a textbook relative to a benchmark to make it easy to compare their quality (Ansary and Babaii, 2002). This was done by the AAAS in their content analysis of mathematics textbooks, where the results were visually represented by a horizontal bar graph in Figure 3.17. A rating of 3 (the benchmark) for a specific criterion is represented by a full bar to the right.

3.9.4 Existing analysis instruments

The project that has received the most attention and acclaim in the sphere of science textbook evaluation is Project 2061 of the AAAS, as discussed in Section 3.2. They developed a five-step analysis procedure for science education (New Jersey Curriculum Framework, 1998). The evaluation procedures were developed and tested over a period of three years in collaboration with more than 100 scientists, mathematicians, educators and curriculum developers, with funding from the National Science Foundation.

A preliminary content inspection is used as screening procedure, as discussed in Section 3.9.3.1. This is followed by an analysis of content for a number of selected learning goals and finally the materials are analysed for instructional effectiveness. The AAAS provides indicators and a rating scheme for every criterion (as mentioned in Sections 3.9.3.2 and 3.9.3.3).

Typical Sightings for *Algebra 1: Explorations and Applications*



CRITERION	PAGE(S)	ACTIVITY	IDEA SET
I.1 Conveying Unit Purpose	T12	Chapter overview	Variables
I.2 Conveying Lesson Purpose	55	Learn how to... So you can...	Operations
I.3 Justifying Sequence of Activities	54t	Mathematical Connection	Functions
II.1 Specifying Prerequisite Knowledge	325t	Warm-Up Exercises	Functions
II.2 Alerting Teacher to Student Ideas	37t	Checking Key Concepts	Variables
II.3 Assisting Teacher in Identifying Ideas	30t	Warm-Up Exercises	Variables
II.4 Addressing Misconceptions	58t, 155t	Common Error	Operations
III.1 Providing Variety of Contexts	69-70	Exercises and Applications; Connection	Functions
III.2 Providing Firsthand Experiences	436-441	Exploring Polynomial Equations	Operations
IV.1 Justifying Importance of Standards Ideas	420-423; 420t-423t	Polynomial Functions; Interview Notes	Operations
IV.2 Introducing Terms and Procedures	119-122	Finding Equations of Lines	Variables
IV.3 Representing Ideas Accurately	326	Example 1	Functions
IV.4 Connecting Standards Ideas	262	Connection	Variables
IV.5 Demonstrating/Modeling Procedures	79-80	Example 1; Example 2	Functions
IV.6 Providing Practice	423-428	Adding and Subtracting Polynomials	Operations
V.1 Encouraging Students to Explain Their Reasoning	39	#38	Variables
V.2 Guiding Interpretation and Reasoning	90-91	Portfolio Project	Functions
V.3 Encouraging Students to Think about What They've Learned	176; 176t	Exploration; Exploration Note	Operations
VI.1 Aligning Assessment	187	#27-30, 37	Variables
VI.2 Assessing through Applications	182/182t; 186-187	Portfolio Project; Assessment	Operations
VI.3 Using Embedded Assessment	68t 83/83t	Closure Question #27/Assessment Note	Functions

Figure 3.17: Results for the analysis of the content of mathematics textbooks, including the sightings that justify the ratings (American Association for the Advancement of Science, 2002, 1)

AAAS Project 2061 Middle Grades Science Textbooks Evaluation

Summary of Instructional Analysis Ratings in Physical Science

Instructional Categories	Glencoe Life, Earth, and Physical Science (Glencoe/McGraw-Hill)	Macmillan/McGraw-Hill Science (Macmillan/McGraw-Hill)	Middle School Science and Technology (Kendall/Hunt)	Prentice Hall Science (Prentice-Hall)	PRIME Science (Kendall/Hunt)	Science 2000 (Decision Development Corp.)	Science Insights (Addison-Wesley)	Science Interactions (Glencoe/McGraw-Hill)	SciencePlus (Holt, Rinehart & Winston)	Matter and Molecules (Michigan State University)
	Textbook Series									
I. PROVIDING A SENSE OF PURPOSE										
Conveying unit purpose	Poor	Fair	Fair	Fair	Poor	Satisfactory	Poor	Poor	Poor	Poor
Conveying lesson purpose	Fair	Fair	Satisfactory	Poor	Poor	Fair	Poor	Poor	Fair	Very Good
Justifying activity sequence	Poor	Poor	Satisfactory	Poor	Fair	Satisfactory	Poor	Poor	Fair	Satisfactory
II. TAKING ACCOUNT OF STUDENT IDEAS										
Attending to prerequisite knowledge and skills	Poor	Poor	Poor	Poor	Fair	Poor	Poor	Poor	Poor	Satisfactory
Alerting teacher to commonly held student ideas	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Excellent
Assisting teacher in identifying own students' ideas	Poor	Poor	Poor	Poor	Poor	Fair	Poor	Poor	Satisfactory	Satisfactory
Addressing commonly held ideas	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Satisfactory
III. ENGAGING STUDENTS WITH RELEVANT PHENOMENA										
Providing variety of phenomena	Fair	Poor	Poor	Poor	Satisfactory	Poor	Poor	Satisfactory	Satisfactory	Satisfactory
Providing vivid experiences	Fair	Poor	Satisfactory	Poor	Very Good	Fair	Poor	Satisfactory	Satisfactory	Excellent
IV. DEVELOPING AND USING SCIENTIFIC IDEAS										
Introducing terms meaningfully	Fair	Fair	Very Good	Poor	Very Good	Fair	Fair	Satisfactory	Satisfactory	Excellent
Representing ideas effectively	Satisfactory	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Fair
Demonstrating use of knowledge	Fair	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Excellent
Providing practice	Fair	Poor	Poor	Poor	Fair	Poor	Poor	Fair	Satisfactory	Very Good
V. PROMOTING STUDENT THINKING ABOUT PHENOMENA, EXPERIENCES, AND KNOWLEDGE										
Encouraging students to explain their ideas	Poor	Poor	Satisfactory	Poor	Poor	Satisfactory	Poor	Fair	Satisfactory	Excellent
Guiding student interpretation and reasoning	Fair	Poor	Satisfactory	Poor	Poor	Fair	Poor	Poor	Fair	Excellent
Encouraging students to think about what they've learned	Fair	Poor	Fair	Poor	Poor	Poor	Poor	Poor	Poor	Poor
VI. ASSESSING PROGRESS										
Aligning assessment to goals	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Fair	Very Good	Very Good
Testing for understanding	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Fair	Satisfactory	Satisfactory
Using assessment to inform instruction	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Satisfactory

■ = Poor (0-1); ⬓ = Fair (1.5); □ = Satisfactory (2); ⬒ = Very Good (2.5); ▣ = Excellent (3)

Figure 3.18: A visual representation of a textbook evaluation

The questions that guide the content analysis were recorded in Section 3.9.1. The instructional support of textbooks is analysed according to criteria grouped in seven categories as shown in Table 3.2.

The AAAS instrument is limited in that it does not consider aspects like the cost, design or layout of science textbooks; it does not consider the relative importance of the different criteria to the overall quality and it is not set within the OBE context.

Table 3.3 shows the criteria developed by the National Science Resources Centre in the USA to assess science curriculum material. It consists of three sets or categories of criteria formulated as questions. The first set that considers the pedagogical appropriateness is further divided into three subsets that covers three different issues (National Academy of Sciences, 1997, 73).

The Gauteng Department of Education designed and utilised a National Curriculum Statement compliant evaluation instrument (Mahlaba, 2006). The instrument incorporates weights. Unfortunately the instrument targets textbooks for the Foundation and Intermediate Phases and is generic. The structure of the instrument and the weights are represented in Figure 3.19. All the criteria contribute equally to the overall result. The representation shows the maximum possible rates for criteria and categories and the percentages that the different categories contribute to the overall result.

In Figure 3.20 the hierarchy of the evaluation instrument of the Western Cape Department of Education is represented. The instrument consists of 25 criteria divided into 4 categories.The instrument has a section that focuses on science. The criteria used to judge the instructional qualities of the textbook are exactly the criteria used by the AAAS, but without the indicators and rating scheme.

The Namibian textbook assessment instrument uses a limited form of weighting. Important criteria have maximum ratings of 10 and less important ones a maximum of 5. The total marks are 175 and the number and kind of criteria on level 2 of every branch determine how much that branch contributes to the whole. The percentages in the diagram illustrate how much each branch contributes to the overall quality. The contributions are determined by the limited weighting and the number of criteria on level 2 that the branch represents. The hierarchy is depicted in Figure 3.21.

The instrument used by the South African National Department of Education has both a generic and a science section. The instrument consists of a list of desirable features in 5 categories (4 generic and 1 science). It does not provide indicators or utilise weights or even a rating scheme. It is represented as a horizontal hierarchy in Figure 3.22.

Table 3.2: The AAAS criteria for judging instructional effectiveness (American Association for the Advancement of Science, n.d.b)

I	**PROVIDING A SENSE OF PURPOSE** Conveying unit purpose Conveying lesson/activity purpose Justifying lesson/activity sequence
II	**TAKING ACCOUNT OF STUDENT IDEAS** Attending to prerequisite knowledge and skills Alerting teachers to commonly held student ideas Assisting teachers in identifying their students' ideas Addressing commonly held ideas
III	**ENGAGING STUDENTS WITH RELEVANT PHENOMENA** Providing variety of phenomena Providing vivid experiences
IV	**DEVELOPING AND USING SCIENTIFIC IDEAS** Introducing terms meaningfully Presenting ideas effectively Demonstrating use of knowledge Providing practice
V	**PROMOTING STUDENTS' THINKING ABOUT PHENOMENA, EXPERIENCES, AND KNOWLEDGE** Encouraging students to explain their ideas Guiding student interpretation and reasoning Encouraging students to think about what they have learned
VI	**ASSESSING PROGRESS** Aligning assessment to goals Testing for understanding Using assessment to inform instruction
VII	**ENHANCING THE LEARNING ENVIRONMENT** Providing teacher content support Encouraging curiosity and questioning Support all students

Table 3.3: National Science Resources Centre assessment criteria (National Academy of Sciences, 1997, 73–74)

Criteria for Judging Pedagogical Appropriateness
Addressing the goals of elementary science teaching and learning 1. Do the materials focus on concrete experiences with science phenomena? 2. Do the materials enable children to investigate important science concepts in depth over an extended period? 3. Do the curriculum materials contribute to the development of scientific reasoning? 4. Do the materials stimulate students' interest and relate science learning to daily life? 5. Do the materials build conceptual understanding over several lessons through a logical sequence of related activities? 6. Does the instructional sequence include opportunities to assess children's prior knowledge and experience? **Focusing on inquiry and activity as basis of the learning experience** 1. Does the material focus on student inquiry and engage students in the processes of science? 2. Does the material provide opportunities for students to gather and defend their own evidence and express their results in a variety of ways? **Using an effective instructional approach** 1. Does the material include a balance of student-directed and teacher-facilitated activities as well as discussions? 2. Does the material incorporate effective strategies for the teacher and/or students to use in assessing student learning? 3. Does the teacher's guide suggest opportunities for integrating science with other areas of the curriculum? 4. Do students have the opportunities to work collaboratively and alone?
Criteria for Judging Science Content
1. Is the science content current and accurately represented? 2. Does the content emphasise scientific inquiry? 3. Is the content of the science programme consistent with the National Science Education Standards? 4. Does the background material for teachers address the science content that is taught, as well as common misconceptions? 5. Is the treatment of content appropriate for the grade level? 6. Is the content free of bias? 7. Is the writing style for students and teachers interesting and engaging, and is scientific language used appropriately? 8. Is science represented as an enterprise connected to society?
Criteria for Judging Presentation and Format
1. Are the print materials for students well-written, developmentally appropriate, and compelling in content? 2. Are the directions for implementing activities clear in both the teacher's guide and the student material? 3. Are the suggestions for instructional delivery in the teacher's guide adequate? 4. Are the materials free of ethnic, cultural, racial, economic, age, and gender bias? 5. Are appropriate strategies provided to meet the special needs of diverse populations? 6. Are lists of materials for each activity provided, as well as a complete set of materials and information about reasonably priced replacement material? 7. Are safety precautions included where needed? 8. Are instructions for using laboratory equipment and materials clear and adequate?

	Level 1			Level 2	
Textbook	Compliance with RNCS	16	7,8%	Outcomes covered	4
				Outcomes in content and activities	4
				In line with constitution (sensitive issues)	4
				Caters for learners with varied abilities	4
	Assessment	20	9,8%	Diverse contexts	4
				Allows learners to integrate + apply know. and skills	4
				Different methods	4
				Assists learners to set own goals	4
				Targets thinking skills on different levels	4
	Integration of learning Areas	8	3,9%	Principles integrated: social justice etc	4
				Integrate other Learning Areas	4
	Approach	12	5,9%	Teaching strategies appropriate for content	4
				Various activities	4
				Activities encourage active participation	4
	Progression and Relationship	8	3,9%	Logical progress	4
				Activities meaningful related	4
	Context	16	7,8%	Start from learner's prior knowledge	4
				Contexts: Discover world beyond their own	4
				Balance familiar and new information	4
				Recognise various social structures	4
	Content	20	9,8%	Unbiased, no stereotypes	4
				Correlate with learner age and level	4
				Different cultural groups represented	4
				Accurate, up-to-date etc	4
				SA context	4
	Relationships content/illustrat	12	5,9%	Activities related to text and illustrations	4
				Address same outcome in act/text/ill	4
				Activities vary and prompt participation	4
	Language and Style	28	13,7%	Vocabulary clear, accessible and age appropriate	4
				New concepts introduced clear and logically	4
				No race discrimination in language	
				Gender sensitive language	4
				Religion and culture sensitive language	4
				Not stereotype language	4
				For students from various language backgrounds and abiliti	4
	Layout and Design	24	11,8%	Clear headings	4
				Clearly marked activities	4
				Print size and type suitable	4
				Adequate space and margins	4
				Appealing and attractive	4
				Holds learner interest throughout	4
	Pictures and Illustrations	12	5,9%	Pictures relevant to activities	4
				Picture details as described in text	4
				Race and gender representative	4
	Educator Support	28	13,7%	Components of package cross-referenced	4
				Adaptable	4
				Contains overview of curriculum	4
				Contains guidance to facilitate activities	4
				Contains guidance to application of assessment	4
				Contains solutions to problems	4
				Contains guidance to appropriate methods and strategies	4

Figure 3.19: The Gauteng Department of Education textbook analysis instrument in which all the criteria contribute equally to the overall result (Mahlaba, 2006)

	Level_1	Level_2	Level_3
Textbook	Adherence to NCSP	Compatible with NCS	
		Covers curriculum	
		Supports attainment of critical outcomes	
		Encourages active learning	
		Encourages critical thinking	
		Suitable assessment tasks	
	Good teaching material	Interactive and interesting	
		Content up to date	
		Structures material logically	
		Content accurate	
		User-friendly	
		Attractive	
		Appropriate concept level	
		Sturdy and re-usable	
		Not in workbook form	
		Appropriate vocabulary level	
		Teaching guide available	
		Encourages media use	
		Cost close to R100	
	Transformational issues	Sensitive to cultural groups	
		Avoids racist and sexist stereotyping	
		Represents all cultures	
		Acknowledges prior experiences	
	Physical science	Content	Balanced 3 curriculum knowledge strand
		Instructional qualities	Identifies a sense of purpose
			Builds on prior ideas
			Encourages scientific enquiry
			Develops scientific ideas
			Promotes science as social enterprise
			Provides assessment
			Enhances learning environment

Figure 3.20: Hierarchical representation of the criteria and categories used by the Western Cape Department of Education to evaluate textbooks (Western Cape Department of Education, 2005)

Level 1				Level 2		
Textbook	Physical characteristics	20	(11%)	How durable?	5	(2,86%)
				Typeface and size appropriate?	5	(2,86%)
				Appealing layout and appearance?	5	(2,86%)
				Cost reasonable?	5	(2,86%)
	Content	80	(46%)	Approach consistent with syllabus?	10	(5,71%)
				Caters for relevant knowledge objectives?	10	(5,71%)
				Caters for relevant skills objectives?	10	(5,71%)
				Reflects current knowledge and culture?	10	(5,71%)
				Factually accurate?	10	(5,71%)
				How free from biases?	5	(2,86%)
				Reflects various ethnic/cultural groups?	5	(2,86%)
				Free from sex stereotypes?	5	(2,86%)
				Encourages positive gender attitude?	5	(2,86%)
				Encourages positive environmental issues attitude?	5	(2,86%)
				Encourages positive population issues attitude?	5	(2,86%)
	Pedagogical aspects	60	(34%)	How understandable?	5	(2,86%)
				Assessment devices helpful to teacher?	5	(2,86%)
				Assessment devices helpful to learner?	5	(2,86%)
				Level differentiation?	5	(2,86%)
				Develops problem solving skills?	5	(2,86%)
				Easy for teacher to manage?	5	(2,86%)
				Activities stimulating and rewarding?	5	(2,86%)
				Are figures etc. appropriate?	5	(2,86%)
				Is there a TOC and index?	5	(2,86%)
				Special equipment needed?	5	(2,86%)
				Learner-centered approach?	5	(2,86%)
				Appropriate length	5	(2,86%)
	Language levels	15	(9%)	Appropriate reading level?	5	(2,86%)
				Appropriate comprehension level?	5	(2,86%)
				Challenging concepts explained?	5	(2,86%)

Figure 3.21: Hierarchical representation of the criteria and categories used by the Namibian Ministry of Education's National Institute for Educational Development (NIED) to evaluate textbooks (National Institute for Educational Development of the Namibian Ministry of Education, 2005)

	Level 1	Level 2
Textbook	Content/context	All outcomes and assessment standards
		All content & appropriately sequenced
		Suitable spaced and appropriate weights to LOs
		Current and up to date
		Integrates assessment standards within subject
		Theory and application integrated
		Sensitive to diversity e.g. culture, religion, gender, etc.
		Meaningful activities for groups, individual and pairs
		Level appropriate for grade
		Language and vocabulary appropriate for grade
		Key concepts and terms clearly defined
		Language and vocabulary appropriate for subject
	Learning activities & assessment	Activities and assessment derived from LOs and Ass
		Assessment activities appropriate to subject
		Assessment aligned with Subject Assessment Guidelines
		Variety
		Different levels of complexity
		Clearly formulated and unambiguous
		Provide for daily assessment
		Allow for expanded opportunities
		Reflect integration of Ass
	Layout,design & overall quality	Structured: Headings and subheadings
		Font and typeface clear and easy to read
		Illustrations and diagrams clear, relevant and unbiased
		Paper of good quality securely bound
		Index with chapter and page numbers
	Teacher Guide	Clear, systematic guidance on use of textbook
		Provide examples of work schedule and framework
		Includes examplar assessment plan
		Provides memoranda, checklists, rubrics etc. of assessment
		Provides suggested answers, etc. for activities and exercises
	Physical Sciences	Covers basic concepts extensively and accurately
		Consistent with SI units and symbols
		Act&Ass promote\assess application of competencies
		Guide `hypothesis testing' approach to Ass
		Provide framework for learner research projects
		Includes latest trends in scientific inquiry and research
		Enable learner practical activities
		Illustrations and diagrams have detailed captions, etc.
		Volume of content appropriate for 4 hours per week

Figure 3.22: The South African representation of the criteria and categories used by the South African National Department of Education to evaluate textbooks

Table 3.4: Comparison of existing textbook evaluation instruments

Instrument	Science/Generic	Categories	Criteria	Indicators	Rating scheme	Weights	OBE	SA
National Department of Education	S	5	40	No	No	No	Yes	Yes
Gauteng Department of Education	G	12	51	No	No	Yes	Yes	Yes
Western Cape Department of Education	S	4	34	No	No	No	Yes	Yes
Mozambique Department of Education	G	5	21	No	No	No	No	No
Namibian Department of Education	S	4	30	Yes	No	Yes	No	No
U.S. Department of Education	S	4	8	Yes	No	No	No	No
AAAS	S	7	22	Yes	Yes	No	No	No

All the instruments that have been discussed in this section are included in the Appendix B. Table 3.4 was compiled to compare the characteristics of the various instruments.

The table clearly shows that none of the available instruments that utilises indicators, weights and rating schemes are appropriate for science textbooks in the South African context. A new instrument for this purpose is needed.

3.10 Designing evaluation instruments

Many textbook analysis systems are reported without any justification for the choice of criteria, or the assigned weights (Brown, 1998). Ansary and Babaii (2002) even contend that textbook selection to date has been made in haste and with a lack of systematically applied criteria. In this regard Fetsko (1992, 133) argues as follows:

> Time spent in designing the analysis instrument will pay great dividends throughout the process and should help the committee make the best possible decision.

3.10.1 Documented approaches to designing textbook evaluation instruments

Whether an evaluation instrument only consists of a checklist of important characteristics that should be present in a textbook or is more complex, the first step in the design of a textbook evaluation instrument is the formulation of criteria. The literature survey revealed four approaches to formulating criteria for textbook analysis, namely:

- criteria selected from the literature (publications of the results of research on texts, textbooks and learning of science) (Pepin and Haggerty, 2003; Bernier, 1996)
- criteria derived from stakeholder concerns (teachers, parents, et cetera) (MacDonald, 2006)
- criteria derived from expert opinion (Ansary and Babaii, 2002) and
- criteria taken or adapted from existing instruments (Davies, 2003; Goldsmith et al., 2000)

A *literature survey* was the point of departure in the work of Pepin and Haggerty (2003) on analysis of mathematics textbooks. They developed a 'schedule' for the analysis of mathematics textbooks by drawing on ideas from literature.

The AAAS (Kesidou and Roseman, 2002, 524) used *expert opinion* to shape their criteria. They developed their textbook evaluation system in a pragmatic way. Three groups of experts in science teaching and learning research and curriculum development got together to help formulate, negotiate, test, and refine a set of criteria.

According to MacDonald (2006) *stakeholder concerns* is a relatively new approach to criteria determination. Evaluation has evolved over time from neutral outsiders measuring, describing and judging what is true, to evaluation orientated and organised around the concerns of stakeholders. The first stakeholder to come to mind is the state. The Constitution of the Republic of South Africa gives an indication of the state's interest. For example, one of the core aspects of the Constitution is to heal the divisions of the past and establish a society based on democratic values, social justice and fundamental human rights (Department of Education of South Africa, 2003a). In the set of criteria for textbook evaluation the state's interest is reflected in the criteria that require a good textbook to promote these values. For example, scientific applications used in examples, activities and assessment are situated in different contexts. The way in which the textbook introduces the content can promote positive or negative attitudes towards these issues. For example, law enforcement officers carrying out their duty can be depicted as efficient or incompetent. Science textbooks are also uniquely suited to encourage learners to identify and critically evaluate the impact of scientific knowledge on socio-economic and environmental issues.

The second obvious stakeholder to consider is the Department of Education. Mikk ((2000)) contends that the most important guideline for the composing of criteria must be the education department's goals as represented in the curriculum documents. The policy documents and National Curriculum Statement for Physical Science (Department of Education of South Africa, 2003a), therefore, could be used as point of departure in the development of a set or hierarchy of criteria for the evaluation of textbooks for Physical Science.

Mikk (2000) even suggests that every goal of education must be represented in the set of criteria and that the weight attributed to the goal must be in accord with how frequently it is mentioned in the curriculum document. For example, if thinking abilities is mentioned 20 times among 200 educational goals, the weight attributed to thinking abilities is 10% of the whole. This approach can only be considered if the curriculum document was written to represent the goals in the desired ratio. This may not always be the case. For example, in the National Curriculum Statement for Physical Science (Department of Education of South Africa, 2003a), a large portion of the text is dedicated to assessment. This is not necessarily an indication that assessment is considered to be the most important aspect. The extensive use of assessment, especially formative assessment, in education is new to most teachers and the curriculum document discusses it in detail to inform teachers. Some aspects may not be mentioned because it is considered obvious.

Criteria should include the following educational principles spelled out in the National Curriculum Statements (Department of Education of South Africa, 2003a):

- social transformation
- outcomes-based education
- high knowledge and high skills
- integration and applied competence
- progression
- articulation and portability
- human rights, inclusivity, environmental and social justice
- valuing indigenous knowledge systems
- credibility, quality and efficiency

The subject Physical Science imposes further demands on good science textbooks. In this regard the following learning outcomes specified for Physical Science in the National Curriculum Statements must be reflected in the chosen evaluation criteria:

- scientific inquiry and problem solving in a variety of scientific, technological, socio-economic and environmental contexts
- the construction and application of scientific and technological knowledge
- the nature of science and its relationship to technology, society and the environment

Other stakeholders whose interests must be considered include parents, teachers, learners and society. The interests of the learner and the teacher in textbook quality were discussed in in Section 2.3.3 of Chapter 2. Criteria that relate to industry concerns are the development of learners' ability to transfer and apply their knowledge in new contexts.

Existing checklists and instruments can serve as indirect sources of information on *experts' opinions*, and can, therefore, serve as point of departure and touchstone when a new set of criteria must be developed for a specific context (Davies, 2003; Goldsmith et al., 2000). Ansary and Babaii (2002) utilised the second approach in a novel way. In the context of English Language Education they attempted to identify some theory-neutral, universal, and broad consensus-reached characteristics for English language textbooks. They utilised the opinions of experts as it is manifested in previously published textbook reviews and checklists. They scrutinised ten textbook reviews and ten evaluation checklists and selected from them a universal set of common-core characteristics.

Bernier (1996) *combined two of the approaches* mentioned above. She developed an instrument for evaluating educational material used in medical settings for patient education. The approach followed for the development of the Bernier Instructional Design Scale (BIDS) was thoroughly documented and justified. Bernier generated 90 instructional design and learning principles from a review of literature in the disciplines of nursing, health education, instructional design, and educational psychology. Thirteen content experts then rated the principles on a five point scale from nonessential to essential. If 80% or more of the experts rated the principle 4 or 5 on the scale it was included in the Bernier Instructional Design Scale (BIDS). The 37 principles that met the condition were arranged by sub-domain categories and given a 4 category rating scale.

3.10.2 Textbook analysis as multiple criteria decision making

Deciding which textbook represents the best choice for a specific situation is a decision that is influenced by multiple objectives or criteria. A checklist of desirable characteristics, like the National Department of Education's evaluation instrument (Department of Education of South Africa, n.d.) represents a sample of the objectives or criteria that can be considered. This may not be enough to distinguish between two books. Winston illustrates the complexity of multiple objectives or criteria decisions with the following example (Winston, 2004, 785):

> In determining which job offer to accept, a job seeker (call her Jane) might choose between the offers by determining how well each one meets the following four objectives:
>
> Objective 1 High starting salary
> Objective 2 Quality of life in city where the job is located
> Objective 3 Interest in work
> Objective 4 Job location near family and relatives
>
> When multiple objectives are important to a decision maker, it may be difficult to choose between alternatives. For example, one job offer may offer the highest starting salary, but it may score poorly on the other three objectives. Another job may meet objectives 2–4 but have low starting salary. In such a case, it may be difficult for Jane to choose between job offers.

In Winston's example Jane will now have to consider how important each of her objectives is. She may decide to discard some of the objectives and keep only the most important one or two. Jane may also decide to consider the relative importance of every objective. This creates a much more complex problem. Likewise, when we consider textbook quality it is necessary to take into account how much every characteristic or criterion influences or contributes to the quality of the textbook. Saaty (1996, 1) maintains that:

> Breaking a problem down into its constituent parts or components, in a framework of a hierarchy or a feedback network, and establishing importance or priority to rank the alternatives is a comprehensive and general way to look at the problem mathematically. This kind of concern has been loosely called multi-criteria decision making.

A textbook evaluation instrument embodies the decision making process up to the point of the final application and evaluation. The instrument is the culmination of the considerations that precede the actual evaluation process. In future evaluators can use the instrument to save the time involved in the "breaking down" and "establishing of priorities." The instrument of the Gauteng Department of Education (Mahlaba, 2006) is an example of the result of the preliminary "breaking down" and "establishing of priorities." Unfortunately no justification for their assignment of weights is available and the reliability and validity of the instrument are not addressed.

Weights can be assigned directly to attributes according to its contribution to the quality of the textbook. Unfortunately human beings find it difficult to rank more than seven (plus or minus two) items at any time and there may be as many as two hundred attributes to consider (Miller, 1956). Techniques are needed to assist with the determining of the weights. The field of the Operations Research and Management Science may provide the answer. This idea will be explored in the next section.

3.10.2.1 Multiple Criteria Decision Making (MCDM) techniques

Operations Research and Management Science use scientific approaches in decision making. Techniques for Multi-Criteria Decision Making (MCDM) have been developed within the field of Operations Research and Management Science to deal with complex decision making problems (Winston, 2004, 1).

MCDM techniques use mathematical models. These models represent the actual situation and may be used to make better decisions. They can also be used to derive scientifically justifiable weights for objectives or criteria (Winston, 2004). MCDM techniques have been applied to a range of decision making problems in a variety of fields.

There are a number of MCDM techniques available like the Simple Multiple Attribute Rating Theory (SMART), the Multi Attribute Utility Theory (MAUT) and the Analytic Hierarchy Process (AHP). Any of these, or elements from them, can be used in decision making processes. The Analytic Hierarchy Process seems to be the most appropriate for use in the textbook evaluation situation and will be explored in the next section.

3.10.2.2 Analytic Hierarchy Process (AHP)

The Analytic Hierarchy Process was developed by Thomas Saaty at the Wharton School of Business. It has been used on many educational problems in the past. For example, Friar, Matson and Matson (1998) used it to evaluate undergraduate curricula; Blin and Tsoukiàs (2001) used it to evaluate software quality and Lee, McCool and Napieralski (2000) used it to assess adult learning preferences. It is deemed a valuable decision making tool since it "allows a fully documented and transparent decision to be made with full accountability" (Davidson and Labib, 2003, 207).

The AHP breaks the problem down into individual criteria and builds a decision hierarchy (Roy, 1996) where the number of levels in the hierarchy depends on the complexity of the problem (Friar et al., 1998). Thereafter, the importance of criteria is compared pair-wise. The data is recorded in a matrix with both the columns and rows representing the individual criteria. An eigenvalue mathematical method is then used to determine both the relative priority of each criterion and a "consistency ratio."

Thomas Saaty's Analytic Hierarchy Process (AHP) has five characteristics that makes it especially applicable to this study:

1. The AHP utilises a *hierarchical modelling* of the problem by forming groups and subgroups of criteria. A hierarchical representation of the textbook quality problem is a very natural approach to the problem. For example, several hierarchical representations have been used in Section 3.9.3.1 to compare the different instruments.

2. The AHP utilises pairwise comparisons "to *derive* ratio scale priorities or weights as opposed to arbitrarily *assigning* them" (as is done with SMART) (Forman and Selly, 2002, 43). The use of pairwise relative comparisons in the AHP is based on the assumption that it is easier and more realistic for decision makers to compare elements in pairs than to directly assign weights to the attributes (Iz, 1991). These comparisons are then used to mathematically determine the weights (Forman and Selly, 2002).

3. Both *quantitative and qualitative considerations* can be incorporated in the model (Friar et al., 1998).

4. Logical thought should enable humans to be consistent in their thinking, but we seldom are. The AHP is developed to utilise mathematical techniques to give an acceptable result even if the data, derived by pair-wise comparison of attributes, is only consistent within certain bounds. The AHP allows for inconsistency and it even provides a *measure of the consistency* of the data (Winston, 2004).

5. *Computerised decision support* is available for individual decision makers and also for groups. A group of experts (teachers and teacher trainers) used as data sources consists of individuals with different perceptions regarding the

importance of criteria. Expert Choice is one of the computer programs that can assist groups in implementing the AHP.

3.11 Conclusion

This chapter was devoted to a discussion of the different approaches to and methods used in textbook assessment. Textbook analysis was indicated as the most suitable approach for selection of textbooks. The use of a textbook assessment instrument was motivated and the various components of such an instrument were discussed. It was indicated that an appropriate instrument that can ensure reliable and valid results should incorporate weighted criteria and indicators of these criteria, as well as transparent evaluation and recording procedures.

Approaches that can be utilised in the development of a textbook assessment instrument were discussed and it was pointed out that the process of development as well as the final instrument must be thorough, justifiable and produce valid and reliable evaluation results. The AHP was identified as the most suitable method for the determination of weights that represent the contribution of the different criteria to the overall quality of the textbook.

The design of the experimental study will be discussed in the next chapter.

Chapter 4

Research design

4.1 Introduction

In a world where science and technology are transforming day-to-day living and influencing the future, science education has become imperative for all learners. Textbooks play an important role in this science education and also in the development of the large portion of science teachers in South Africa who have limited science qualifications or teaching experience. It is only logical that the most appropriate textbooks should be selected and used in our classrooms.

To provide the background knowledge for an investigation into textbook evaluation and selection, Chapter 2 focussed on education and science education and the role of the textbook within this context. The construct "quality" as embodied in science education textbooks and the evaluation and measurement of this quality were probed in Chapter 3. An investigation into the existing evaluation instruments for science education textbooks identified a need for an instrument with both criteria and weights that are valid and reliable.

The investigation into existing evaluation instruments confirmed that textbook evaluation is a multi-criteria problem. Criteria are not equally important when the quality of the textbook must be determined. Textbook evaluation instruments traditionally handle the problem in one of two ways. The first is to categorise criteria as important enough to be considered or not important enough to consider. The criteria deemed important enough to consider are then treated as equally important. The second approach used is to determine in some way how important the criteria are in relation to each other and applying weights to adjust the contribution of every criterion to the total score for textbook quality.

The Analytic Hierarchy Process (AHP), a technique for Multi-Criteria Decision Making (MCDM), was utilised to develop a new instrument for the evaluation of South African science education textbooks.

4.2 Hypothesis

The overarching hypothesis that directed the experimental study states that it is possible to develop a suitable instrument for the assessment of the quality of science education textbooks in the South African Education. To verify this hypothesis the following notions were formulated:

1. the AHP can be successfully adapted to constitute an appropriate process to be utilised in the development of a textbook evaluation instrument and
2. an instrument can be developed that gives transparent, reliable and valid evaluation results.

The evaluation instrument is limited to science textbooks and does not include teachers'guides.

4.3 Development of the Instrument for the Evaluation of Science Education Textbooks (IESET)

The AHP that was used to inform the development of the IESET, breaks the evaluation process down into individual criteria and consequently builds a decision hierarchy. Therefore, the formulation of criteria and ordering it in a hierarchical form was the first stage in the development of the instrument.

The second stage was deriving weights that represent and account for the relative contributions of specific criteria to the quality of the textbook as a whole. This was done according to the pair-wise comparison method used by the AHP process.

In parallel with this process a rubric was developed that is used as rating scheme and score sheet in the final instrument. This is not an essential element of the AHP, but was incorporated to increase the reliability of the results of the instrument.

The criteria, weights and rubric were then combined to give the Instrument for the Evaluation of Science Education Textbooks (IESET). The IESET was implemented on a pilot scale, adjusted where necessary and applied on a larger scale to test for validity, reliability and ease of application.The basic stages in the development and testing process is summarised in the following flowchart.

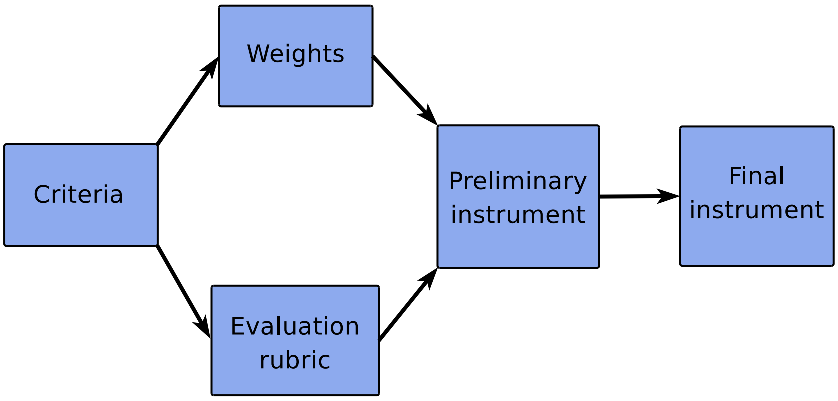


Figure 4.1: Diagram of the planned development and testing of a textbook evaluation instrument

The various stages of the development and testing of the IESET is broken down in greater detail in Table 4.1.

The remainder of the chapter is devoted to an explanation of the different phases of the development process and instrument testing.

4.3.1 Formulation of criteria

There are many aspects or characteristics of a textbook that can influence its success in supporting the teacher and learner. Valid results about the quality of the textbook can only be obtained by the use of criteria that are based on thorough research. This research must include studies on the characteristics of the textbook that contribute to learner achievement of the science education goals. These characteristics were identified and relevant criteria formulated. The literature survey pointed out a number of approaches that can be utilised in this regard, as discussed in Section 3.6. Three of these approaches were combined to develop the criteria for the IESET:

- With the first approach criteria were selected from the relevant *literature*. Chapters 2 and 3 reported the results of the literature survey on research on how children learn from text, with special reference to the science education context. From these results relevant researched-based criteria were derived.
- The second approach was used to add criteria derived from stakeholder concerns. The State and the Department of Education are the most obvious stakeholders in this regard. The Constitution of the Republic of South

	Process	Data source	Aim
Qualitative research	Literature survey Formulation of criteria	Literature on: textbook evaluation and existing analysis instruments	Discover relevant indicators/criteria Discover relations between indicators
	Ordering of criteria in a hierarchical structure (tree)	Literature on: textbook evaluation and existing analysis instruments	To obtain a hierarchical model of the construct "science textbook quality"
	Semi-structured interviews	Two experts in: - outcomes-based education - science education - instructional design	Determine the content validity of the criteria of the IESET
Quantitative research	Pairwise comparisons and data analysis with Expert Choice: AHP computer software	Panel of eight experts in: - outcomes-based education - science education - instructional design Independent facilitator	Determine the relative importance of indicators/criteria of textbook quality. Add weights to indicators to complete the instrument
	Pilot implementation, analysis of data and adjustment of instrument	Four Teachers Two Textbooks	Compare results to determine reliability Analyse instrument sensitivity Make adjustments to instrument
	Larger scale implementation, analysis of data	Twelve Teachers Two textbooks	Determine reliability
Qualitative	Semi-structured interviews	Teachers who have used the textbooks	Compare data of teachers experience with instrument to verify the validity of the instrument

Table 4.1: Development and testing of IESET

Africa was considered as representative of the State's interest, while policy documents gave a clear perspective on the Department of Educations' concerns, as discussed in Section 3.6.

- *Existing evaluation instruments* were scrutinised in the final approach. The existing instruments served as indirect sources of information on *experts' opinions*. This process was complicated by the difference in formulation of the criteria and criteria that were not clearly defined.

The approaches used to formulate the criteria generated research-based criteria that are valid and contribute to the validity of the whole instrument and, therefore, to the validity of the evaluation results delivered by the instrument.

Criteria that were found to be so vital that any textbook that does not meet these criteria, are not considered for selection, and are combined to give a collection of *screening criteria*. The first step in the implementation of the IESET is to analyse textbooks for their compliance with the screening criteria. Only textbooks that satisfied these minimum criteria were then evaluated according to the weighted criteria of the IESET.

4.3.2 Development of the quality model

The criteria were ordered to develop a hierarchical model of the quality of science education textbooks in accordance with the AHP, as discussed in Section 3.10.2.2. The categories used in existing instruments suggested the branches on the first level of the model. The hierarchy was necessary for the implementation of the AHP, but it also provided a clearer visual picture of the criteria than a plain list of criteria. The hierarchy also made it possible to give the results for the separate categories.

To illustrate the process on a scale that is easy to comprehend an instrument developed by MEAS, referred to in Section 3.9.3.1, will be used as example. The final hierarchy of the IESET was similar to, but more complex and comprehensive than the model of MEAS depicted in Figure 3.6 and Figure 3.7.

4.3.3 Content validity of the quality model

Interviews were conducted with two experts in the field of science education to determine whether they consider the criteria and the hierarchical model a valid representation of aspects that indicate quality in science textbooks in the South

African education system. This gave an indication of the validity of the content of the model.

4.3.4 Development of a rubric for the rating of criteria

The validity and reliability of the evaluation results of the IESET is of paramount importance and were considered in every stage of the development of the instrument. If analysts have to award a rating between 3 and 0 for a specific criterion, different analysts could make different subjective decisions. A rubric that can guide analysts in the rating of textbooks can ensure that different analysts look for the presence of the same indicators (as discussed in Section 3.9.3.2). This makes the process more objective and increases the reliability of the ratings and the final results.

A rubric was created similar to the History and Social Science Textbook and Instructional Materials Review used by Virginia (shown in Figure 3.10). The rubric incorporates the indicators that must be present to satisfy every criterion, as well as a rating scheme. This was based, where appropriate, on research reports on specific aspects discovered during the literature survey. The rubric is used by analysts during textbook evaluation to guide the awarding of ratings. The rubric describes what indicators must be present to justify a specific rating. In the rubric the description of a textbook that is awarded a rating of 3 for a criterion is actually the benchmark against which the textbooks are measured.

The rubric also acts as scoring sheet. Analysts record the ratings on the rubric by indicating each rating with a mark in the relevant column of every row. Every row represents a criterion. The first column identifies the criterion and columns 2 to 5 contain the rating scheme for that criterion. Analysts are not burdened by separate rubrics and scoring sheets. An analyst who knows the content and learning outcomes will only need the textbook and the score sheet to complete the evaluation. The scores are then transferred for computer computation and presentation of results.

4.3.5 Deriving the weights of criteria

Not all aspects that contribute to textbook quality contribute equally. Simply adding the ratings for all the criteria would assume that each criterion makes the same contribution to the quality of the textbook. The solution is to use weights that represent the relative importance of every criterion to the overall quality. During the computation of results the specific ratings are multiplied by the weights of the criteria to incorporate their relative contributions to the quality.

To determine the weights of the criteria a panel of eight experts were convened. The panel was asked to do pairwise comparisons of the importance of different criteria to the overall quality. Experienced teachers and educator trainers were asked to participate in the research. This was a *purposeful non-random sampling* that is "a useful method of getting information from a sample of the population that you think knows most about a subject" (Walliman, 2005, 279).

The method used was adapted from the AHP. It is based on the assumption that it is easier and more realistic for decision makers to compare elements in pairs than to directly assign weights to the attributes. The panel did pairwise comparisons of the importance of different criteria to the overall quality. The comparisons were made in group sessions with the aid of a facilitator.

On every level of the hierarchy the importance of each characteristic within a specific branch was compared with the importance of every other characteristic. The rubric, with brief descriptions of the indicators of every criterion, was provided to the panel members prior to the process. Respondents used a scale of real numbers to systematically assign preferences. Each judgement was expressed as the ratio of the importance of one criterion compared to another criterion (Forman and Selly, 2002).

An example of the planned questionnaire (see Appendix C) used during the session is depicted in Figure 4.2.

The opinions of the panel members were processed with the Expert Choice AHP computer package. The consistency index was calculated for every individual (and the aggregated results) to ensure that they were within acceptable limits. The weights can be represented as percentages to show how much the group perceives an attribute to contribute to an attribute in the next level and finally to the quality of the textbook. For example, consider the random weights used in Figure 4.3. The attribute 'Index' contributes 20% towards the higher level attribute 'Structure', that in turn contributes 40% to 'Layout', which carries a weight of 20% in the evaluation of the textbook. Thus the contribution of 'Index' to the measure of quality of the textbook is $\frac{20}{100} \times \frac{40}{100} \times \frac{20}{100} = 1,6\%$.

The percentages are actually indicators of the *validity* of the individual criteria. An unimportant criterion will have a low weight and, therefore, the rating for this criterion will make only a small contribution to the textbook's overall score. In this way the derived weights of the criteria ensure the validity of the overall quality score.

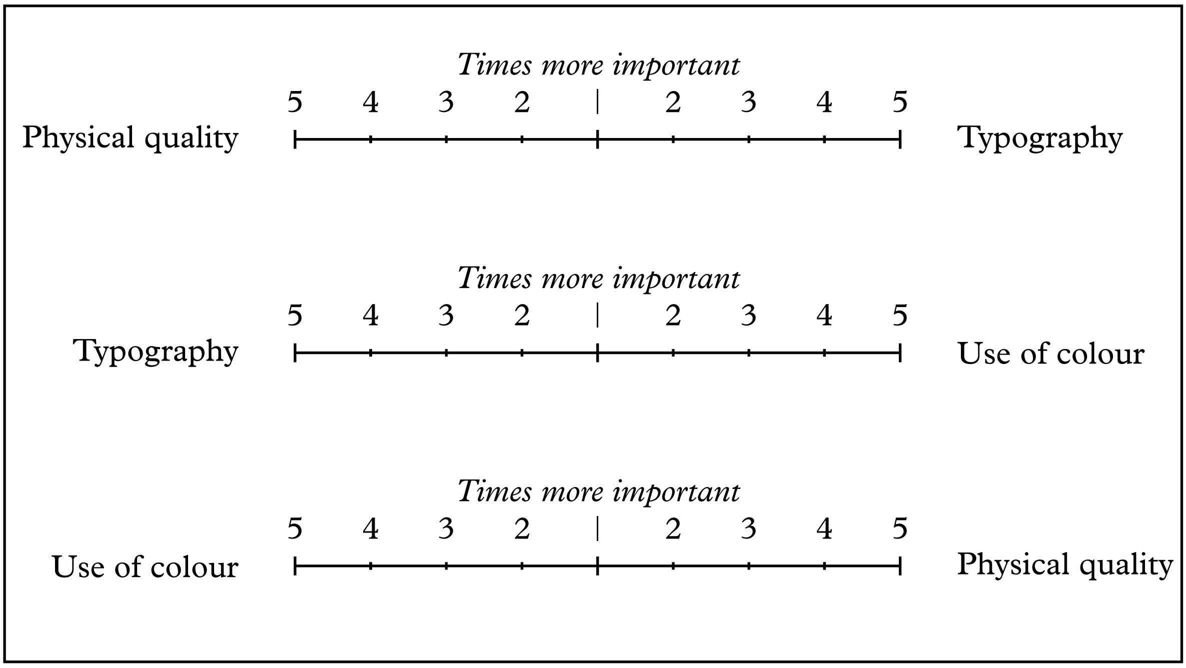


Figure 4.2: Extract from the questionnaire used for pairwise comparisons

The weights derived in this process were displayed and later used in spreadsheet format. To illustrate the format, random weights have been attributed to the hierarchy of the MEAS represented in Figure 3.6 and it is displayed in the planned spreadsheet format in Figure 4.3.

4.4 Construction of the instrument

The evaluation instrument consists of

- instructions to standardise the use of the instrument;
- a list of screening criteria as discussed in Section 3.9.3.1;
- the rubric for every criterion to guide the assigning of ratings for that criterion and also act as score sheet; and
- a spreadsheet with the hierarchy of criteria, weights and relevant formulæ for the final calculation already entered.

The implementation of the IESET starts with the evaluation of a textbook against the screening criteria. Textbooks that satisfy the screening criteria are thereafter rated on every criterion on a scale from 0 to 3, according to the rubric. For every criterion the rubric specifies the indicators that must be present for a textbook to satisfy the criterion. Every criterion is represented by a row in the rubric. The analysts record their ratings by marking the cell entry on the rubric that corresponds with the level of quality the textbook exhibits for every criterion.

	Level 1	Weight 1	Level 2	Weight 2	Level 3	Weight 3
Textbook	Content	50%	Outcomes	40%	Explicit?	45%
					All?	15%
					Integrate?	40%
			Accuracy	60%	Facts correct?	75%
					Relevant?	15%
					Range?	10%
	Layout	20%	Structure	40%	Consistent structure?	50%
					Contents?	10%
					Index?	20%
					Cross-references	20%
			Linguistic	30%	Level correct?	35%
					Readability	35%
					Terminology	30%
			Technical	30%	Physical quality	45%
					Typography	50%
					Use of colour	5%
	Didactics	30%	Pedagogic	40%	Differentiating	45%
					Pre-knowledge	25%
					Context	30%
			Practical	40%	Activities?	35%
					Exercises	60%
					Group work?	5%
			Auxiliary	20%	Tables?	30%
					Diagrams?	30%
					Pictures relevant?	40%

(The inconsistent use of question marks with criteria is as published by the MEAS group.)

Figure 4.3: An illustration of the planned spreadsheet format using the MEAS hierarchy of criteria and random weights

The ratings of different analysts can be combined or individually entered into the last column of the spreadsheet. The quality measure of the textbook on every subgroup or category and the textbook as a whole can thereafter automatically be calculated and displayed.

To give a clear idea of this phase of the evaluation procedure the ratings for a hypothetical textbook was entered in the spreadsheet used as example in Figure 4.3 and the result displayed in Figure 4.4.

The hypothetical textbook represented by the ratings in Figure 4.4 achieved, inter alia, a value of 80% ($\frac{2}{3}\times\frac{45}{100}+\frac{2}{3}\times\frac{15}{100}+\frac{3}{3}\times\frac{40}{100}$) against the benchmark for outcomes. The teacher can, therefore, assume that the textbook does address most of the outcomes adequately. For easy interpretation these strengths and weaknesses of the textbook can be represented as percentages on a modified bar graph. As an example Level 2 of the results in Figure 4.4 is represented as a bar graph in Figure 4.5.

	Root value	Level 1	Weight 1	Value 1	Level 2	Weight 2	Value 2	Level 3	Weight 3	Rating
Textbook	**71.17%**	Content	50%	87%	Outcomes	40%	80%	Explicit?	45%	**2**
								All?	15%	**2**
								Integrate?	40%	**3**
					Accuracy	60%	92%	Facts correct?	75%	**3**
								Relevant?	15%	**2**
								Range?	10%	**2**
		Layout	20%	53%	Structure	40%	43%	Consistent structure?	50%	**1**
								Contents?	10%	**2**
								Index?	20%	**0**
								Cross-references	20%	**3**
					Linguistic	30%	67%	Level correct?	35%	**1**
								Readability	35%	**3**
								Terminology	30%	**2**
					Technical	30%	53%	Physical quality	45%	**0**
								Typography	50%	**3**
								Use of colour	5%	**2**
		Didactics	30%	57%	Pedagogic	40%	72%	Differentiating	45%	**3**
								Pre-knowledge	25%	**2**
								Context	30%	**1**
					Practical	40%	53%	Activities?	35%	**1**
								Exercises	60%	**2**
								Group work?	5%	**1**
					Auxiliary	20%	33%	Tables?	30%	**0**
								Diagrams?	30%	**2**
								Pictures relevant?	40%	**1**

Figure 4.4: Example of a spreadsheet with the results of an evaluation of a textbook

4.5 Pilot implementation of the instrument

A pilot implementation of the completed instrument were conducted to identify any problems in the application procedure or unclear items in the rubric and gave an preliminary indication of its reliability and validity.

Four well-qualified teachers with extensive teaching experience and knowledge of OBE, Physical Science learning outcomes and the core syllabus evaluated two different science textbooks, using the IESET. The *reliability* of the instrument can be assessed by comparing the results of different analysts on each of the textbooks (Nogova and Huttova, 2006). The teachers were also interviewed to gather information on their evaluation experience with the instrument. The instrument was adjusted in the light of the information gained in the pilot implementation.

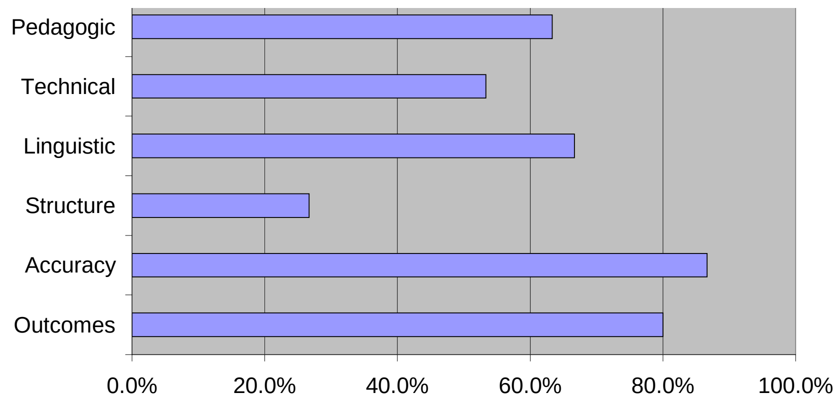


Figure 4.5: Visual representation of the results of level 2 for the hypothetical textbook and MEAS criteria

4.6 Implementation of the instrument on a larger scale

The instrument was then implemented on a larger scale and the reliability verified by comparing the results of different analysts for the same textbook. Twelve teachers (with exemplary qualifications and experience) used the IESET to evaluate two textbooks. The implementation of the IESET was followed by semi-structured interviews with teachers who have used the textbooks and were able to provide information on how well it supports the learning process. The *validity* of the IESET was confirmed by triangulation. Triangulation is often used to indicate that more than two methods are used in a study with a view to double-checking results. The results of the evaluations were compared with the opinions of experts or teachers that had used the textbooks in their classes to see how well the results of the textbook assessment with the IESET correlate with the teachers' experience.

4.7 Conclusion

This chapter described the experimental study that was designed to verify the hypothesis formulated in Section 4.2. The development process incorporated elements of the AHP process (hierarchy of criteria and pair-wise comparisons for determining weights). The results of the development and testing of the IESET are reported in the next chapter.

Chapter 5

Results of the empirical study

5.1 Introduction

The importance of good textbooks for science education makes it imperative to have good methods for selecting textbooks. It was found that the available textbook assessment instruments are not appropriate for physical science textbooks in the South African context (see Section 3.9.4). To address this problem an investigation into the relevant literature was undertaken. The investigation lead to the formulation of the hypothesis, stated in Section 4.2, that it is possible to develop a suitable instrument for the assessment of the quality of science education textbooks in the South African context. The hypothesis was extended to encompass the statements that elements from the Analytical Hierarchy Process (AHP) can be successfully adapted to give an appropriate process for the development of a textbook evaluation instrument and that such an instrument can be developed to give transparent, reliable and valid evaluation results.

This chapter contains a report of the development of the Instrument for the Evaluation of Science Education Textbooks (IESET) and the testing to verify all aspects of the hypothesis.

5.2 The development of the Instrument for the Evaluation of Science Education Textbooks (IESET)

To verify the hypothesis that elements from the AHP can be successfully adapted to give an appropriate process for the development of a textbook evaluation instrument, every stage of the process was considered, as well as the final version of the IESET.

5.3 Hierarchy of criteria

In accordance with the AHP the development of the IESET started with the identification and formulation of relevant criteria. As discussed in Section 4.3.1 criteria were selected from the literature; taken or adapted from existing instruments and derived from stakeholder concerns.

The following *screening criteria* were identified that are absolutely vital to the quality of any textbook:

- it contains no offensive text or pictures
- it contains at least 80% of the core knowledge
- the paper and binding will last for the relevant period

During the implementation of IESET, the first step is to determine if the textbooks satisfy the screening criteria. Any textbook that does not satisfy these criteria can not even be considered as a candidate for selection and will not undergo further assessment, saving unnecessary time and effort.

For the general assessment, the different approaches yielded 58 criteria that contribute to the quality of a textbook. These criteria were structured to give a hierarchy on four levels. The number of levels in the hierarchy reflects the complexity of the evaluation of textbook quality. Level 0 represents the quality of the textbook as a whole. This initial hierarchy is depicted in Figure 5.1.

To verify that the criteria and hierarchy are representative of textbook quality (content validity) the hierarchy of criteria was presented to two experts in science education. Semi-structured interviews were held and the experts were given the opportunity to study the hierarchy and to disapprove of categories or criteria or to suggest new ones. The experts were also given the opportunity to move criteria to other categories or branches of the hierarchy. In light of their contributions the criteria were reduced to 55, but the original categories or branches were retained.

In the science branch in Level 2 the criterion "promote science as inquiry" was removed. The experts contended that this criterion is represented both by the learning outcome branch (the first learning outcome) and by the nature of science branch. In the assessment branch the *formative and summative opportunities* criterion was removed, because teachers can choose to use the assessment opportunities as formative or summative assessment. In the pictures branch the crite-

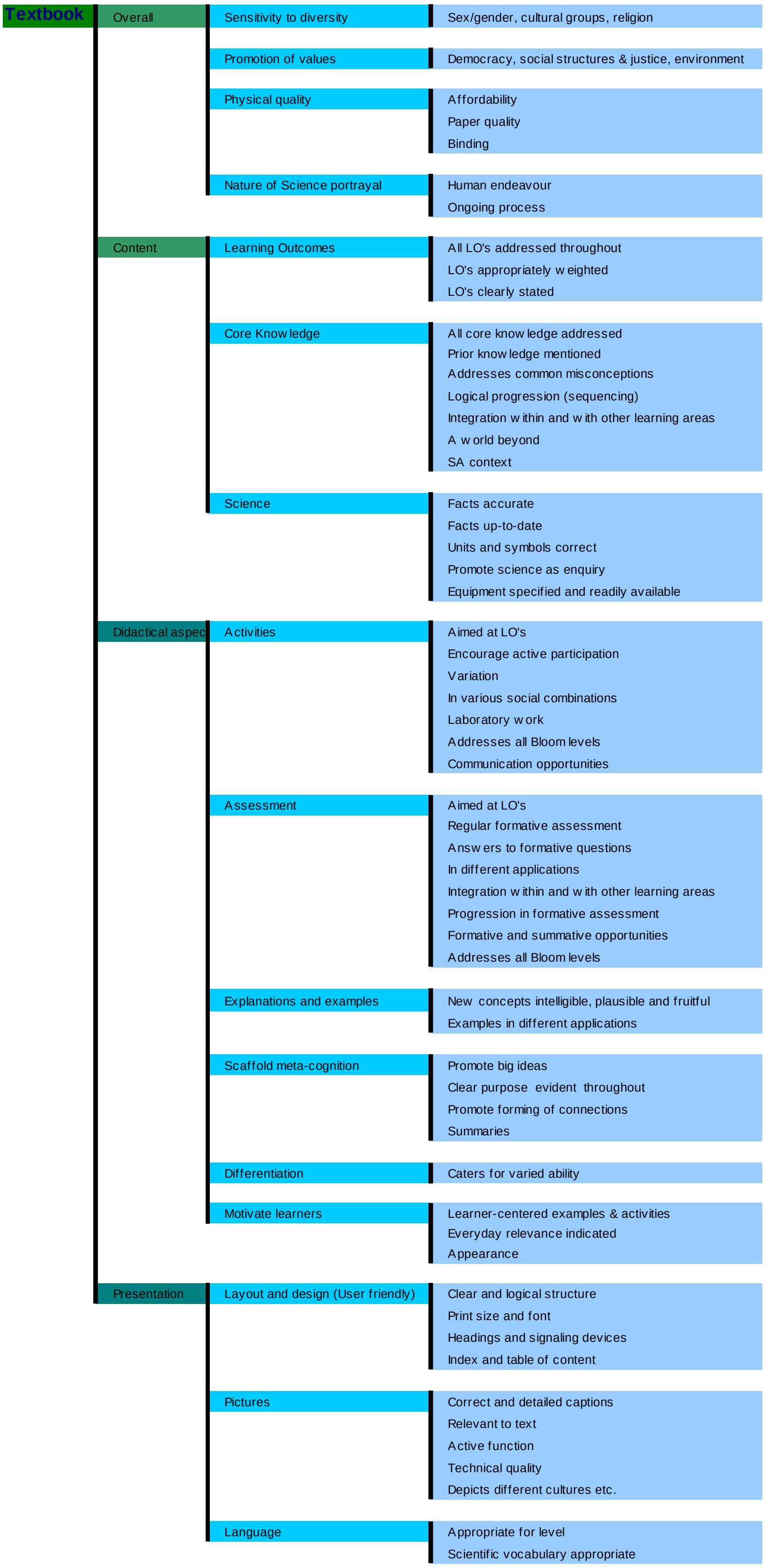


Figure 5.1: Initial hierarchy of the characteristics that influence textbook quality with 58 criteria in Level 3

rion "depict different cultures etc." was removed, because it is represented by the "sensitivity to diversity" branch. The final hierarchy is depicted in Figure 5.2.

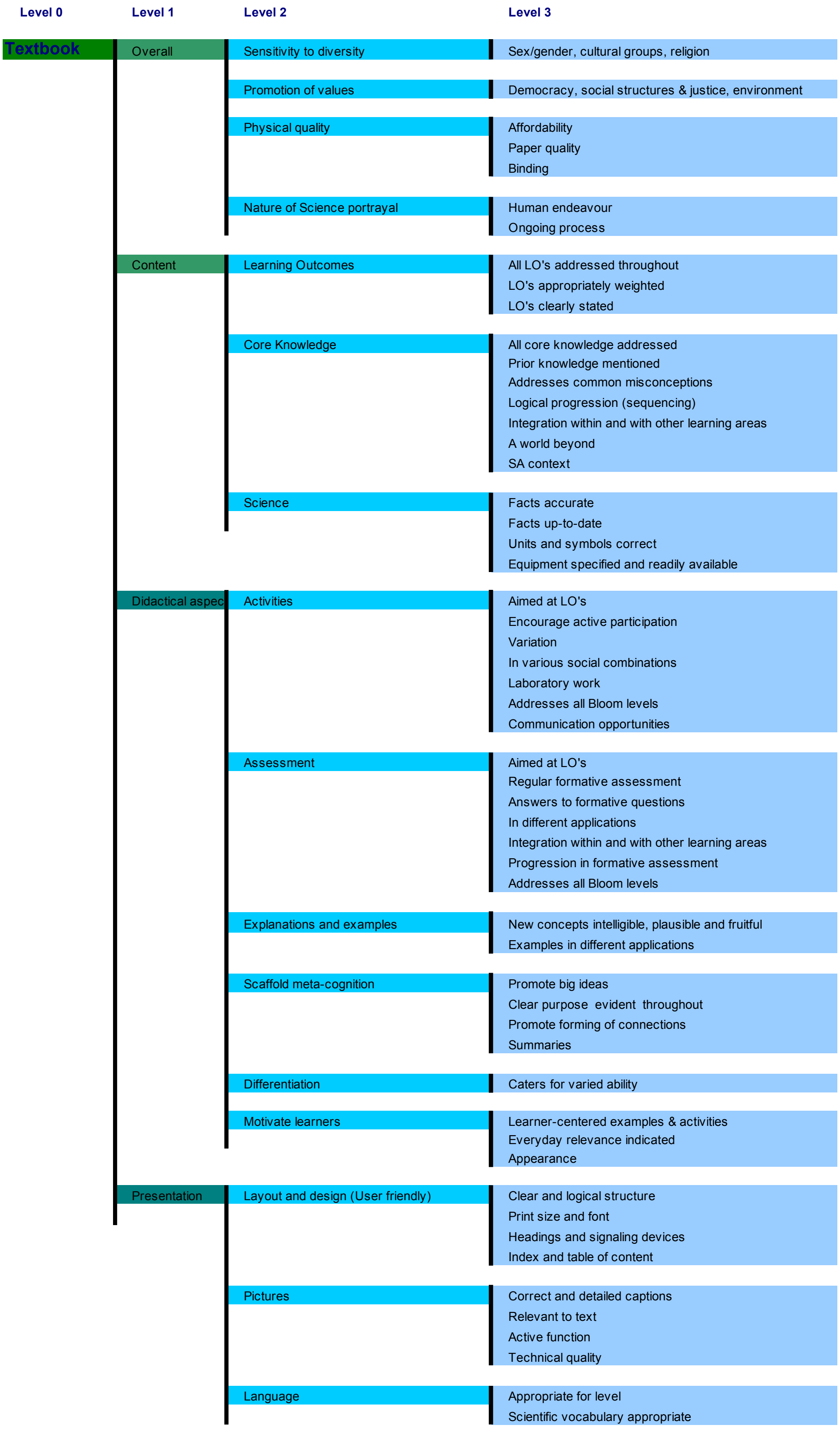


Figure 5.2: Final hierarchy of the characteristics that influence textbook quality with regard to science textbooks in the South African context

Level 1 incorporates four branches: the overall quality, the quality of the content, the didactical aspects and the presentation. Each of these three branches divides into three or more smaller branches on Level 2.

The *overall* quality is determined by the textbook's sensitivity to diversity, its promotion of values, the physical quality and its portrayal of the nature of science. The first three of these can be found in all of the existing textbook evaluation instruments. The last one is more specific to science and not expected in any of the generic instruments. It is present in the instruments used by the Western Cape and US Departments of Education and the inclusion of this branch is fully motivated by the results of the literature survey discussed in Section 2.2.1.

The quality of the *content* on the first level is broken down into the quality of the learning outcomes, the core knowledge and the scientific aspects on the second level. The quality of the core knowledge (Level 2) in the textbook is derived from 7 criteria on Level 3, all present in existing instruments. The contribution of the handling of prior knowledge and misconceptions were discussed in Section 2.2.4.1. Integration, the emphasis on South African context and the introduction of learners to a world beyond their own, is specifically addressed in the policy documents of the South African Department of Education.

The quality of the *didactical aspects* is subdivided into six categories on Level 2. The learner is guided by the explanation and examples, the activities and the assessment to master the core knowledge and reach the desired outcomes. Each of these elements can incorporate strategies to motivate learners, scaffold metacognition and provide differentiation opportunities.

The quality of the *presentation* on Level 1 is influenced by the quality of the language, pictures and layout included on Level 2. Each of these is influenced by a number of aspects reflected by the criteria on Level 3.

The hierarchy of criteria that was created to represent the aspects that influence textbook quality is not the only hierarchy. The validity of this hierarchy was confirmed by two subject experts. They confirmed that all relevant criteria are included, that none of the criteria that have been included are irrelevant and that the ordering of the criteria in the hierarchy is a valid representation of relationships between the criteria. The criteria are not equally important for textbook quality and the next stage of the development addressed this issue.

5.4 Deriving weights according to the AHP

To take the relative contributions of individual criteria to the quality of a textbook into account, the relative weights of these criteria in the overall result of the evaluation were determined. This stage of the process was performed in accordance with the AHP. To determine the weights of the criteria a panel of eight experts were used. Two separate sessions with groups of four were convened. The sessions lasted two to three hours and were facilitated by a decision scientist. The panels were asked to do pairwise comparisons of the importance of different criteria to the overall quality. Each criterion was compared to every other criterion in the same branch on the same level. The complete questionnaire is included in Appendix C. Every member of the panel was provided with

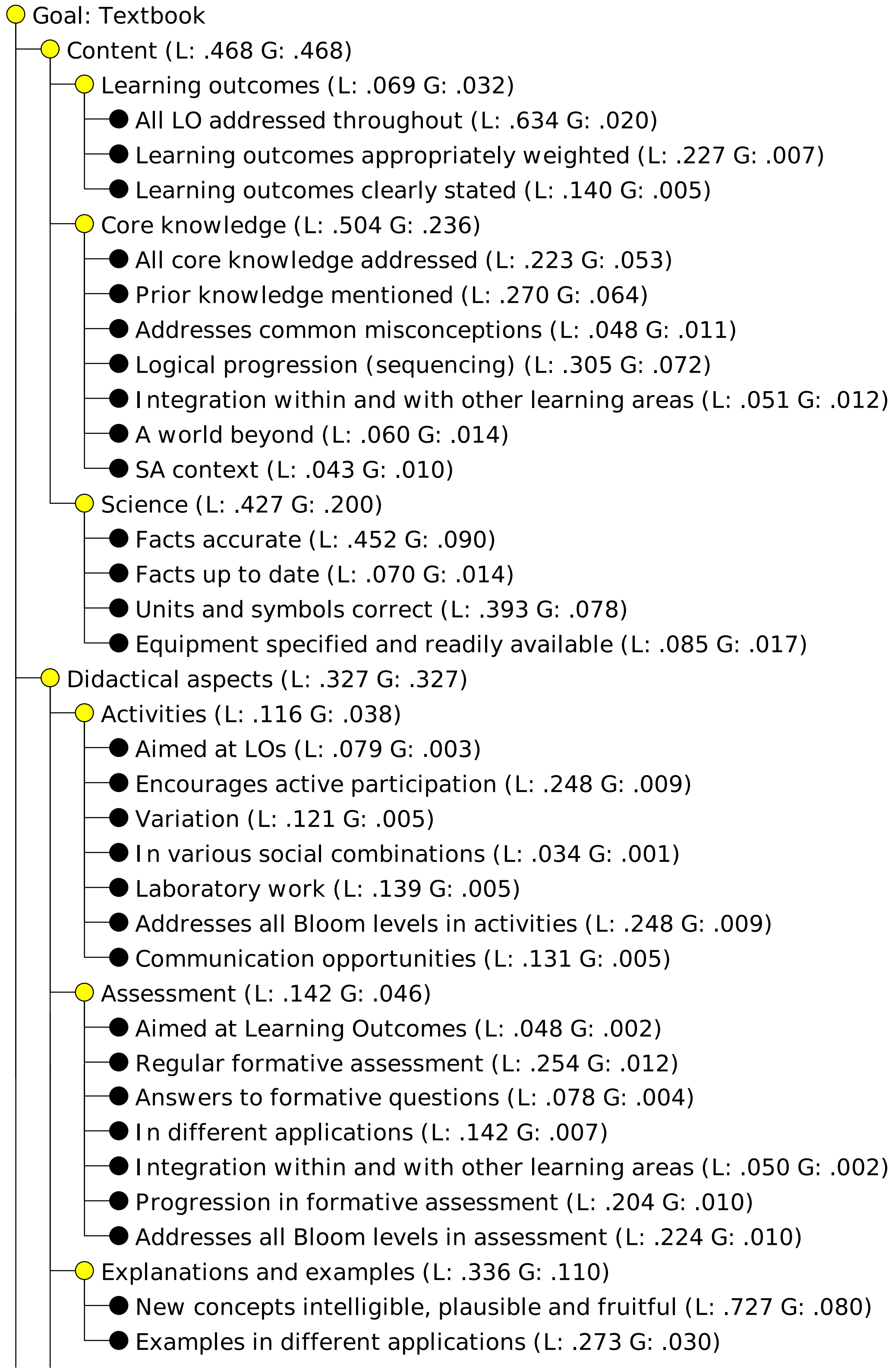
- the hierarchy of criteria to orientate them;
- the questionnaire that consisted of 16 pages with all the comparisons; and
- the rubric that shows exactly what every criterion entails.

Respondents used a relation scale of real numbers (0–5) to systematically assign preferences. Each judgement was expressed as the ratio of the importance of one criterion compared to another criterion. The respondents indicated the judgements on the questionnaire. The questionnaire is included in Appendix C. The data were imported into Expert Choice, a commercial computer program that is specifically designed to calculate weights according to the AHP.

The data were recorded in a matrix in the Expert Choice computer application, with both the columns and rows representing the individual criteria. The computer used an eigenvalue based mathematical method to determine both the relative priority of each criterion for the group as well as an "inconsistency index" for every member of the panel and the combined data. The inconsistency index is a measure of how consistent the data source is in the comparisons. If criterion A is considered twice as important as B and B is twice as important as C, then, to be consistent, A should be four times as important as C. The AHP allows for inconsistency and it even provides a *measure of the consistency* of the data (Winston, 2004). The inconsistency indices for every individual and the group as a whole were found to be lower than 10% or 0,1 which is acceptable in the AHP (Saaty, 1987).

The weights that were derived are reported in Figure 5.3. Both the contribution to the local branch (L) and to the quality of the textbook as a whole are shown.

It is interesting to note that the these expert teachers already indicated in September 2009 that they considered the quality of the presentation of the learning out-



(a) First part – continued on next page

Figure 5.3: The hierarchy of criteria with the weights that represent their contributions to the quality of the textbook as a whole (G) and to the relevant branch (L)

- Scaffold metacognition (L: .199 G: .065)
  - Promote big ideas (L: .094 G: .006)
  - Clear purpose evident throughout (L: .235 G: .015)
  - Promote forming of connections (L: .281 G: .018)
  - Summaries (L: .390 G: .025)
- Differentiation (L: .147 G: .048)
  - Caters for varied ability (L: 1.000 G: .048)
- Motivate learners (L: .059 G: .019)
  - Learner centred examples and activities (L: .290 G: .006)
  - Everyday relevance indicated (L: .404 G: .008)
  - Appearance (L: .306 G: .006)

- Overall (L: .069 G: .069)
  - Sensitivity to diversity (L: .060 G: .004)
    - Sex, gender, culture and religion (L: 1.000 G: .004)
  - Promotion of values (L: .205 G: .014)
    - Democracy, social structures and justice, environment (L: 1.000 G: .014)
  - Physical quality (L: .393 G: .027)
    - Affordability (L: .261 G: .007)
    - Paper quality (L: .318 G: .009)
    - Binding (L: .421 G: .011)
  - Nature of Science portrayal (L: .342 G: .024)
    - Human endeavour (L: .386 G: .009)
    - Ongoing process (L: .614 G: .014)
- Presentation (L: .136 G: .136)
  - Layout and design (L: .336 G: .046)
    - Clear and logical structure (L: .288 G: .013)
    - Print size and font (L: .176 G: .008)
    - Headings and signalling devices (L: .344 G: .016)
    - Index and Table of Content (L: .192 G: .009)
  - Pictures (L: .196 G: .027)
    - Correct and detailed captions (L: .150 G: .004)
    - Relevant to text (L: .186 G: .005)
    - Active function (L: .418 G: .011)
    - Technical quality (L: .245 G: .007)
  - Language (L: .468 G: .064)
    - Appropriate for level (L: .242 G: .015)
    - Scientific vocabulary appropriate (L: .758 G: .048)

(b) Second part – continued from previous page

Figure 5.3: The hierarchy of criteria with the weights that represent their contributions to the quality of the textbook as a whole (G) and to the relevant branch (L)

comes not nearly as important as the core knowledge. They maintained that the learning outcomes should be implicitly present in the way that the core knowledge is addressed, but saw no need for the textbook to state the learning outcomes explicitly. The weights derived from their comparisons indicate that the presentation of the learning outcomes only contributes 3,2% to the quality, while the core knowledge contributes 23,6%. Interestingly, in November 2009 Minister Motshekga declared OBE "dead."

In Figure 5.4 the criteria are ordered according to their contribution to the quality of the textbook as a whole. This corresponds well with the remarks by teachers, mentioning the fact that for them the accuracy of the facts is the most important criterion.

5.5 Development of a rubric for rating of criteria

The rubric is an element that is not present in the AHP. The rubric was included in the development of the IESET in order to increase the reliability of the ratings and the final results. Many of the ratings are highly subjective. This can lead to different interpretations and unreliable ratings. When analysts rate textbooks according to the rubric they will be able to give more consistent ratings and this contributes to the reliability of the instrument.

The rubric that was developed designates the indicators that must be present to satisfy every criterion. The indicators are based on the results of research on specific aspects of learning science identified during the literature survey.

The complete rubric can be found in Appendix D. The rubric is presented as a table with five columns and a row for each criterion.

The rubric is used by analysts as rating scheme and scoring sheet. They record the ratings on the rubric with a mark in the relevant column of every row. The rubric covers five A4-pages and should not appear too daunting to analysts.

5.6 Initial instrument and pilot implementation

The initial IESET was used by four teachers to evaluate two Grade 12 Physical Science textbooks by different authors: *Fisiese Wetenskappe Graad 12 (Teorie en Werkboek)* (Olivier, 2009) and *Fisiese Wetenskappe Graad 12* (Lucas, 2008)). The textbooks were chosen because they are readily available, are used by many

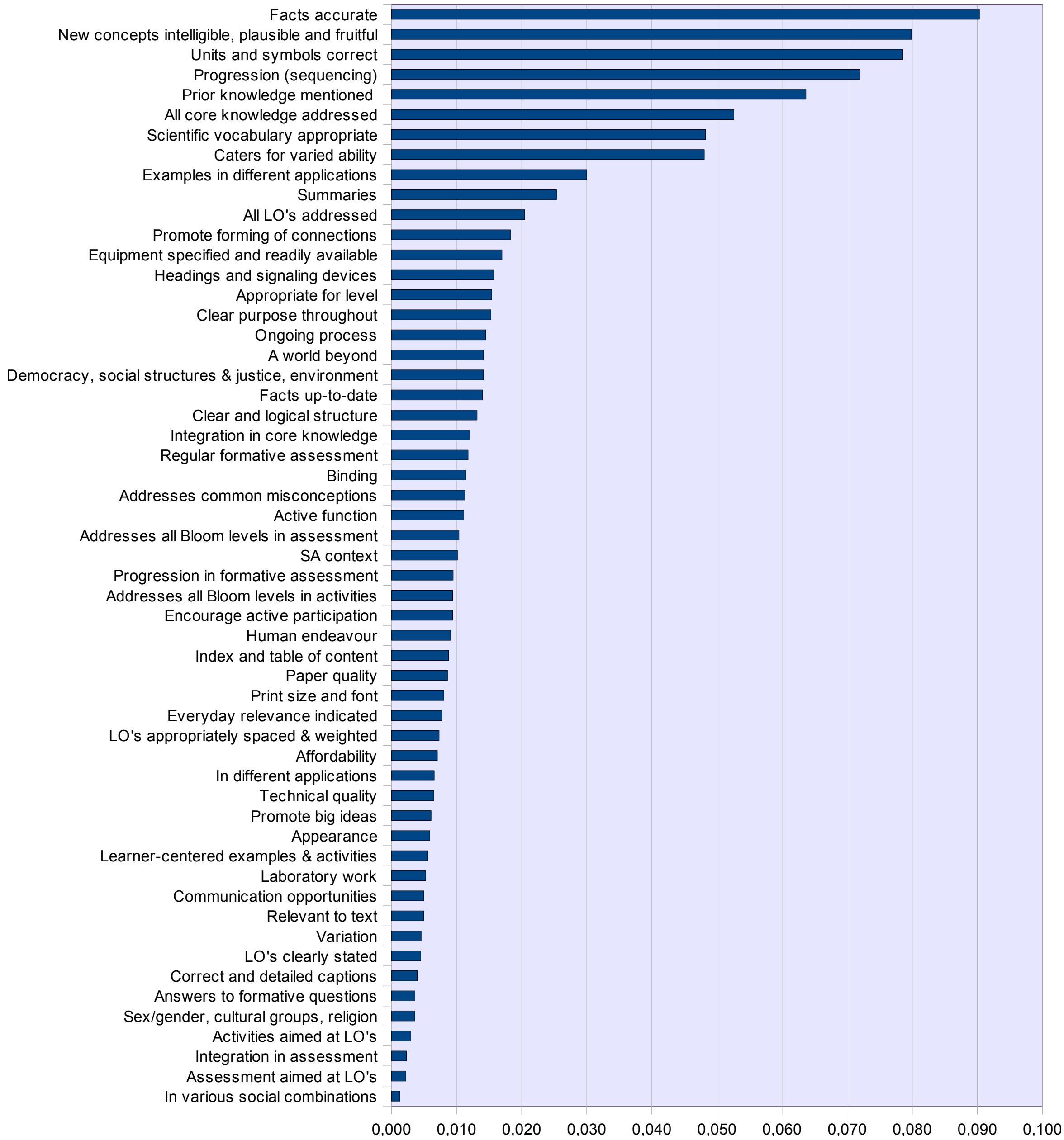


Figure 5.4: The criteria of the IESET ordered according to their weights or their contributions to the overall quality of the textbook

schools in the Pretoria area and they satisfied the screening criteria. The teachers were well qualified, with at least four year university degrees with chemistry and/or physics as major. All had more than ten years of experience in science teaching. No training in textbook assessment was offered. Only a brief explanation of the use of the IESET was provided. The four teachers did the evaluations on separate occasions.

To compute the result for the group it was necessary to aggregate the values in the data set to obtain a number that in some sense represents the "middle" of the set of values. For different interpretations of the term "middle" one resorts to different measures of central tendency. Although the *arithmetic mean* is the best known of the different measures of central tendency, it is not the preferred measure when the aim is to give a "typical" or representative value. In a small sample the arithmetic mean is very sensitive to outliers, and hence is susceptible to manipulation. This is especially relevant when one tries to aggregate value judgements. A single highly positive or highly negative judgement can change the average significantly.

Two common alternative measures of central tendency are used to overcome the problem of sensitivity to outliers. The first is the so-called *trimmed mean* where a fixed percentage of both the highest and lowest values are discarded. This is a good approach when one works with large data sets, since with an appropriate choice of the percentage, one can almost surely eliminate the effect of outliers. For smaller data sets, however, this is not a practical approach, since it discards a significant proportion of the information available.

The second approach is to use the *median*. (The median is simply the middle value of the data set when it is ranked in order.) It has the advantage that it gives a measure of central tendency that takes all measurements into consideration, but does not depend on the magnitude of individual measurements. It is hence much less susceptible to manipulation (Rubin, 2009; Vercellis, 2009; Dietz and Kalof, 2009). In normally distributed data with no or balanced outliers the arithmetic mean and median will coincide. The median of the ratings from the different analysts were, therefore, used to represent the combined rating for the group.

The pilot application scores for the textbook as a whole and the branches on the first level are represented in Table 5.1. The table shows the individual teachers' scores (A to D), as well as the scores for the group (derived from the median of the ratings). The results are also depicted as small graphs in Figure 5.5. The graphs present trends and variations in the scores on Level 1 for the two textbooks in the pilot application of the IESET. This visual representation provides a quick comparison of the teachers' evaluation results.

The results obtained by the various analysts compared well. All four analysts (and the group) ranked the two textbooks in the same order with regards to the quality of the textbook and the branches on the second level. This suggested that the IESET results are reliable.

The teachers were asked for comments on the process and the wording of the rubric. In general the teachers reported that they found the instrument easy to use. On their recommendation some of the descriptions of indicators were changed to make them clearer. For example, the criterion "All core knowledge

Branch	Median	A	B	C	D
Textbook Olivier	77,47	72,76	72,15	81,98	80,94
Textbook Lucas	63,34	62,70	60,39	66,54	67,43
Overall Olivier	69,75	76,25	66,67	73,07	77,67
Overall Lucas	59,13	57,57	60,69	69,00	63,09
Content Olivier	81,66	72,67	70,64	84,45	88,74
Content Lucas	69,73	69,39	62,96	73,49	75,46
Didactics Olivier	75,80	73,62	76,56	81,34	74,39
Didactics Lucas	54,14	52,55	54,53	55,37	58,08
Presentation Olivier	70,97	69,26	69,49	79,52	71,46
Presentation Lucas	65,60	66,70	65,48	68,19	64,50

Table 5.1: The results for Levels 1 and 2 of the pilot application of IESET by four teachers (A to D)

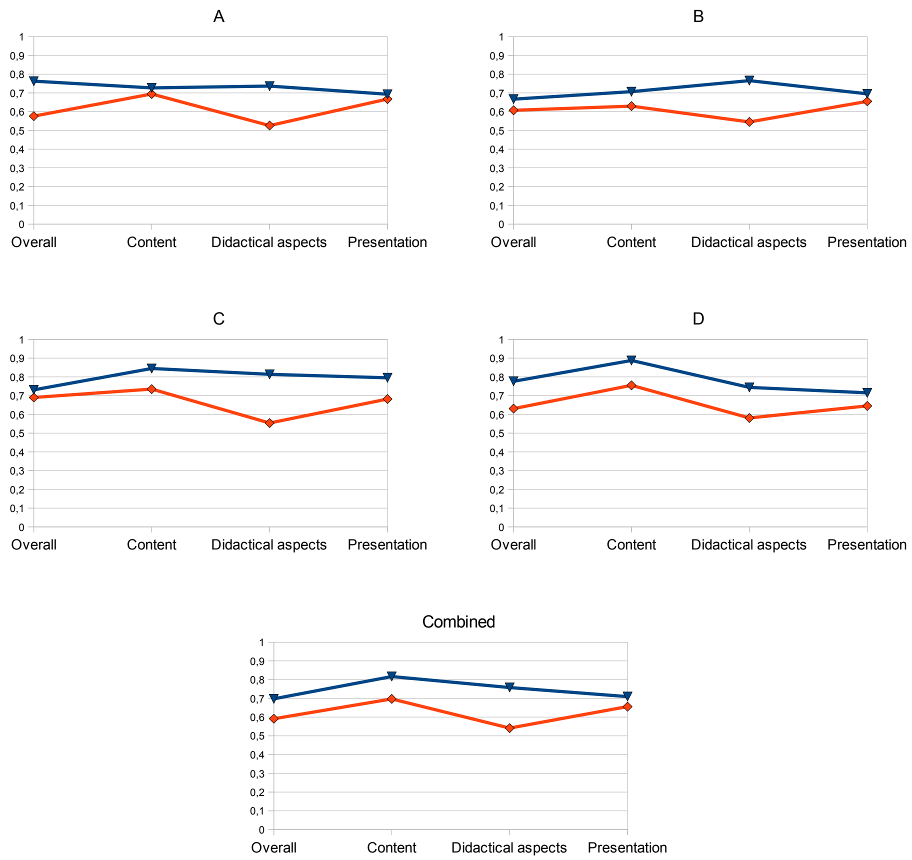


Figure 5.5: Visual representation of the pilot application results on Level 2 for every analyst and group: Lucas (2008) in red and Olivier (2009) in blue

included" the description of the third column was changed from "Include almost all of the core knowledge" to "All topics present, but a few details are not included." Distinguishing between topics and details provides an easier method to identify the presence of the core knowledge. The rubric in Appendix D is the final improved version.

The transfer of ratings from the rubrics to the data sheets proved an easy process and the computation time was negligible.

5.7 Final IESET and implementation

The adapted instrument was applied by twelve experienced teachers from various schools in Pretoria: Menlopark High School (3 teachers), Afrikaans Hoër Seunskool (2 teachers), Afrikaans Hoër Meisieskool (1 teacher), Cornwall Hill College (2 teachers), Pretoria Girls High School (2 teachers), private (2 teachers). Many of the teachers are heads of their departments and all of them are well qualified and have many years experience in science education, especially in the FET band.

The teachers were provided with the two textbooks and a rubric. The same two textbooks used in the pilot application were also used in the larger scale application. The teachers were not provided with any training in the use of the IESET. They received only brief explanations and instructions, but all their questions were answered. Due to their busy schedules, the teachers opted to complete the evaluations at their schools on separate occasions. The researcher was present at all the evaluations in order to answer questions as they arose.

The time needed by the teachers to complete the evaluation of a textbook varied between 15 minutes and one hour. Some teachers were familiar with both books and completed the evaluation without opening the textbooks, while others needed to look at the textbook on a few occasions. Some teachers arrived at judgements quickly, while others spent some time deliberating their decisions.

After the evaluations every teacher's ratings were transferred from the rubrics to a spreadsheet that calculated the scores for the quality of the textbook regarding every branch and the textbook as a whole. The teachers' ratings for the individual criteria were also combined on a single spreadsheet that calculated the median for every rating and used it to compute the scores for the textbook and also for every branch.

Figure 5.6 and Figure 5.7 show the ratings of the twelve teachers (A to L) for every criterion, as well as the median (discussed in Section 5.6). The figures also

Lucas	Level 1	Level 2	Level 3	Median	A	B	C	D	E	F	G	H	I	J	K	L
Textbook 66,40%	Overall 70,75%	Sensitivity to diversity 67%	Sex/gender, cultural groups, religion	2	3	1	2	2	2	2	3	3	2	2	2	2
		Promotion of values 67%	Democracy, social struct. & justice, environment	2	3	2	2	2	2	3	2	3	2	2	2	2
		Physical quality 86%	Affordability	3	3	3	3	3	2	3	3	3	3	3	3	3
			Paper quality	3	2	3	3	3	3	3	3	2	3	3	2	3
			Binding	2	1	1	2	3	2	2	2	2	1	2	2	2
		Nature of Science portrayal 56%	Human endeavour	2	1	2	2	1	2	1	2	2	1	1	2	2
			Ongoing process	1,5	2	1	2	1	2	1	2	1	1	1	2	2
	Content 71,29%	Learning Outcomes 67%	All LO's addressed throughout	2	1	2	3	1	2	3	2	2	2	3	2	2
			LO's appropriately weighted	2	1	1	1	2	2	2	2	2	1	2	2	2
			LO's clearly stated	2	1	2	2	2	2	1	1	2	2	1	2	2
		Core Knowledge 56%	All core knowledge addressed	2	2	2	2	3	2	2	2	3	2	2	3	1
			Prior knowledge mentioned	2	3	2	2	2	2	2	2	2	3	2	2	1
			Addresses common misconceptions	1,5	2	2	1	1	2	2	2	2	1	1	1	1
			Logical progression (sequencing)	1,5	2	2	2	1	2	1	2	2	1	1	1	1
			Integration within and with other learning areas	1	2	1	2	2	2	1	2	1	1	1	1	1
			A world beyond	1	1	1	2	1	2	1	2	1	1	1	1	2
			SA context	1,5	2	1	1	1	2	2	2	1	1	1	2	2
		Science 90%	Facts accurate	2,5	2	3	2	3	3	2	3	2	3	2	3	2
			Facts up-to-date	3	2	2	3	3	2	3	3	3	3	3	3	3
			Units and symbols correct	3	3	3	2	3	2	3	3	3	3	2	3	2
			Equipment specified and readily available	2	2	2	2	2	2	2	2	2	2	2	3	3
	Didactical aspects 57,23%	Activities 74%	Aimed at LO's	2	2	2	1	1	3	2	2	2	1	2	2	2
			Encourage active participation	3	2	3	3	3	3	3	3	2	2	2	2	3
			Variation	2	2	1	2	2	2	2	2	1	1	2	2	1
			In various social combinations	1	0	0	0	1	1	1	1	1	0	0	1	1
			Laboratory work	2	2	2	1	3	2	2	3	2	2	2	3	2
			Addresses all Bloom levels	2	2	2	2	2	2	2	2	2	2	2	2	2
			Communication opportunities	2	2	2	2	2	2	2	2	1	1	1	2	1
		Assessment 68%	Aimed at LO's	2	2	3	1	2	2	2	2	2	1	2	2	2
			Regular formative assessment	2	3	2	1	2	2	2	2	2	1	2	2	2
			Answers to formative questions	2,5	2	3	3	2	2	3	2	3	3	2	2	3
			In different applications	2	2	2	2	2	2	2	2	2	2	2	2	2
			Integration within and with other learning areas	2	1	2	2	2	2	2	2	2	2	2	2	2
			Progression in formative assessment	2	1	2	2	2	1	2	2	2	2	2	2	2
			Addresses all Bloom levels	2	1	1	1	2	2	2	2	2	2	2	2	2
		Explanations and examples 67%	New concepts intelligible, plausible and fruitful	2	1	2	2	1	2	2	2	2	2	2	2	2
			Examples in different applications	2	2	1	2	1	2	2	2	2	2	2	1	2
		Scaffold meta-cognition 39%	Promote big ideas	1,5	1	2	1	2	2	1	2	1	1	1	2	2
			Clear purpose evident throughout	2	2	2	3	2	3	3	2	3	2	2	2	2
			Promote forming of connections	2	2	2	2	2	2	3	2	3	2	2	3	2
			Summaries	0	0	1	0	1	0	0	0	0	0	0	0	1
		Differentiation 33%	Caters for varied ability	1	1	1	1	1	2	2	1	2	1	1	2	2
		Motivate learners 67%	Learner-centered examples & activities	2	2	2	2	3	3	2	2	2	2	2	2	2
			Everyday relevance indicated	2	3	2	2	3	2	2	2	2	1	2	2	2
			Appearance	2	2	2	2	2	3	2	3	3	2	1	2	2
	Presentation 69,43%	Layout and design (User friendly) 60%	Clear and logical structure	2	3	2	2	2	3	3	2	2	2	2	2	2
			Print size and font	2	3	2	2	3	2	3	3	2	2	2	2	2
			Headings and signaling devices	2	1	1	2	3	2	3	2	2	2	2	2	2
			Index and table of content	1	1	1	1	1	1	1	1	1	1	1	1	1
		Pictures 92%	Correct and detailed captions	3	2	3	3	3	3	3	3	3	3	3	2	3
			Relevant to text	3	3	3	3	3	3	3	3	3	3	3	2	3
			Active function	3	3	3	3	3	2	2	3	3	3	3	2	3
			Technical quality	2	1	2	2	3	2	2	2	2	3	2	2	2
		Language 67%	Appropriate for level	2	2	2	2	3	3	2	3	2	2	2	2	2
			Scientific vocabulary appropriate	2	2	2	3	2	2	3	3	3	2	2	2	2

Figure 5.6: The combined evaluation results for Lucas (2008)

show the combined scores for the quality of the book as a whole and for every branch.

The combined results as well as the separate results of the twelve teachers (A to L) for the quality of the textbook as a whole as well as the results of branches on Level 1 are summarised in Table 5.2.

Although the individual ratings varied, the evaluation results of all teachers indicated that the textbook by Olivier (2009) is of higher quality than the textbook by Lucas (2008) (as whole and in the branches of Level 1). This verifies the reliability of the instrument.

Olivier	Level 1	Level 2	Level 3	Median	A	B	C	D	E	F	G	H	I	J	K	L
Textbook 82,39%	Overall 81,08%	Sensitivity to diversity 67%	Sex/gender, cultural groups, religion	2	3	2	2	2	2	2	2	3	2	2	2	2
		Promotion of values 100%	Democracy, social struct. & justice, environment	3	3	2	3	2	3	3	2	3	2	2	3	3
		Physical quality	Affordability	3	2	2	3	3	3	3	3	2	2	3	3	2
		86%	Paper quality	3	2	3	3	3	3	3	2	2	3	3	3	3
			Binding	2	2	2	2	3	2	3	2	2	2	2	2	2
		Nature of Science portrayal	Human endeavour	2	3	2	2	2	2	2	3	3	2	3	2	2
		67%	Ongoing process	2	3	2	2	2	2	2	3	3	2	2	2	2
	Content	Learning Outcomes	All LO's addressed throughout	3	2	2	3	3	2	3	3	3	2	3	3	3
	84,19%	96%	LO's appropriately weighted	2,5	2	2	2	3	2	3	3	3	2	3	3	2
			LO's clearly stated	3	3	3	3	3	3	3	3	3	3	3	2	3
		Core Knowledge	All core knowledge addressed	3	2	3	3	3	3	3	3	3	3	3	3	3
		83%	Prior knowledge mentioned	3	3	3	3	3	2	3	3	2	3	2	2	3
			Addresses common misconceptions	2	2	1	2	2	1	2	2	2	1	2	1	2
			Logical progression (sequencing)	2	3	3	2	2	2	2	3	2	3	3	2	2
			Integration within and with other learning areas	2	2	1	2	2	2	2	3	2	2	1	2	2
			A world beyond	2	2	1	2	2	2	2	2	2	1	2	2	2
			SA context	2	2	1	1	2	2	2	2	2	2	1	2	2
		Science	Facts accurate	2	2	2	2	1	2	2	2	2	2	2	2	2
		84%	Facts up-to-date	3	3	2	3	3	3	3	3	3	2	3	2	3
			Units and symbols correct	3	3	2	2	3	3	3	3	3	3	2	3	2
			Equipment specified and readily available	2,5	3	3	2	3	2	2	3	3	2	2	2	3
	Didactical aspects	Activities	Aimed at LO's	3	2	2	3	3	3	3	3	3	2	3	2	3
	75,44%	89%	Encourage active participation	3	3	3	3	3	3	3	3	2	3	2	2	3
			Variation	2	2	2	2	2	1	2	2	2	2	2	2	2
			In various social combinations	1	0	0	1	1	1	1	1	1	1	0	1	1
			Laboratory work	3	3	2	3	3	2	3	3	3	3	2	3	3
			Addresses all Bloom levels	3	3	2	2	3	3	2	3	3	2	3	3	2
			Communication opportunities	2	2	2	2	3	2	3	3	3	2	2	2	3
		Assessment	Aimed at LO's	3	2	2	3	3	3	3	3	3	3	2	2	3
		91%	Regular formative assessment	3	3	2	3	3	3	3	3	3	3	3	3	3
			Answers to formative questions	1	1	2	1	1	1	1	1	1	1	1	1	2
			In different applications	2,5	3	2	3	2	2	3	3	2	2	3	2	3
			Integration within and with other learning areas	2	2	2	3	2	2	3	3	2	2	2	2	2
			Progression in formative assessment	3	3	2	3	3	2	3	3	3	2	2	3	3
			Addresses all Bloom levels	3	3	2	3	3	3	2	3	3	2	3	3	2
		Explanations and examples	New concepts intelligible, plausible and fruitful	2,5	3	2	2	3	3	3	3	2	2	3	2	2
		79%	Examples in different applications	2	3	3	2	3	2	3	3	2	2	2	2	2
		Scaffold meta-cognition	Promote big ideas	2	2	2	1	2	3	3	3	3	2	2	2	2
		53%	Clear purpose evident throughout	3	3	2	3	2	2	3	3	3	2	3	2	3
			Promote forming of connections	2,5	2	2	3	2	2	3	3	3	2	3	2	3
			Summaries	0	0	0	0	0	0	0	0	0	0	0	0	0
		Differentiation 67%	Caters for varied ability	2	2	1	2	1	2	2	2	2	1	2	2	2
		Motivate learners	Learner-centered examples & activities	2	3	2	2	3	2	3	3	3	2	2	2	2
		90%	Everyday relevance indicated	3	3	2	3	3	2	3	3	3	2	2	2	3
			Appearance	3	3	2	3	2	2	3	3	3	2	3	3	2
	Presentation	Layout and design (User friendly)	Clear and logical structure	3	3	2	3	3	3	3	3	3	2	3	2	3
	93,59%	87%	Print size and font	3	3	2	3	3	2	3	3	2	2	3	3	3
			Headings and signaling devices	3	3	2	3	2	3	3	3	2	3	2	2	3
			Index and table of content	1	1	2	1	1	2	1	1	1	1	1	2	1
		Pictures	Correct and detailed captions	2,5	3	2	2	3	3	3	2	3	2	2	2	3
		89%	Relevant to text	3	3	3	3	3	3	3	3	3	2	3	3	3
			Active function	3	3	3	3	3	3	3	3	3	3	2	2	3
			Technical quality	2	2	2	2	3	2	3	3	2	2	2	2	3
		Language	Appropriate for level	3	3	3	3	3	3	3	2	3	2	3	3	2
		100%	Scientific vocabulary appropriate	3	3	2	3	3	2	3	3	3	3	2	3	3

Figure 5.7: The combined evaluation results for Olivier (2009)

Visual representation of the results makes it easier to compare the two textbooks. Therefore, the results of the twelve teachers for the branches on Level 1 are also visually represented by the small graphs in Figure 5.8. For example, the trends depicted in the sparklines show that books can be improved with regard to their didactical aspects. By comparing the lines for the two books it is obvious that Olivier (2009) consistently earned higher scores on all the branches of Level 1.

Branch	Median	A	B	C	D	E	F	G	H	I	J	K	L
Textbook Olivier	82,39	85,09	71,32	77,61	79,41	76,80	85,61	88,64	79,46	74,84	76,97	75,00	77,98
Textbook Lucas	66,40	63,87	66,67	65,61	66,93	70,16	69,66	73,37	71,80	65,97	59,82	69,41	60,75
Overall Olivier	81,08	86,90	70,83	81,08	79,77	81,08	86,60	81,49	86,90	70,83	78,65	81,08	77,67
Overall Lucas	70,75	69,00	59,74	74,25	68,37	70,83	69,68	76,25	71,92	57,34	62,85	70,09	74,25
Content Olivier	84,19	84,45	78,22	77,01	78,62	76,52	83,85	91,04	80,52	84,18	77,26	77,19	78,94
Content Lucas	71,29	73,49	75,59	67,07	74,32	73,10	67,41	79,37	74,42	75,19	60,29	76,86	53,81
Didactics Olivier	75,44	81,87	60,28	71,29	75,31	73,94	83,75	85,77	72,70	58,95	77,16	66,25	70,89
Didactics Lucas	57,23	47,23	55,82	55,19	50,34	64,32	65,42	58,49	64,09	52,50	54,53	60,64	64,24
Presentation Olivier	93,59	94,08	74,37	93,10	91,83	82,43	95,68	90,92	88,25	82,91	74,69	85,44	91,90
Presentation Lucas	69,43	68,19	65,58	81,25	80,63	73,70	87,57	87,00	81,25	71,03	69,43	64,50	69,43

Table 5.2: Results of the application of the IESET on every textbook as a whole and on the second level

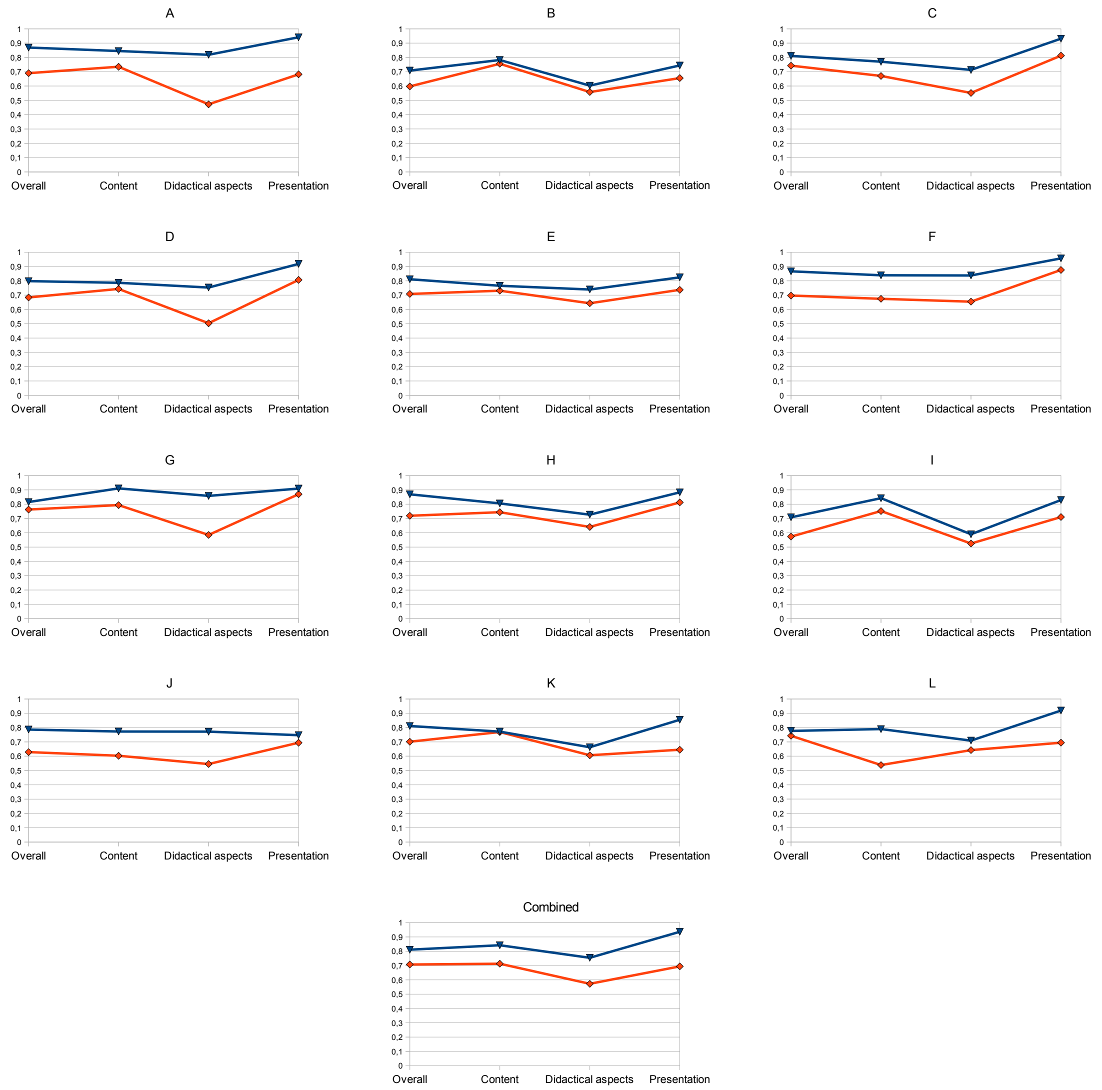


Figure 5.8: Visual representation of the results on Level 2 for every analyst and group: Lucas (2008) in red and Olivier (2009) in blue

The visual representation provides easy access to information about the textbook that teachers can use during textbook selection. Figure 5.9 shows that Olivier (2009) is regarded by the group as superior in every branch on Level 1. A teacher with limited core knowledge should choose this textbook for the quality of its content. A teacher with limited experience, but good core knowledge should choose this textbook for the quality of its didactical aspects. The combined results for the separate categories on Level 1 are displayed visually in Figure 5.9.

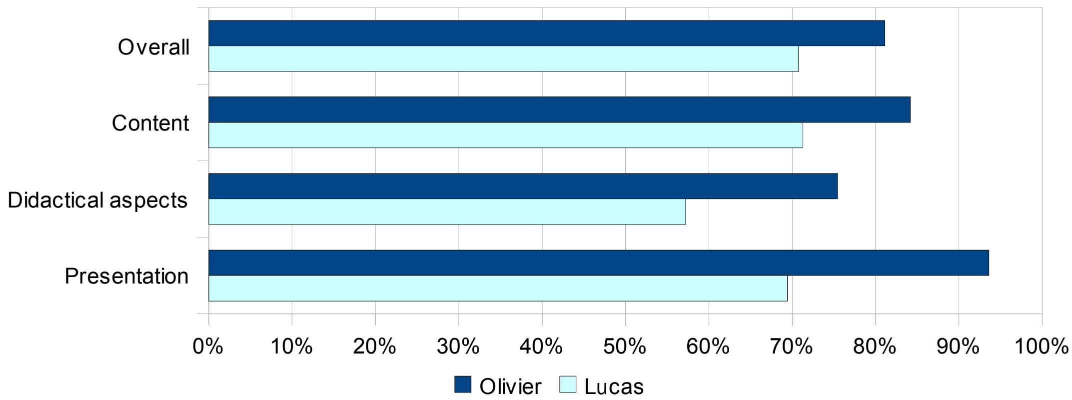


Figure 5.9: The combined evaluation results of Level 1 for Olivier (2009) and Lucas (2008)

The results for Level 2 are presented in Figure 5.10. Although Lucas (2008) obtained higher scores in the quality of the *science* aspects of its content, this branch did not contribute enough to the *content* branch in Level 1 to ensure that it is regarded as the best of the books with regard to the quality of the content. The science aspects contributes 42,7% to the content, while the core knowledge and learning outcomes contribute respectively 50,4% and 6,9%, as shown in Figure 5.3.

Structured interviews were conducted with four of the twelve teachers who have used the textbooks and who were able to compare the quality of the textbooks in the classroom context. After the second interview no new data were gained from the interviews. In the third and fourth interviews the teachers simply repeated the first two teachers' sentiments regarding the two textbooks. At that point "saturation of data" occurred and additional interviews were considered unnecessary (Bowen, 2008). It was apparent that the teachers all considered the textbook by Olivier as superior. Therefore, the results of the textbook analysis obtained with the IESET were confirmed by the experience of the teachers. This verified the validity of the evaluation results derived from implementation of the IESET.

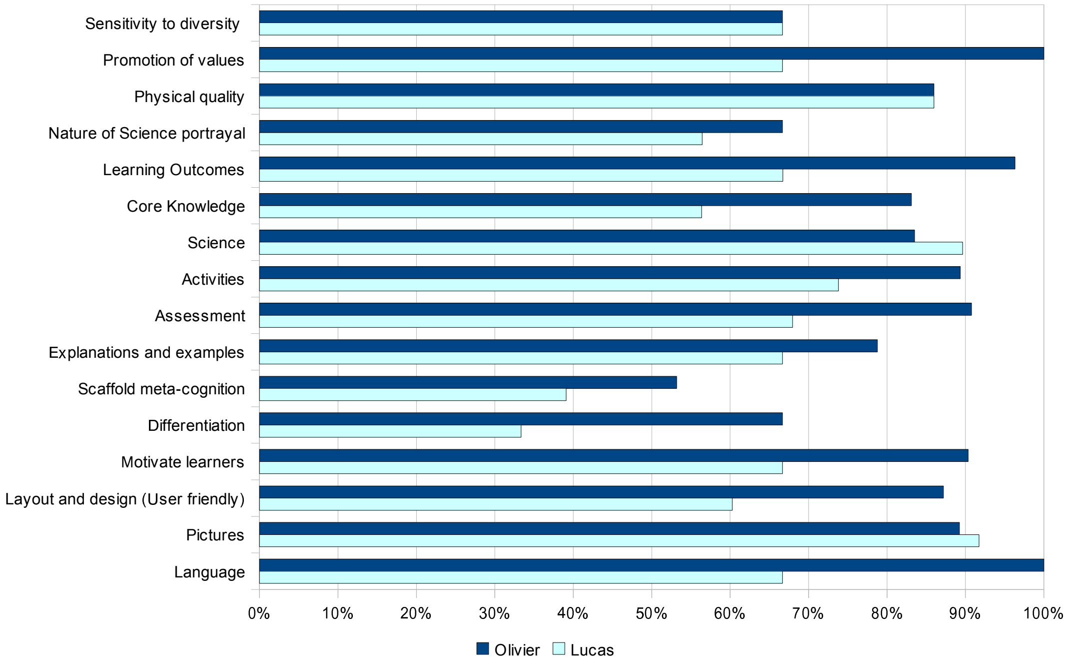


Figure 5.10: The combined evaluation results for Level 2 for Olivier (2009) and Lucas (2008)

5.8 Validity of IESET

The IESET is a valid instrument because the evaluation results correctly predict how well the textbook is able to support the learners in science education.

The following aspects are relevant to the validity of the IESET:

- The *process of criteria formulation*. Criteria were not chosen based on speculation or the opinion of the researcher, but derived from research on learning reported in the literature; from the Constitution of South Africa and education policy documents and from existing assessment instruments.

- The *content validity* of the IESET was *confirmed by two experts* in science education during semi-structured interviews. They studied the hierarchical structure of criteria and confirmed that all the aspects that influence textbook quality were present and that the hierarchical ordering was an appropriate representation of the relationships.

- The *use of weights* contribute to the validity of the IESET. Criteria that are more important to textbook quality were attributed higher weights. The final result will therefore be more valid.

- The use of *pair-wise comparisons* to derive weights. Weights were not awarded by arbitrarily *assigning* them, but through the scientifically justifiable process of *pair-wise comparisons* used by the AHP. The *consistency index* for the comparisons were within acceptable levels.

- *Triangulation of the results* of the final implementation of the IESET and the quality of the textbooks according to teachers'experience. Both the IESET and the teachers indicated the same textbook as superior. This correlation confirms the validity of the IESET.

5.9 Reliability of IESET

A reliable assessment instrument will produce the same results on re-test or with the assessment of different analysts, and will produce similar results with similar textbooks, so it is consistent in its methods and criteria.

The reliability or repeatability of the IESET was influenced or confirmed by the following:

- The procedure for the application of the IESET utilises the *same set of criteria* in every application, formulated for the specific context.
- The use of a *rubric* guides analysts in the awarding of rates. Different analysts will more consistently award similar ratings when they apply the rubric *rating scheme*, than without one.
- All the analysts individually and the median of the individual ratings indicated the same textbook as of better quality. This confirms the reliability of the IESET.

5.10 Conclusion

The hypothesis formulated in Section 4.2 has been verified. It is possible to develop a suitable instrument for the assessment of the quality of textbooks in Physical Science in the South African education system. The IESET was developed and it yielded transparent, reliable and valid evaluation results.

The research also concluded that the AHP can be successfully adapted to establish an appropriate process to be utilised in the development of a textbook evaluation instrument. The AHP ensured that the development process was fully documented and ensured transparency and accountability.

Chapter 6

Summary, conclusions and recommendations

6.1 Introduction

Science and technology are constantly changing our living environment, the things we do and the way we do it. Science education has become of vital importance to prepare every learner for life in this ever-changing context. Science education is also vital for the development of future scientists to keep the momentum of the growth of scientific knowledge and its applications in the future.

Science textbooks play an important role in science education (as shown in Chapter 2) and, therefore, the quality of the science textbooks has a direct influence on the quality of teaching and learning. The National Department of Education does a national assessment of textbooks for the FET band in order to compile a national catalogue. From this catalogue, teachers then select textbooks for use in their classrooms. A good selection process is essential to ensure that teachers use the best textbooks available. From the different approaches to textbook assessment, textbook analysis was identified (in Chapter 3) as the most appropriate approach for textbook selection.

The use of a textbook evaluation instrument contributes to the transparency and reliability of the evaluation results. The existing instruments are limited, since their focus is on other subjects, phases or contexts and not on physical science education in South Africa. The need for an appropriate textbook evaluation instrument for this context was the motivation for this research.

The research identified and explored a number of concepts (and relationships between them) in science education, textbooks and the assessment of textbook quality. The most important of these findings are mentioned in the next section.

6.2 Findings

6.2.1 Science education

The introduction indicated why science education is important. Science education in this study refers to physical science education. Research on what counts as science, what the goals of science education are and how these goals can be reached contributed to the background for this study. The following are some of the important findings:

- Science, as a discipline, focusses on reality, but scientific knowledge is not limited to our observation of reality. It also includes constructs (theories and models) that have been invented in attempts to interpret and explain observed phenomena. When new evidence or better explanations become available, scientific knowledge changes. Scientific knowledge is therefore subject to change, but it should always concur with experimental and observational evidence and should make accurate predictions, when appropriate. Science education must foster learners' understanding of this nature of science.

- The goals of science education include scientific literacy for all and the preparation of future scientists. The balance between these goals moves from greater focus on scientific literacy in earlier years to understanding and applying more advanced scientific concepts in the final school years to prepare learners for further study in scientific areas.

- In the present science education paradigm constructivism and especially conceptual change is the most prominent learning theory. The learners' prior knowledge is recognised as point of departure for teaching strategies that aim to facilitate assimilation or conceptual change in learning new concepts and relationships.

- Science education must build learners' capacity for life-long learning, especially in scientific contexts. To prepare learners to live in a world that keeps changing at a rapid pace and thus keeps requiring new knowledge and skills, learners must be able to adapt and keep learning throughout their lives. Science education should not only equip learners with scientific knowledge, but it should foster the development of learning habits, like meta-cognition and self assessment. Meta-cognition, or being aware of your own level of understanding and constantly assessing that level of understanding, will enable learners to construct their own structure of new scientific knowledge and the technological applications that the future may hold.

From these general findings regarding science education, more specific findings about science education, with special reference to South Africa, were made.

6.2.2 Science education in OBE

The introduction of OBE in South Africa changed the whole educational landscape and, therefore, the role of the textbook. A survey of the relevant literature on OBE and its implementation produced many findings of which the most important are the following:

- OBE aims towards a system where all students become learners, by using strategies that cater for learners with different learning styles and strategies. The focus is on the long-term outcomes that must be reached by using subject content. Learner-centred and activity-based learning programmes aim to keep learners engaged in the learning process and learners have to demonstrate their mastering of learning outcomes.
- OBE was introduced in South Africa as a result of political change. It was introduced as a radical change from all educational strategies and policies of the past. Outcomes were seen as more important than core knowledge. Learner-centred activity-based teaching strategies were introduced in such a way that physical engagement flourished instead of cognitive engagement. Recently the balance has moved back to focus on subject knowledge and tried and trusted universally accepted teaching strategies. OBE as it was originally introduced has died (Motshekga, 2009). Although it has not yet been replaced by a new policy, many changes have been made to the original policy.

6.2.3 Textbooks

The literature study revealed that science education has always been associated with the use of textbooks and textbooks have an important role to play in science education. The following findings in this regard deserve mention:

- Textbooks provide the content in organised and logically sequenced fashion in a single source. It provides both the core knowledge and the activities needed to understand and apply the core knowledge. The included knowledge represents the scientific knowledge deemed necessary to develop scientific literacy and provide a solid foundation for further study in scientific areas.

- The way in which the knowledge is presented in textbooks should illustrate the nature of science and facilitate conceptual assimilation or conceptual change necessary for the learners to understand and apply the concepts.

- Textbooks can support teachers in their planning and their teaching. The textbook is a *written curriculum* that links the *intended curriculum* (articulated in the National Curriculum Statements) to the *enacted curriculum* or *implemented curriculum* (that is actually experienced in the classroom). As such it can interpret the intended curriculum and help teachers to transform the intended curriculum into activities or implemented curriculum.

- With any change in educational policy, like the introduction of OBE in South Africa, textbooks can aid curriculum implementation. Unfortunately introductory OBE training and initial policy documents created the impression that learning programmes and supporting media can only be learner-centred and activity-based if the teachers design their own. This implied that the Department of Education did not approve the use of textbooks. The teachers did not know how to transform the intended curriculum into appropriate classroom activities and they were denied the support that textbooks could provide.

  The interpretation of OBE in South Africa has changed since the introduction of Curriculum 2005 (the OBE curriculum introduced in 1997). Curriculum 2005 was replaced by the Revised National Curriculum Statements and in 2010 still more changes will be effected (Department of Education of South Africa, 2009). One of the changes is the focus on textbooks. The Department of Education has recognised that it is impossible to expect from teachers to design their own supporting media. Most teachers do not have the time, the resources or the skills that are necessary for the job. Teachers should use their time to prepare and teach, using materials that were designed by teams of experts in the subject field. The Department of Education has committed itself to changing teachers' and learners' views of textbooks from 2010. The use of textbooks will be actively approved and encouraged and learners will be supplied with textbooks.

- Textbooks can not transmit knowledge. In accordance with the constructivistic learning theory, learners must construct their own knowledge. In order to do that the learners must be able to decode the text and construct their new knowledge. Text comprehension theories like the Comprehension

Integration (CI) model of Kintsch (Kintsch and Kintsch, 2005) indicate that it is important that the language in textbooks must consist of words and sentences that the learners can decode and understand. The text must be coherent to enable learners to integrate the sentences to construct the message and it must connect with the learners' prior knowledge to assist with conceptual assimilation or change.

6.2.4 Assessment of textbook quality

The literature survey offered two complimentary definitions for the quality of a textbook.

- The first definition states that a textbook is of good quality if the learners who use the textbook attain the learning outcomes they are supposed to reach.
- According to the second definition a good textbook is one that has the characteristics that will enable learners to reach the outcomes.

These two definitions link directly to the different approaches to the assessment of textbook quality. The two categories of assessment methods indicated by the definitions hinge respectively on

- determining whether the learning outcomes have been attained by learners using the textbook (first definition) or
- determining if the textbook has the characteristics that makes it fit to guide learners to attain the outcomes (second definition).

A classification of assessment methods suggested by Mikk (2000) distinguishes between:

- analysis of textbooks
- experimental investigation (measuring outcomes)
- respondent opinions

A new more comprehensive and explicit classification was developed by considering the context in which the assessment is done and the type of data gathered. This classification system was discussed in detail in Section 3.7.1.

The relevant literature revealed several merits and limitations of the various types of textbook assessment discussed in Section 3.7.2. The merits of textbook assessment (through analysis) that were identified justified the choice of this method as the focus of the empirical research in this study. The main advantages of textbook analysis are summarised in the following statements:

- Textbook analysis can *predict* the efficacy of textbooks before they are used in schools and inform the selection process.
- Textbook analysis is a *thorough and systematic* process that considers all relevant criteria.
- Textbook analysis is a *transparent process*, where the criteria and procedure are prescribed in detail (in a textbook analysis instrument) for all shareholders to examine.

The literature survey identified the following elements as important parts of any instrument or system for textbook evaluation:

- a predetermined preferred set of characteristics of a textbook
- a system within which one may ensure objective, quantified assessment
- a rating method that can provide the possibility of a comparative analysis
- a simple procedure for recording and reporting the evaluator's opinion
- a mechanism by which the universal scheme may be adapted and/or weighted to suit the particular requirements of any teaching situation
- a rating trajectory that makes possible a quick and easy display of the judgements on each and every criterion
- a graphic representation to provide a visual comparison between the evaluator's preferred choices as an archetype and their actual realisation in a particular textbook

An investigation of existing instruments that are available for textbook evaluation showed that they are not appropriate for physical science education in South Africa. An appropriate instrument would contribute to textbook selection and, therefore, to science education.

6.2.5 Development and testing of a textbook evaluation instrument for science education textbooks

The study verified that it is possible to develop a suitable instrument for the assessment of the quality of science education textbooks in the South African education system. It confirmed the notions that the AHP can be successfully adapted to derive an appropriate process for the development of a textbook evaluation instrument and that an instrument can be developed that gives transparent, reliable and valid evaluation results.

The following aspects of the research substantiated the notion that the AHP could be adapted to provide an appropriate process for the development of an instrument for textbook assessment:

- The AHP breaks the problem down into individual criteria and builds a hierarchical structure of criteria, which proved to be a natural ordering of criteria. In the AHP this is a deliberate ordering with a definite purpose, but it turned out to be a natural successor to the categorising of criteria found in all the existing assessment instruments. The process elevated criteria development from a speculative operation to a thorough, thoughtful process.

- The AHP provided a scientifically justifiable process for the determination of weights. The weights of criteria were not *assigned* arbitrarily, but were derived according to the pair-wise comparisons process dictated by the AHP. The panel of experts found it easier to compare elements in pairs than to directly assign weights to the attributes. The use of these weights increased the validity of results of the IESET. The contributions of individual criteria to the final score is in relation to its contribution to the textbook's quality.

- The AHP made it possible to combine subjective criteria and objective criteria in the assessment process.

- The elements of the AHP that were used yielded an instrument development process that can be easily applied in different contexts to develop instruments for specific contexts.

The study confirmed that an instrument can be developed that gives transparent, reliable and valid evaluation results for science textbooks in SA. The development and testing of IESET illustrated the following:

- The IESET is appropriate for *science education textbooks*. The literature survey on science education provided the landscape from which the criteria that are relevant to science education textbooks were derived.

- The IESET is appropriate for *OBE in South Africa*. The approaches to the formulation of *criteria* utilised in the development of the IESET included identification of relevant criteria from the Constitution of South Africa and education policy documents. It also formulated criteria derived from the literature survey that investigated research reports on OBE in SA and the rest of the world.

- The *transparency* of the results is situated in the availability of the weights and the rubric. Stakeholders can see exactly what criteria are used, how ratings are awarded and what the weights of the individual criteria are. The results include the scores on every criterion, every branch and the textbook as a whole.

- The *reliability* of the IESET is situated in the consistency of the ratings awarded by different analysts due to the guidance of a *rubric rating scheme*. All the analysts individually, as well as the median of the individual ratings, indicated the same textbook as of better quality. This *correlation* confirms the reliability of the IESET.

- The following aspects contributed to, or confirmed the *validity* of the IESET:

  - The IESET utilises *research-based criteria* derived from research on learning reported in the literature; from the Constitution of South Africa and education policy documents and from existing assessment instruments.

  - The *content validity* of the IESET was confirmed by two experts in science education. They confirmed that all the aspects that influence textbook quality were present and that the hierarchical ordering was an appropriate representation of the relationships between the criteria.

  - The *use of weights* contributes to the validity of the IESET. Criteria that have a greater influence on textbook quality were attributed higher weights. The final result will therefore be more valid.

  - The use of *pair-wise comparisons* to derive weights is a scientifically justifiable process and during the calculation of the weights by the Expert Choice program the *inconsistency index* was calculated by the program and found to be within acceptable levels. This confirmed the validity of the derived weights.

  - *Triangulation of the results* of the final implementation of the IESET and the quality of the textbooks according to teachers' experience confirmed the validity of the IESET results. Both the IESET and the teachers'experience identified the same textbook as being of higher quality.

6.3 Limitations

Practical considerations limited the *number of teachers* involved in the implementation of the IESET. The application was limited to 12 teachers from Pretoria.

Application on a larger scale might provide the opportunity for statistical analysis of the data and more confirmation of the reliability of the IESET.

The application was limited to two textbooks with which the teachers are *familiar*. Future investigations should include books that are unknown to the teachers in order to verify that this factor does not influence the reliability or validity of the results.

The IESET was developed for science textbooks in the South African context. Although Minister Angie Motshekga (Motshekga, 2009) declared OBE in South Africa "dead" (Motshekga, 2009), it was not replaced by a new education policy. If the learning outcomes become irrelevant in future, the current IESET might have to be amended by using the original pairwise comparison information. The criteria involving learning outcomes can be removed from the matrix and the weights recalculated with the Expert Choice program to yield a suitable instrument.

6.4 Recommendations

The IESET was implemented by well qualified experienced teachers, who were well-versed in science education and learning outcomes. To improve the reliability of the results *training* of analysts can be added to the procedure. Discussing examples of textbooks with specific ratings for every criterion during training may improve analysts' ability to judge subjective criteria.

The application of the IESET to a number of textbooks may require a lot of *time* from teachers. To make the application of the IESET less time consuming, criteria with very low global weights can probably be removed without compromising the validity of the results. Further investigation is needed to confirm this hypothesis.

The use of a *5-point rating scale* can be investigated. It may prove to discriminate better between textbooks than the 3-point rating scale. The rubric would, however, become very clumsy if two columns were added. It would, perhaps, be possible to change the rating of the current columns representing 2 and 3 to 3 and 5 respectively. The teachers can then, for example, indicate a rating of 4 by making their cross on the line between the cells representing 3 and 5.

The *median* was used to represent the combined group ratings. This was to allow for any outlayers in the small samples. When the instrument is applied on a larger scale the use of the average or *arithmetic mean* may be considered.

The implementation of IESET on a *national level* can provide scores on individual criteria, branches and overall. If the results were available teachers would not need to do their own assessments. They would be able to consider the scores and choose a textbook that is appropriate for their context. A teacher with limited subject knowledge will choose one with a high score for core knowledge, while a teacher with good subject knowledge, but limited experience, may prefer a textbook with high scores for activities and assessment.

Thorough textbook assessment can have a positive effect on science education in our country and can indirectly motivate publishers to provide learners with better textbooks. The IESET should be made *available to publishers* to guide textbook authors.

6.5 Conclusions

The development process derived from the AHP has proved valuable in the development of the IESET and can be strongly recommended for use in other contexts. Elements of the AHP contributed significantly to the validity of the results obtained by the IESET.

Implementation of the IESET illustrated that the instrument was easy to use and generated transparent, reliable and valid assessment results. It can be used successfully by teachers or education providers during the selection of textbooks and for textbook authors and publishers as a reference to textbook quality.

Bibliography

Abd-El-Khalick, F. and Lederman, N. G.: 2000, Improving Science teachers' conception of Nature of Science: A critical review of the literature, *International Journal of Science Education* **22**(7), 665–701.

Abruscato, J.: 2004, *Teaching Children Science: Discovery Methods for the Elementary and Middle Grades*, second edn, Allyn & Bacon, Boston, Massachusetts.

Abusharbain, E., n.d., Textbook evaluation [online]. Available from: `http://www.siue.edu/~eabusha/TEXTBOOKEVAL.doc` [last accessed on 2008-07-18].

Ade-Ridder, L.: 1989, Textbook decisions: Making an informed choice, *Family Relations* **38**, 231–240.

Adolfsson, L. and Henriksson, W.: 1999, Different methods – different results: How the Swedish results in Mathematics vary depending on methodological approach, *Educational Research and Evaluation* **5**(2), 127–138.

Alabama Department of Education, 2009, Textbook appraisal form [online]. Available from: `http://ftp.alsde.edu/document/54/Evaluation\%20Form\%20Textbooks\%20Gen%eral.doc` [last accessed on 2009-02-23].

Aladejana, F. and Aderibigbe, O.: 2007, Science laboratory environment and academic performance, *Journal of Science Education and Technology* **16**, 500–506.

Aldridge, J. M., Laugksch, R. C. and Fraser, B. J.: 2006, School-level environment and outcomes-based education in South Africa, *Learning Environments Research* **9**, 123–147.

Aleixandre, M. P. J.: 1994, Teaching evolution and natural selection: A look at textbooks and teachers, *Journal of Research in Science Teaching* **31**(5), 519–535.

Alexander, P. A.: 1998, The nature of disciplinary and domain learning: The knowledge, interest, and strategic dimensions of learning from subject matter text, *in* C. R. Hynd (ed.), *Learning from Text Across Conceptual Domains*, Lawrence Erlbaum Associates, Mahwah, New Jersey, Chapter 12, pp. 263–287.

Alexander, P. A. and Kulikuwich, J. M.: 1994, Learning from Physics text: A synthesis of recent research, *Journal of Research in Science Teaching* **31**(9), 895–911.

Allington, R. L.: 2002, You can't learn much from books you can't read, *Educational Leadership* pp. 16–19.

Amaral, O. M. and Garrison, L.: 2007, Missing the forest for the trees, *Journal of Science Education and Technology* **16**, 155–169.

American Association for the Advancement of Science, 2002, Typical sightings for algebra 1: Explorations and applications [online]. Available from: `http://www.p2061.org/publications/textbook/algebra/summary/Alg1Expl/ts_%Alg1E.pdf` [last accessed on 2009-04-09].

American Association for the Advancement of Science, n.d.a, AAAS Project 2061 – advancing Science literacy for all [online]. Available from: `www.project2061.org` [last accessed on 2008-02-02].

American Association for the Advancement of Science, n.d.b, AAAS Project 2061 Middle Grades Science textbooks evaluation: Summary of instructional analysis ratings in Physical Science [online]. Available from: `http://www.project2061.org/publications/textbook/mgsci/report/chart-phy%sical.pdf` [last accessed on 2010-03-13].

Ametller, J. and Pinto, R.: 2002, Students' reading of innovative images of energy at secondary school level, *International Journal of Science Education* **24**(3), 285–312.

Anagnostopoulos, D.: 2005, Testing, tests, and classroom texts, *Journal of Curriculum Studies* **37**(1), 35–63.

Anderson, L. M.: 1993, Auxiliary materials that accompany textbooks: Can they promote "higher-order" learning?, *in* B. K. Britton, A. Woodward and M. Binkley (eds), *Learning from Textbooks: Theory and Practice*, Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp. 135–160.

Anderson, R. D. and Helms, J. V.: 2001, The ideal of standards and the reality of schools: Needed research, *Journal of Research in Science Teaching* **38**(1), 3–16.

Ansary, H. and Babaii, E.: 2002, Universal characteristics of EFL/ESL textbooks: A step towards systematic textbook evaluation, *The Internet*

TESL Journal **VIII**(2). Available from: http://iteslj.org/Articles/Ansary-Textbooks/ [last accessed on 2010-02-07].

Appelbaum, P. and Clark, S.: 2001, Science! Fun? A critical analysis of design/content/evaluation, *Journal of Curriculum Studies* **33**(5), 583–600.

Apple, M. W. and Christian-Smith, L. K.: 1991, The politics of the textbook, *in* M. W. Apple and L. K. Christian-Smith (eds), *The Politics of the Textbook*, Routledge, New York, pp. 1–21.

Arif, M., Smiley, F. and Kulonda, D.: 2005, Business and education as push-pull processes: An alliance of philosophy and practice, *Education* **125**(4), 602–614.

Armbuster, B. B., Osborn, J. H. and Davidson, A. L.: 1985, Readability formulas may be dangerous to your textbooks, *Educational Leadership* pp. 18–20.

Arriassecq, I. and Greca, I. M.: 2007, Approaches to the teaching of special relativity theory in high school and university textbooks of Argentina, *Science & Education* **16**, 65–86.

Asimov, I.: 1981, *A Choice of Catastrophes*, Arrow Books, London.

Asmal, K.: 1999, Address at the Royal Netherlands Embassy Learning Support Project on 15 November 1999. Available from: http://www.info.gov.za/speeches/1999/991122436p1006.htm.

Asmal, K.: 2002, Speech by Professor Kader Asmal, Minister of Education at the SABA annual general meeting on 20 August 2002. Available from: http://www.info.gov.za/speeches/2002/02120616111001.htm.

Astleitner, H., n.d., Media assessment: MEAS – evaluation and consulting [online]. Available from: http://www.sbg.ac.at/erz/meas/home.htm [last accessed on 2006-03-09].

Ausiello, D. A.: 2007, Science education and communication, *The Journal of Clinical Investigation* **117**(10), 3128–3130. Available from: http://www.jci.org [last accessed on 2008-01-03].

Bailey, M.: n.d., Learning, technology and educational transformation: From philosophy to pedagogy. Available from: http://education.ed.pacificu.edu/bcis/workshop/philosophy.html [last accessed on 2009-06-12].

Bakonis, E.: 1998, Analysis of textbooks: Theory and practice, *Technical Report Newsletter 1*, Textbook Research and Information Centre for the Baltic

Countries (VTIC). Available from: `http://www.pedagogika.It/vtic/newsletter.htm` [last accessed on 2005-08-09].

Ball, D. L. and Feiman-Nemser, S.: 1988, Using textbooks and teachers' guides: A dilemma for beginning teachers and teacher educators, *Curriculum Inquiry* **18**(4), 401–423.

Bancroft, J.: 2002, A methodology for developing Science teaching materials for pupils with learning difficulties, *Support for Learning* **17**(4), 168–175.

Basson, I.: 2002, Physics and Mathematics as interrelated fields of thought development using acceleration as an example, *International Journal of Mathematical Education in Science and Technology* **33**(5), 679–690.

Bazler, J. A. and Simonis, D. A.: 1991, Are high school Chemistry textbooks gender fair?, *Journal of Research in Science Teaching* **28**(4), 353–362.

Beck, I. L. and McKeown, M. G.: 2001, Inviting students into the pursuit of meaning, *Educational Psychology Review* **13**(3), 225–241.

Bell, B.: 2000, Formative assessment and Science education: A model and theorizing, *in* R. Millar, J. Leach and J. Osborne (eds), *Improving Science Education: The Contribution of Research*, Open University Press, Buckingham, pp. 48–61.

Bennett, J. and Kennedy, D.: 2001, Practical work at the upper high school level: The evaluation of a new model of assessment, *International Journal of Science Education* **23**(1), 97–110.

Bensaude-Vincent, B.: 2006, Textbooks on the map of Science studies, *Science & Education* **15**, 667–670.

Bernier, M. J.: 1996, Establishing the psychometric properties of a scale for evaluating quality in printed educational materials, *Patient Education and Counselling* **29**, 283–299.

Best, R. M., Rowe, M., Ozuru, Y. and McNamara, D. S.: 2005, Deep-level comprehension of Science texts: The role of the reader and the text, *Topics in Language Disorders* **25**(1), 65.

Bisseker, C.: 2005, Time, teachers and textbooks, *Financial Mail* (28 October 2005). Interview with Jonathan Jansen.

Black, P. and Hughes, S.: 2003, Using narrative to support young children's learning in Science and Design and Technology, *in* D. Davies and A. Howe (eds), *Teaching Science, Design and Technology in the Early Years*, David Fulton, London, Chapter 4, pp. 38–50.

Blin, M. J. and Tsoukiàs, A.: 2001, Multicriteria methodology contribution to the software quality evaluation, *Software Quality Journal* **9**(2), 113–132.

Blumenfeld, P. C., Kempler, T. M. and Krajcik, J. S.: 2006, Motivation and cognitive engagement in learning environments, *in* R. K. Sawyer (ed.), *The Cambridge Handbook of The Learning Sciences*, Cambridge University Press, New York, pp. 475–488.

Boostrom, R.: 2001, Whither textbooks?, *Journal of Curriculum Studies* **33**(2), 119–145.

Boscolo, P. and Mason, L.: 2003, Topic knowledge, text coherence, and interest: How they interact in learning from instructional texts, *The Journal of Experimental Education* **71**(2), 126–148.

Boulter, C. and Gilbert, J.: 1996, Texts and contexts: Framing modelling in the primary Science classroom, *in* G. Welford, J. Osborn and P. Scott (eds), *Research in Science Education in Europe: Current Issues and Themes*, The Falmer Press, London, pp. 177–188.

Bowen, G. A.: 2008, Naturalistic inquiry and the saturation concept: a research note, *Qualitative Research* **8**(1), 137–152.

Brackenridge, J. B.: 1989, Education in Science, History of Science, and the textbook – necessary vs. sufficient conditions, *Interchange* **20**(2), 71–80.

Brändström, A.: 2005, *Differentiated tasks in Mathematics textbooks: An analysis of the levels of difficulty*, Master's thesis, Luleå University of Technology, Sweden.

Brandt, C.: 2005, Examining the "script" in Science education: Critical literacy in the classroom, *in* A. J. Rodriguez and R. S. Kitchen (eds), *Preparing Mathematics and Science Teachers for Diverse Classrooms*, Lawrence Erlbaum Associates, Mahwah, New Jersey, Chapter 12, pp. 243–261.

Bransford, J. D., Vye, N., Bateman, H., Brophy, S. and Roselli, R.: 2004, Vanderbilt's Amigo3 project: Knowledge of how people learn enters cyberspace, *in* T. Duffy and J. Kirkley (eds), *Learner-centered theory and practice in distance education*, Lawrence Erlbaum Associates, Mahwah, New Jersey, Chapter 10, pp. 209–234.

Brickhouse, N. W., Dagher, Z. R., Shipman, H. L. and Letts, W. J.: 2000, Why things fail: Evidence and warrants for belief in a college astronomy course, *in* R. Millar, J. Leach and J. Osborne (eds), *Improving Science Education: The Contribution of Research*, Open University Press, Buckingham, pp. 11–26.

Britton, B. K., Dusen, L. V., Gülgöz, S. and Glynn, S. M.: 1989, Instructional text rewritten by five expert teams: Revisions and retention improvements, *Journal of Educational Psychology* **81**(2), 226–239.

Britton, B. K., Gülgöz, S. and Glynn, S.: 1993, Impact of good and poor writing on learners: Research and theory, *in* B. K. Britton, A. Woodward and M. Binkley (eds), *Learning from Textbooks: Theory and Practice*, Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp. 1–46.

Broer, N. A., Aarnoutse, C. A. J., Kieviet, F. K. and Leeuwe, J. F. J. V.: 2002, The effect of instructing the structural aspect of text, *Educational Studies* **28**(3), 213–238.

Brooks, V.: 2002, *Assessment in Secondary Schools*, Open University Press, Buckingham.

Brown, J. B.: 1998, Textbook evaluation form, *The Language Teacher Online* **21**(11). Available from: `http://langue.hyper.chubu.ac.jp/jalt/pub/tlt/97/oct/brown.html` [last accessed on 2003-02-15].

Budiansky, S.: 2001, The trouble with textbooks, *Prism* pp. 24–27. Available from: `http://www.project2061.org/newsinfo/research/articles/asee.htm` [last accessed on 2003-02-10].

Burns, B.: 2006, I don't have to count syllables on my fingers anymore: Easier ways to find readability and level books, *Illinois Reading Council Journal* **34**(1), 34–40.

Butler, M. B. and Nesbit, C.: 2008, Using Science notebooks to improve writing skills and conceptual understanding, *Science Activities* **44**(4), 137–140.

Cain, K. and Oakhill, J.: 2006, Assessment matters: Issues in the measurement of reading comprehension, *British Journal of Educational Psychology* **76**, 679–708.

California Department of Education, n.d., Criteria for evaluating K8 Science instructional materials [online]. Available from: `http://www.cde.ca.gov/cfir/science/scicriteria.html` [last accessed on 2003-03-03].

Carey, S.: 2000, Science education as conceptual change, *Journal of Applied Developmental Psychology* **21**(1), 13–19.

Carney, R. N. and Levin, J. R.: 2002, Pictorial illustrations still improve students' learning from text, *Educational Psychology Review* **14**(1), 5–26.

Carr, D.: 2003, *Making Sense of Education. An introduction to the Philosophy and Theory of Education and Teaching*, RoutledgeFalmer, London.

Center for Curriculum Materials in Science (CCMS), n.d., CCMS mission and principles [online]. Available from: `http://www.sciencematerialscenter.org/about_mission.htm` [last accessed on 2000-12-09].

Chall, J. S. and Conard, S. S.: 1991, *Should Textbooks Challenge Students?: The Case for Easier or Harder Textbooks*, Teachers College Press, New York.

Chambliss, M. J.: 1994, Evaluating the quality of textbooks for diverse learners, *Remedial & Special Education* **15**(6), 348–363.

Chambliss, M. J.: 2001, Analyzing Science textbook materials to determine how "persuasive" they are, *Theory into Practice* **40**(4), 255–264.

Chambliss, M. J. and Calfee, R. C.: 1998, *Textbooks for Learning: Nurturing Children's Minds*, Blackwell, Massachusetts.

Chandler, D. L., 1999, Textbooks flunk out [online]. Available from: `http://www.harvard.edu/cfa/ep/pressrel/textbooks.html` [last accessed on 2003-03-27].

Chatterji, M.: 2003, *Designing and Using Tools for Educational Assessment*, Allyn & Bacon, Boston, Massachusetts.

Cheung, D.: 2007, Facilitating Chemistry teachers to implement inquiry-based laboratory work, *International Journal of Science and Mathematics Education* **6**, 107–130.

Chiappetta, Sethna, G. H. and Fillman, D. A.: 1993, Do Middle School Life Science textbooks provide a balance of scientific literacy themes?, *Journal of Research in Science Teaching* **30**(7), 787–797.

Cicourel, A. V.: 1985, Text and discource, *Annual Review of Anthropology* **14**, 159–185.

Cisar, S. H., n.d., Standards-based textbook evaluation guide [online]. Available from: `http://title3.sde.state.ok.us/languages/textbookevaluation-3.htm` [last accessed on 2001-05-29].

Clark, R. E.: 2001, Media are “mere vehicles”: The opening argument, *in* R. E. Clark (ed.), *Learning with Media*, Information Age Publishing, Conneticut, Chapter 1, pp. 1–12.

Clark, R. E. and Salomon, G.: 2001, Why should we expect media to teach anyone anything?, *in* R. E. Clark (ed.), *Learning with Media*, Information Age Publishing, Conneticut, Chapter 3, pp. 37–70.

Clericuzio, A.: 2006, Teaching Chemistry and chemical textbooks in France. from Beguin to Lemery, *Science & Education* **15**, 335–355.

Clough, M. P.: 2006, Learners' responses to the demands of conceptual change: Considerations for effective Nature of Science instruction, *Science Education* **15**, 463–494.

Cocking, R. R., Mestre, J. P. and Brown, A. L.: 2000, New developments in the science of learning: Using research help students learn Science and Mathematics, *Journal of Applied Developmental Psychology* **21**, 1–11.

Coetzer, I. A.: 2001, A survey and appraisal of outcomes-based education (OBE) in South Africa with reference to progressive education in America, *Educare* **30**(1), 73–93.

College of Education: New Mexico State University, n.d., Archived textbook evaluations summary chart [online]. Available from: `http://pt3.nmsu.edu/texts/chart.html` [last accessed on 2005-09-09].

Commisso, A.: 2004, Iraqi textbook quality improvement program for primary and secondary education Mathematics and Science textbooks: Review, training, printing and distribution process: A case study, *Technical Report April 2004*, UNESCO.

Conzales, P., Guzman, J. C., Partelow, L., Pahlke, E., Jocelyn, L., Kastberg, D. and Williams, T.: 2004, *Highlights from the Trends in International Mathematics and Science Study (TIMMS) 2003*, U.S. Department of Education, National Centre for Educational Statistics, Washington, D.C.

Cook, A. and Tulip, D.: 1992, The importance of selected textbook features to Science teachers, *Research in Science Education* **22**, 91–100.

Cordova, D. I. and Lepper, M. R.: 1996, Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice, *Journal of Educational Psychology* **88**(4), 715–730.

Costa, J., Caldeira, H., Gallástegui, J. R. and Otero, J.: 2000, An analysis of question asking on scientific texts explaining natural phenomena, *Journal of Research in Science Teaching* **37**(6), 602–614.

Costu, B., Ayas, A., Niaz, M., Unal, S. and Calik, M.: 2007, Facilitating conceptual change in students' understanding of boiling concept, *Journal of Science Education and Technology* **16**, 524–536.

Council of State Science Supervisors, n.d., Rubric for evaluating essential features of classroom inquiry in instructional materials [online]. Available from: `http://www.inquiryscience.com/documents/materials.htm` [last accessed on 2002-06-12].

Cowie, B. and Bell, B.: 1999, A model of formative assessment in Science education, *Assessment in Education* **6**(1), 101–116.

Crawford, B. A.: 2007, Learning to teach Science as inquiry in the rough and tumble of practice, *Journal of Research in Science Teaching* **44**(4), 613–642.

Crawford, M. B.: 2005, Science education and liberal education, *The New Atlantis* pp. 49–60. Available from: `http://www.thenewatlantis.com/publications/science-education-and-libera%l-education` [last accessed on 2009-11-30].

Cronjé, J.: 2006, Paradigms regained: Toward integrating Objectivism and Constructivism in instructional design and the learning sciences, *Educational Technology Research and Development* **54**(4), 387–416.

Crystal Springs Books, n.d., Textbook evaluation form [online]. Available from: `http://www.sde.com/Conferences/Differentiated-Instruction/downloads/Dif%ferentiating_Textbooks/TextbookEval.pdf` [last accessed on 2006-01-30].

Dake, L. S.: 2007, Student selection of the textbook for an introductory Physics course, *The Physics Teacher* **45**, 416–419.

Dale, E. and Chall, J. S.: 1948, A formula for predicting readability, *Educational Research Bulletin* **27**, 37–53.

Dall'Alba, G., Walsh, E., Bowden, J., Martin, E., Masters, G., Ramsden, P. and Stephanou, A.: 1993, Textbook treatments and students' understanding of acceleration, *Journal of Research in Science Teaching* **30**(7), 621–635.

Daniel, J. and Menon, M.: 2005, ODL and ICTs for teacher development in Sub-Saharan Africa: The experience of the Commonwealth of Learning, Paper presented at the *Botswana College of Distance and Open Learning (BOCODOL)* workshop in Gaborone, Botswana on 1 September 2005.

Daniels, H. and Zemelman, S.: 2003, Out with textbooks, in with learning, *Educational Leadership* **61**(4), 36–40.

Darabi, A. A., Nelson, D. W. and Paas, F.: 2007, Learner involvement in instruction on a complex cognitive task: Application of a composite measure of performance and mental effort, *Journal of Research on Technology in Education* **40**(1), 39–48.

Davidson, G. G. and Labib, A. W.: 2003, Learning from failures: design improvements using a multiple criteria decision-making process, *Proceedings of the Institute for Mechanical Engineers* **217** Part G, pp. 207–216.

Davies, D.: 2003, Developing children's scientific knowledge, skills and attitudes, *in* D. Davies and A. Howe (eds), *Teaching Science, Design and Technology in the Early Years*, David Fulton, London, Chapter 9, pp. 120–135.

Davies, D. and Ward, S.: 2003, Young children as scientists, designers and technologists, *in* D. Davies and A. Howe (eds), *Teaching Science, Design and Technology in the Early Years*, David Fulton, London, Chapter 2, pp. 10–24.

Davis, E. A. and Linn, M. C.: 2000, Scaffolding students' knowledge integration: Prompts for reflection in KIE, *International Journal of Science Education* **22**(8), 819–837.

Davis, J. D.: 2009, Understanding the influence of two Mathematics textbooks on prospective secondary teachers' knowledge, *Journal of Mathematics Teacher Education* **12**, 365–389.

Davis, K.: 2003a, "Change is hard": What Science teachers are telling us about reform and teacher learning of innovative practices, *Science Education* **87**(3), 3–30.

Davis, S.: 2003b, Selecting challenging instructional materials for Middle-Grades Science: A digest for state, district, and school leaders, *Technical report*, Center for Science Education.

De Beer, J.: 2002, The NEPAD road to 2061: Reflections on Science education in Southern Africa. Paper read at the American Association for the Advancement of Science (AAAS) meeting, Washington, 4 October 2002.

De Berg, K. C. and Greive, C. E.: 1999, Understanding the siphon: An example of the development of pedagogical content knowledge using textbooks and the writings of early scientists, *Australian Science Teachers Journal* **45**(4), 19–26.

DeBolt, G. P.: 1992, Consumers of textbooks: Concerns from the classroom, *in* J. G. Herlihy (ed.), *The Textbook Controversy*, Ablex Publishing Corporation, Norwood, New Jersey, Chapter 14, pp. 137–145.

Dekker, A. J.: 1983, Die handboekskrywer se probleme, *in* R. A. Krüger (ed.), *Handboeke in die Onderwys*, Butterworth, Durban, pp. 21–24.

Department of Education of South Africa: 1997, Outcomes-based education in South Africa. Background information for educators.

Department of Education of South Africa: 1999, Report of the review committee on Curriculum 2005, Pretoria.

Department of Education of South Africa: 2001a, *Draft Revised National Curriculum Statement for Grades R–9, (Schools) for Natural Sciences*. Gazette no 22559 of 5 August 2001.

Department of Education of South Africa: 2001b, *National Strategy for Mathematics, Science and Technology Education in General and Further Education and Training*, Department of Education, Pretoria.

Department of Education of South Africa: 2002a, *Draft National Curriculum Statement for Grades 10–12, (Schools) for Physical Sciences*, Department of Education, Pretoria.

Department of Education of South Africa: 2002b, *Revised National Curriculum Statement for Grades R–9, (Schools) for Natural Sciences*, Department of Education, Pretoria.

Department of Education of South Africa: 2003a, *National Curriculum Statement Grades 10–12 (General) Physical Sciences*, Department of Education, Pretoria.

Department of Education of South Africa: 2003b, *Reviewing of the Financing, Resourcing and Costs of Education in Public Schools*, Department of Education, Pretoria.

Department of Education of South Africa: 2005, National Curriculum Statement grades R–9: Parents' guide, Supplement to the Sunday Times of 16 January 2005, Pretoria.

Department of Education of South Africa: 2006, *National Curriculum Statement Grades 10–12 (General) Physical Sciences Content*, Department of Education of South Africa.

Department of Education of South Africa: 2009, *Curriculum News. Improving the quality of learning and teaching. Planning for 2010 and beyond*, Department of Education of South Africa, Pretoria.

Department of Education of South Africa, n.d., Guiding criteria for selecting textbooks [online]. Available from: `http://www.thutong.org.za/resourcedetails.asp?col=6&id=18429&instanceid%=18552&pop=1` [last accessed on 2007-12-25].

Department of Education of the Western Cape, n.d., Textbooks are important to OBE and the RNCS [online]. Available from: `http://wwwcurriculum.wcape.school.za/site/104/news/view/215` [last accessed on 2003-06-27].

Dietz, T. and Kalof, L.: 2009, *Introduction to Social Statistics*, Wiley-Blackwell, Chichester, West Sussex.

Dillemans, R., Lowyck, J., Van der Perre, G., Claeys, C. and Elen, J.: 1998, *New Technologies for Learning: Contribution of ITC to Innovation in Education*, Leuven University Press, Leuven (Belgium).

Dimopoulos, K., Koulaidis, V. and Sklaveniti, S.: 2005, Towards a framework of socio-linguistic analysis of Science textbooks: The Greek case, *Research in Science Education* **35**, 173–195.

DiSessa, A. A.: 2006, A history of conceptual change research, *in* R. K. Sawyer (ed.), *The Cambridge Handbook of The Learning Sciences*, Cambridge University Press, New York, pp. 265–281.

Driver, R.: 1994, The fallacy of induction in Science teaching, *in* R. Levinson (ed.), *Teaching Science*, Routledge, London and New York, pp. 41–48.

Driver, R., Asoko, H., Leach, J., Mortimer, E. and Scott, P.: 1994, Constructing scientific knowledge in the classroom, *Educational Researcher* **23**(7), 5–12.

Dutch, S. I.: 2005, Why textbooks are the way they are, *Academic Questions* **18**(4), 34–48.

Eltinge, E. M. and Roberts, C. W.: 1993, Linguistic content analysis: A method to measure Science as inquiry in textbooks, *Journal of Research in Science Teaching* **30**(1), 65–83.

Erduran, S., Bravo, A. A. and Naaman, R. M.: 2007, Developing epistemologically empowered teachers: examining the role of Philosophy of Chemistry in teacher education, *Science & Education* **16**, 975–989.

ESchool News Online, n.d., New forum helps Science teachers evaluate textbooks [online]. Available from: `http://sciencetextcentral.org` [last accessed on 2005-04-01].

Ewing, B. F.: 2006, "Go to the page and work it from there." Young people's experiences of learning Mathematics from a text, *Australian Senior Mathematics Journal* **20**(1), 8–14.

Fetsko, W. J.: 1992, Approaching textbook selection systematically, *in* J. G. Herlihy (ed.), *The Textbook Controversy*, Ablex Publishing Corporation, Norwood, New Jersey, Chapter 13, pp. 129–135.

Fishman, B. J. and Davis, E. A.: 2006, Teacher Learning Research and the Learning Sciences, *in* R. K. Sawyer (ed.), *The Cambridge Handbook of The Learning Sciences*, Cambridge University Press, New York, pp. 535–550.

Flores, F., Lopez, A., Gallegos, L. and Barojas, J.: 2000, Transforming Science and learning concepts of Physics teachers, *International Journal of Science Education* **22**(2), 197–208.

Flowerday, T. and Schraw, G.: 2000, Teacher beliefs about instructional choice: A phenomenological study, *Journal of Educational Psychology* **92**(4), 634–645.

Fordham, N. W., Wellman, D. and Sandmann, A.: 2002, Taming the text: Engaging and supporting students in social studies reading, *Sosial Studies* **93**(4), 149–158.

Foreign Policy, n.d., Libyan children to get wired [online]. Available from: `http://blog.foreignpolicy.com/node/1933` [last accessed on 2006-11-10].

Forman, E. H. and Selly, M. A.: 2002, *Decision by Objectives*, World Scientific Publishing Company, Singapore.

Franssen, H. A. M.: 1989a, Het ontwikkelen van schoolboeken nader bekeken, *in* A. Reints, R. A. De Jong and N. A. J. Lagerweij (eds), *Om de Kwaliteit van het Leermiddel*, Zwijsen, Tilburg, pp. 31–36.

Franssen, H. A. M.: 1989b, Procedures voor het kiezen van schoolboeken, *in* A. Reints, R. A. De Jong and N. A. J. Lagerweij (eds), *Om de Kwaliteit van het Leermiddel*, Zwijsen, Tilburg, pp. 98–109.

Freeman, E. B. and Person, D. G.: 1998, *Connecting Informational Children's Books with Content Area Learning*, Allyn & Bacon, Boston, Massachusetts, Chapter 1, pp. 1–24.

Friar, L., Matson, J. O. and Matson, J. E.: 1998, An undergraduate curriculum evaluation with the Analytic Hierarchy Process, *FIE Conference*, pp. 992–997.

Fry, E.: 2002, Readability versus leveling, *The Reading Teacher* **56**(3), 286–291.

Furtak, E. M.: 2006, The problem with answers: An exploration of guided scientific inquiry teaching, *Science Education* **90**, 453–467.

Galili, I. and Tseitlin, M.: 2003, Newton's first law: Text, translations, interpretations and Physics education, *Science & Education* **12**, 45–73.

Garcia-Barros, S., Martinez-Losada, C., Vega, P. and Mondelo, M.: 2001, The ideas of Spanish primary teachers on how to develop an understanding of processes in Science and their support in textbooks, *in* H. Beherendt, H. Dahncke, R. Duit, W. Gräber, M. Komorek and A. Kross (eds), *Research in Science Education – Past, Present, and Future*, Kluwer Academic, Dordrecht, pp. 149–154.

Garinger, D., 2003, Textbook evaluation [online]. Available from: `http://www.teflweb-j.org/v1n3/v1n3.html` [last accessed on 2003-02-24].

Gauteng Department of Education: 2006, Circular 26/2006: Resource allocation for public ordinary schools for 2006/2007. Available from: `http://www.education.gpg.gov.za/Legislation/circulars/CircularsArchives%.htm` [last accessed on 2009-12-01].

Gerber, B. L., Cavallo, A. M. L. and Marek, E. A.: 2001, Relationships among informal learning environments, teaching procedures and scientific reasoning ability, *International Journal of Science Education* **23**(5), 535–549.

Giles, T. D. and Still, B.: 2005, A syntactic approach to readability, *Journal of Technical Writing & Communication* **35**(1), 47–70.

Glynn, S. M. and Takahashi, T.: 1998, Learning from analogy-enhanced Science text, *Journal of Research in Science Teaching* **35**(10), 1129–1149.

Goldsmith, L. T., Mark, J. and Kantrov, I.: 2000, *Choosing a Standards-Based Mathematics Curriculum*, Heinemann, Portsmouth, New Hampshire.

Good, R.: 1993, Science textbook analysis (editorial), *Journal of Research in Science Teaching* **30**(7), 619.

Goodwin-Davey, A. and Davey, C.: 2000, Development of learning resources for outcomes-based distance education, *Progressio* **22**(1).

Gräber, W., Nentwig, P., Becker, H.-J., Sumflecth, E., Pitton, A., Wollweber, K. and Jorde, D.: 2001, Scientific literacy: From theory to practice, *in* H. Behrendt, H. Dahncke, R. Duit, W. Gräber, M. Komorek, A. Kross and P. Reiska (eds), *Research in Science Education – Past, Present, and Future*, Kluwer Academic, Dordrecht, pp. 61–70.

Gross, S. J., Shaw, K. M. and Shapiro, J. P.: 2003, Deconstructing accountability through the lens of democratic philosophies: Toward a new analytic framework, *Journal of Research for Educational Leaders (JREL)* **1**(3), 5–27. Available from: `http://www.education.uiowa.edu/jrel/spring03/spring03_number3.htm` [last accessed on 2010-01-30].

Guisasola, J., Almudi, J. M. and Furio, C.: 2005, The Nature of Science and its implications for Physics textbooks: The case of Classical Magnetic Field Theory, *Science & Education* **14**, 321–338.

Gunel, M., Hand, B. and Prain, V.: 2007, Writing for learning in Science: A secondary analysis of six studies, *International Journal of Science and Mathematics Education* **5**, 615–637.

Guthrie, J. T. and Wigfield, A.: 2005, Roles of motivation and engagement in reading comprehension assessment, *in* S. G. Paris and S. A. Stahl (eds), *Children's Reading Comprehension and Assessment*, Lawrence Erlbaum Associates, Mahwah, New Jersey, pp. 187–213.

Guzzetti, B. J.: 2000, Learning counter-intuitive Science concepts: What have we learned from over a decade of research?, *Reading and Writing Quaterly* **16**, 89–96.

Halloun, I. A.: 2007, Mediated modeling in Science education, *Science & Education* **16**, 653–697.

Hand, B. and Keys, C. N.: 1999, Inquiry investigation. A new approach to laboratory reports, *The Science Teacher* **66**, 27–29.

Hand, B., Prain, V., Lawrence, C. and Yore, L. D.: 1999, A writing in Science framework designed to enhance Science literacy, *International Journal of Science Education* **21**(10), 1021–1035.

Hargraves, E.: 2005, Assessment for learning? Thinking outside the (black) box, *Cambridge Journal of Education* **35**(2), 213–224.

Harp, S. F. and Mayer, R. E.: 1997, The role of interest in learning from scientific text and illustrations: On the distinction between emotional interest and cognitive interest, *Journal of Educational Psychology* **89**(1), 92–102.

Harrison, A. G.: 2001, How do teachers and textbook writers model scientific ideas for students?, *Research in Science Education* **31**, 401–435.

Hattingh, A.: 2009, Meester-wiskundeonderwysers as mentors in ondervoorsiene en benadeelde skole, *Die Suid-Afrikaanse Tydskrif vir Natuurwetenskap en Tegnologie* **28**(4), 340–354.

Henson, K. T.: 2004, *Constructive Methods for Teaching in Diverse Middle-level Classrooms*, Allyn & Bacon, Boston, Massachusetts.

Herlihy, J. G.: 1992a, The nature of the textbook controversy, *in* J. G. Herlihy (ed.), *The Textbook Controversy*, Ablex Publishing Corporation, Norwood, New Jersey, Chapter 1, pp. 3–13.

Herlihy, J. G. (ed.): 1992b, *The Textbook Controversy*, Ablex Publishing Corporation, Norwood, New Jersey.

Hewson, P. and Lemberger, J.: 2000, Status as the hallmark of conceptual learning, *in* R. Millar, J. Leach and J. Osborne (eds), *Improving Science Education: The Contribution of Research*, Open University Press, Buckingham, pp. 110–125.

Hewson, P. W.: 2004, Resources for Science learning: Tools, tasks, and environment, *International Journal of Science and Mathematics Education* **2**, 201–225.

Hodson, D. and Prophet, R. B.: 1994, Why Science curriculum changes: Evolution or social control?, *in* R. Levinson (ed.), *Teaching Science*, Routledge, London and New York, pp. 22–38.

Holliday, W. G.: 2003, Influential research in Science teaching: 1963–present, *Journal of Research in Science Teaching* **40** (Supplement), v–x.

Holliday, W. G., Yore, L. D. and Alvermann, D. E.: 1994, The reading-science learning-writing connection: Breakthroughs, barriers, and promises, *Journal of Research in Science Teaching* **31**(9), 877–893.

Hong Kong Department of Education, 2002, Guiding principles for quality textbooks [online]. Available from: `http://cd.ed.gov.hk/cr_2001/eng/textbook/background1.htm` [last accessed on 2002-09-04].

House, J. D.: 2000, Relationships between instructional activities and Science achievement of adolescent students in Hong Kong: Findings from the Third

International Mathematics and Science Study (TIMSS), *International Journal of Instructional Media* **27**(3), 275–288.

Hsu, P.-L. and Yang, W.-G.: 2007, Print and image integration of Science texts and reading comprehension: A systematic function linguistics perspective, *International Journal of Science and Mathematics Education* **5**, 639–659.

Huang, H.-J.: 2006, Listening to the language of constructing Science knowledge, *International Journal of Science and Mathematical Education* **4**, 391–415.

Huber, R. A. and Moore, C. J.: 2001, A model for extending hands-on Science to be inquiry based, *School Science & Mathematics* **101**(1), 32–42.

Huberman, M. and Middlebrooks, S.: 2000, The dilution of inquiry: A qualitative study, *Qualitative studies in Education* **13**(3), 281–304.

Hubisz, J. L.: 2003, Choosing Middle School Science textbooks: Is North Carolina failing its students?, *Technical report*, North Carolina Education Alliance. Available from: www.nceducationalliance.org [last accessed on 2007-12-22].

Hudson, T.: 1994, Developing pupils' skills, *in* R. Levinson (ed.), *Teaching Science*, Routledge, London and New York, pp. 94–109.

Huitt, W., Hummel, J. and Kaeck, D.: 2000, Assessment, measurement, evaluation, and research, *Educational Psychology Interactive* . Available from: http://chiron.valdosta.edu/whuitt/col/intro/sciknow.html [last accessed on 2007-03-07].

Hynd, C., Holschuh, J. and Nist, S.: 2000, Learning complex scientific information: Motivation theory and its relation to student perceptions, *Reading and Writing Quarterly* **16**, 23–57.

IARTEM, n.d., The international association for research on textbooks and educational media [online]. Available from: http://www.iartem.no/ [last accessed on 2006-10-24].

Iding, M. K.: 2000, Can strategies facilitate learning from illustrated Science texts?, *International Journal of Instructional Media* **27**(3), 289–301.

Iding, M., Klemm, E. B., Crosby, M. E. and Speitel, T.: 2002, Interactive texts, figures and tables for learning Science: Constructivism in text design, *International Journal of Instructional Media* **29**(4), 441–452.

Imuse: digital design company, n.d., Textbook evaluation tool [online]. Available from: http://www.imuse.com/portfolio/TVal_case.html.

Independent Projects Trust, n.d., The IPT and the new changes in education [online]. Available from: http://www.ipt.co.za [last accessed on 2005-07-02].

Issitt, J.: 2004, Reflections on the study of textbooks, *History of Education* **33**(6), 683–696.

Iszak, A. and Sherin, M. G.: 2003, Exploring the use of new representations as a resource for teacher learning, *School Science and Mathematics* **103**(1), 18–27.

Iz, P. H., 1991, Group decision support and multiple criteria optimization [online]. Available from: http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=184201 [last accessed on 2009-11-24].

Izsak, A. and Sherin, M. G.: 2003, Exploring the use of new representations as a resource for teacher learning, *School Science and Mathematics* **103**(1), 18–27.

Jacobs, I.: 2001, The 2000/2001 Annual Report on Education, Gauteng Department of Education. Address by the Gauteng MEC for Education on the presentation of the 2000/2001 Annual Report.

Jansen, J.: 2003, Can research inform policy in developing countries? A South African experience, *International Journal of Educational Development* **23**(1), 85–95.

Jansen, J.: 2007, Learning and leading in a globalized world: The lessons from South Africa, *in* T. Townsend and R. Bates (eds), *Handbook of Teacher Education*, Springer, Chapter 2, pp. 25–40.

Jenkins, E.: 2004, Science, *in* J. White (ed.), *Rethinking the School Curriculum*, RoutledgeFalmer, London, Chapter 13, pp. 165–178.

Jepson, D.: 2002, Fact or fiction: The SBOE's role in textbook adoption, *Technical Report 77-17*, Texas House of Representatives: House Research Organization.

Johnsen, E. B.: 1993, *Textbooks in the Kaleidoscope*, Scandinavian University Press, Oslo.

Johnson, K., n.d., Readability [online]. Available from: http://www.timetabler.com/reading.html [last accessed on 2010-01-21].

Jones, R.: 2000, Textbook troubles, *American School Board Journal* **2000**(12). Available from: http://www.asbj.com/2000/12/1200coverstory.html [last accessed on 2002-01-27].

Jonsson, G., Gustafsson, P. and Enghag, M.: 2007, Context rich problems as an educational tool in Physics teaching – A case study, *Journal of Baltic Science Education* **6**(2), 26–34.

Kan, A.: 2007, An alternative method in the new educational program from the point of performance-based assessment: Rubric scoring scales, *Educational Sciences: Theory & Practice* 7(1), 144–152.

Kearsey, J. and Turner, S.: 1999, Evaluating textbooks: The role of genre analysis, *Research in Science and Technological Education* **17**(1), 35–43.

Keller, J. and Burkman, E.: 1993, Motivation principles, *in* M. Fleming and W. H. Levie (eds), *Instructional Message Design*, 2 edn, Educational Technology Publications, Englewood Cliffs, New Jersey, Chapter 1, pp. 3–53.

Kelly, A. E. and Lesh, R. A.: 2000, Trends and shifts in research methods, *in* A. E. Kelly and R. A. Lesh (eds), *Handbook of Research Design in Mathematics And Science Education*, Lawrence Erlbaum Associates, Mahwah, New Jersey, Chapter 2, pp. 35–44.

Kendeou, P. and Van den Broek, P.: 2005, The effects of misconceptions on comprehension of scientific text, *Journal of Educational Psychology* **97**(2), 235–245.

Kesidou, S. and Roseman, J. E.: 2002, How well do Middle School Science programmes measure up? Findings from Project 2061's curriculum review, *Journal of Research in Science Teaching* **39**(6), 522–549.

Kesidou, S. and Roseman, J. E.: 2003, Project 2061 analysis of Middle-School Science textbooks: A response to Holliday, *Journal of Research in Science Teaching* **40**(5), 535–453.

Khutorskoi, A. V.: 2006, The place of the textbook in the didactic system, *Russian Education and Society* **48**(3), 78–93.

Killen, R.: 2002, Aligning outcomes, teaching strategies and assessment, Paper presented at the *Annual Conference of the South African Society of Educators*, Pretoria, 26–29 September 2002.

Killen, R.: 2003, An introduction to Outcomes-Based Education, Unpublished manuscript, Faculty of Education and Arts, University of New Castle, Australia.

Kindi, V.: 2005, Should Science teaching involve the History of Science? An assessment of Kuhn's view, *Science & Education* **14**, 721–731.

Kintsch, W. and Kintsch, E.: 2005, Comprehension, *in* S. G. Paris and S. A. Stahl (eds), *Children's Reading Comprehension and Assessment*, Lawrence Erlbaum Associates, Mahwah, New Jersey, pp. 71–92.

Klassen, S.: 2006, A theoretical framework for contextual Science teaching, *Interchange* **37**(1–2), 31–62.

Knain, E.: 2001, Ideologies in school Science textbooks, *International Journal of Science Education* **23**(3), 319–329.

Kneller, G. F.: 1964, *Introduction to the Philosophy of Education*, John Wiley & Sons, New York.

Koch, A.: 2001, Training in metacognition and comprehension of Physics texts, *Science Education* **85**(6), 758–768.

Koliopoulos, D. and Constantinou, C.: 2005, The pendulum as presented in school Science textbooks of Greece and Cyprus, *Science & Education* **14**, 59–73.

Kozma, R. B.: 2001, Learning with media, *in* R. E. Clark (ed.), *Learning with Media*, Information Age Publishing, Conneticut, Chapter 8, pp. 137–178.

Krüger, R. A.: 1983, Die kurrikulumdeskundige se standpunt oor die handboek, *in* R. A. Krüger (ed.), *Handboeke in die Onderwys*, Butterworth, Durban, pp. 14–20.

Kulm, G., Roseman, J.-E. and Treistman, M.: 1999, A benchmarks-based approach to textbook evaluation, *Science Books and Films* **35**(4), 147–153. Available from: `http://www.project2061.org/newsinfo/research/textbook/articles/approach%.htm` [last accessed on 2003-02-10].

Ladyman, J.: 2002, *Understanding Philosophy of Science*, Routledge, London.

Land, S.: 2002, The state of book development in South Africa, *Journal of Education* **28**, 93–124.

Lang, M.: 2001, Teacher professionalism and change: Developing a professional self through reflective assessment, *in* H. Behrendt, H. Dahncke, R. Duit, W. Gräber, M. Komorek, A. Kross and P. Reiska (eds), *Research in Science Education – Past, Present, and Future*, Kluwer Academic, Dordrecht, pp. 131–136.

Lederman, L.: 2008, Science education and the future of humankind, *Science News* **173**(16), 1.

Lee, D., McCool, J. and Napieralski, L.: 2000, Assessing adult learning preferences using the Analytic Hierarchy Process, *International Journal of Lifelong Education* **19**(6), 548–560.

Leitte, L.: 2002, History of Science in Science Education: Development and validation of a checklist for analysing the historical content of Science textbooks, *Science & Education* **11**, 333–359.

Lemmer, M., Edwards, J.-A. and Rapule, S.: 2008, Educators' selection and evaluation of Natural Science textbooks, *South African Journal of Education* **28**, 175–187.

Lesh, R. and Lovitts, B.: 2000, Research agendas: Identifying priority problems and developing useful theoretical perspectives, *in* A. E. Kelly and R. A. Lesh (eds), *Handbook of Research Design in Mathematics and Science Education*, Lawrence Erlbaum Associates, Mahwah, New Jersey, pp. 45–71.

Lisichkin, G. V.: 2007, The teaching of the natural science disciplines in the schools, *Russian Education and Society* **49**, 23–32.

Litz, D. R. A.: 2001, Textbook evaluation and ELT management: A South Korean case study, *Asian EFL Journal* pp. 1–52.

Lubben, F., Campbell, B., Kasanda, C., Kapenda, H., Gaoseb, N. and Kanjeo-Marenga, U.: 2003, Teachers' use of textbooks: Practice in Namibian science classrooms, *Educational Studies* **29**(2/3), 109–125.

Lubisi, C., Wedekind, V. and Parker, B.: 1997, *Understanding Outcomes-Based Education: Knowledge, Curriculum and Assessment in South Africa*, South Africa Institute for Distance Education, Bloemfontein.

Lucas, J. M.: 2008, *Fisiese Wetenskappe Graad 12*, JNM, Pretoria.

Luft, J. A. and Patterson, N. C.: 2002, Bridging the gap: Supporting beginning Science teachers, *Journal of Science Teacher Education* **13**(4), 267–282.

Lumby, J. and Foskett, N.: 2005, *Education: Policy, Leadership and Learning*, Sage Publications, London.

Lynch, S., Taymans, J., Watson, W. A., Ochsendorf, R. J., Pyke, C. and Szesze, M. J.: 2007, Effectiveness of highly rated Science curriculum unit for students with disabilities in general education classrooms, *Exceptional Children* **73**(2), 202–223.

Maarschalk, J. and McFarlane, L. R.: 1987, *Vakdidaktiek: Natuur- en Skeikunde*, De Jager-HAUM, Pretoria.

MacDonald, R.: 2006, The use of evaluation to improve practice in learning and teaching, *Innovations in Education and Teaching International* **43**(1), 3–13.

Mahlaba, F. S.: 2006, *An evaluation of the Gauteng Department of Education instrument for evaluating learning and teaching support materials (LTSM)*, Master's thesis, University of Johannesburg.

Malcolm, C. and Alant, B.: 2004, Finding direction when the ground is moving: Science Education Research in South Africa, *Studies in Science Education* **40**, 49–104.

Mann, J. S., Lu, Y. C. and Grzybowski, S. C. W.: 1999, The race of the millennium: CD-rom versus textbook, *Canadian Medical Association Journal* **161**(12), 1527–1529.

Maree, K., Aldous, C., Hattingh, A., Swanepoel, A. and van der Linde, M.: 2006, Predictors of learner performance in Mathematics and Science according to a large-scale study in Mpumalanga, *South African Journal of Education* **26**(2), 229–252.

Marshall, J. D.: 1991, State-level textbook selection reform: Toward the recognition of fundamental control, *in* P. G. Altbach, G. Kelly, H. G. Petrie and L. Weis (eds), *Textbooks in American Society*, State University of New York Press, Albany, New York, pp. 117–133.

Matthews, M. R.: 2002, Constructivism and Science Education: A further appraisal, *Journal of Science Education and Technology* **11**(2), 121–134.

Mayer, R. E.: 1999, Designing instruction for Constructivist learning, *in* C. M. Reigeluth (ed.), *Instructional-Design Theories and Models*, Vol. II, Lawrence Erlbaum Associates, Mahwah, New Jersey, pp. 141–159.

Mayer, R. E.: 2003, The promise of multimedia learning: Using the same instructional design methods across different media, *Learning and Instruction* **13**, 125–139.

McCallum, K.: 2004, The culture of reading. What does it cost?, *Impressions* pp. 1,4 and 7–9. Available from: `http://www.pifsa.org/newsletters/Imp22_200411c.pdf` [last accessed on 2010-04-21].

McEneaney, E. H.: 2003, The worldwide cachet of scientific literacy, *Comparative Education Review* **47**(2), 217–237.

McKenney, S.: 2001, *Computer-Based Support for Science Education Materials Developers in Africa: Exploring Potentials*, PhD thesis, Universiteit Twente.

McKinney, C.: 2005, Textbooks for diverse learners: A critical analysis of learning materials used in South African schools, *Technical report*, HSRC, Cape Town.

McNally, J. G.: 2000, Teaching investigative Science: Preliminary theorizing from the shared reflections of teachers, *International Journal of Science Education* **22**(2), 159–176.

McNamara, D. S., Kintsch, E., Songer, N. B. and Kintsch, W.: 1996, Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text, *Cognition and Instruction* **14**(1), 1–43.

Mellado, V., Bermejo, M. L., Blanco, L. J. and Ruiz, C.: 2007, The classroom practice of a prospective secondary Biology teacher and his conception of the Nature of Science and of teaching and learning of Science, *International Journal of Science and Mathematics Education* **6**, 37–62.

Mesa, V.: 2004, Characterizing practices associated with functions in Middle School textbooks: An empirical approach, *Educational Studies in Mathematics* **56**, 255–286.

Mhlongo, A.: 2006, Schools still without books, *Daily News* p. 3. 22 March 2006.

Miekley, J., 2005, ESL textbook evaluation checklist [online]. Available from: `http://www.readingmatrix.com/reading_projects/miekley/project.pdf` [last accessed on 2010-03-01].

Mikk, J.: 2000, *Textbook Research and Writing*, Peter Lang, Frankfurt am Main.

Mikkilä-Erdmann, M.: 2002, Science learning through text: The effect of text design and text comprehension skills on conceptual change, *in* M. Limón and L. Mason (eds), *Reconsidering Conceptual Change. Issues in Theory and Practice*, Kluwer Academic, Dordrecht, pp. 337–356.

Millar, R.: 1994, What is 'scientific method' and can it be taught?, *in* R. Levinson (ed.), *Teaching Science*, Routledge, New York, pp. 164–177.

Millar, R. and Hames, V.: 2006, Using research to clarify learning goals and measure outcomes, *in* R. Millar, J. Leach, J. Osborne and M. Ratcliffe (eds), *Improving Subject Teaching: Lessons from Research in Science Education*, Routledge, London, pp. 44–59.

Miller, G. A.: 1956, The magical number seven, plus or minus two: Some limits on our capacity for processing information, *The Psychological Review* **63**, 81–97.

Moje, E. B., Collazo, T., Carrilo, R. and Marx, R. W.: 2001, "Maestro, what is 'quality'?": Language, and discourse in project-based Science, *Journal of Research in Science Teaching* **38**(4), 469–498.

Monge-Najera, J. A., Rossi, M. R. and Mendez-Estrada, V. H.: 2001, Internet, multimedia and virtual laboratories in a 'third world' environment, *Open Learning* **16**(3), 279–190.

Montgomery, M. A.: 2006, How to evaluate textbooks and other instructional materials, *Technical Report Rev. 1.1*, EdVantage Consulting. Available from: http://www.edvantage-consulting.com/white_paper.html [last accessed on 2006-11-12].

Motshekga, A.: 2009, Statement by Minister of Basic Education on the curriculum review process to the National Assembly on 5 November 2009. Available from: http://www.ecdoe.gov.za/news_article/140/We-ve-signed-OBEs-death-certif%icate---Motshekga.

Mozambique Ministry of Education, 2002, Operational guide for textbook evaluators and managers [online]. Available from: http://www.uem.mz/unesco/programmes/education/book_sector_dev/mozambiqu%e_manual_26Nov.pdf [last accessed on 2006-02-15].

Mumba, F., Chabalengula, V. M. and Hunter, W.: 2007, Inquiry levels and skills in Zambian high school Chemistry syllabus, textbooks and practical examinations, *Journal of Baltic Science Education* **6**, 50–57.

Murphy, C., Beggs, J., Hickey, I., O'Meara, J. and Sweeney, J.: 2001, National curriculum: Compulsory school Science – is it improving scientific literacy?, *Educational Research* **43**(2), 189–199.

Muther, C. and Conrad, M.: 1988, Textbooks, *Educational Leadership* pp. 79–80.

Muwanga-Zake, J. W. F.: n.d., Is Science Education in a crisis? Some of the problems in South Africa, *Science in Africa* **2**. Available from: http://www.scienceinafrica.co.za/scicrisis.htm [last accessed on 2010-04-02].

National Academy of Sciences: 1997, *Science for All Children*, National Academy Press, Washington, D.C.

National Institute for Educational Development of the Namibian Ministry of Education: 2005, Guide for the submission and evaluation of textbooks and learning materials. Available from: http://www.nied.edu.na/publications/otherresources/

EvaluationofTextbooksandLearningMaterials2005.pdf [last accessed on 2005-02-14].

National Science Board: 2004, *Science and Engineering Indicators 2004*, National Science Foundation, Chapter 1. Available from: http://www.nsf.gov/statistics/seind04/c1/c1h.htm [last accessed on 2009-12-31].

National Science Resources Center of the National Academy of Sciences: 1998, *Resources for Teaching Middle School Science*, National Academies Press.

Nazarova, T. S. and Gospodarik, I. U. P.: 2006, Strategy for the development of the textbook, *Russian Education and Society* **47**(6), 86–102.

Ncokwane, T. and Prabhala, A.: 2006, The high price of reading, *Mail & Guardian* (16 April 2006).

Nelson, I. T., 1996, A tetrahedral view of accounting education: How can we improve the quality of our graduates? [online]. Available from: http://www.usu.edu/account/faculty/nelson/tet.html [last accessed on 2003-11-16].

Nevada Department of Education, n.d., Textbook evaluation report [online]. Available from: http://www.doe.nv.gov/resources/texbookadoption1/textevaluation.attachm%ent/301418/ta02.pdf [last accessed on 2001-08-14].

New Bruinswick Public Schools, n.d., Textbook evaluation form [online]. Available from: http://www.nbps.k12.nj.us/academics/c&i/pdf/TextbookEvaluationForm.pdf [last accessed on 2007-04-03].

New Jersey Curriculum Framework: 1998, *Analyzing and Selecting Science Curriculum Materials*, New Jersey Department of Education, Chapter 4, pp. 39–48. Available from: http://www.state.nj.us/njded/frameworks/science/ [last accessed on 2008-06-14].

New York Times, 2007, Education sayings, education quotes, education quotations [online]. Available from: http://www.just-quotes.com/education_quotes.html [last accessed on 2007-11-11].

Newton, D. P.: 1990, *Teaching with Text: Choosing, Preparing and Using Textual Materials for Instruction*, Kogan Page, London.

Newton, D. P. and Newton, L. D.: 2006, Could elementary Mathematics textbooks help give attention to reasons in the classroom?, *Educational Studies in Mathematics* **64**, 69–84.

Newton, L. D.: 2003, The occurrence of analogies in elementary school Science books, *Instructional Science* **31**(6), 353–375.

Newton, L. D., Newton, D. P., Blake, A. and Brown, K.: 2002, Do primary school Science books for children show a concern for explanatory understanding?, *Research in Science and Technological Education* **20**(2), 227–240.

Ninnes, P.: 2002, Discursive space(s) in Science curriculum materials in Canada, Australia and New Zealand, *Journal of Curriculum Studies* **34**(5), 557–570.

Nitsche, C. G.: 1992, A teacher and his students examine textbooks, *in* J. G. Herlihy (ed.), *The Textbook Controversy*, Ablex Publishing Corporation, Norwood, New Jersey, Chapter 11, pp. 113–120.

Nogova, M. and Huttova, J.: 2006, Process of development and testing of textbook evaluation criteria in Slovakia, *in* E. Bruillard, B. Aamotsbakken, S. V. Knudsen and M. Horsley (eds), *Caught in the Web or Lost in the Textbook? (Eighth International Conference on Learning and Educational Media)*, Eighth International Conference on Learning and Educational Media.

Nola, R. and Irzik, G.: 2005, *Philosophy, Science, Education and Culture*, Springer, Dordrecht.

Norris, S. P. and Phillips, L. M., n.d., Reading as inquiry [online]. Available from: `http://www.ruf.rice.edu/~rgrandy/Norris.doc` [last accessed on 2004-12-12].

NSTA, 2003, The use and adoption of textbooks in Science teaching [online]. Available from: `http://www.nsta.org/485/` [last accessed on 2003-05-19].

Nurtaç, C., Tacettin, P., Samih, B. and Omer, G.: 2006, The conceptual change approach to teaching chemical equilibrium, *Research in Science & Technological Education* **24**(2), 217–235.

Ogan-Bekiroglu, F.: 2007, To what degree do the currently used Physics textbooks meet the expectations?, *Journal of Science Teacher Education* **18**, 599–628.

Oklahoma State Textbook Committee, n.d.a, Evaluation form for Mathematics K-12 [online]. Available from: `http://www.sde.state.ok.us/pro/textbooks/textbook_pdf/MathEval03.pdf` [last accessed on 2005-07-13].

Oklahoma State Textbook Committee, n.d.b, Instructional materials laws & legislation [online]. Available from: `http://www.sde.state.ok.us/pro/textbooks/textbook_pdf/laws.pdf` [last accessed on 2006-02-22].

Olesko, K. M.: 2006, Science pedagogy as a category of historical analysis: Past, present, and future, *Science & Education* **15**, 963–880.

Olivares-Cuhat, G.: 2002, A touch of . . . class!, *The Canadian Modern Language Review* **58**(4), 653–659.

Olivier, A.: 2009, *Fisiese Wetenskappe Graad 12 (Teorie en werkboek)*, Reivilo Uitgewers, Kroondal.

Olivier, C.: 1997, *Outcomes-Based Education and Training Programmes*, OBET Pro, Ifafi.

Orgill, M. and Bodner, G. M.: 2006, An analysis of the effectiveness of analogy use in College-level Biochemistry textbooks, *Journal of Research in Science Teaching* **43**(10), 1040–1060.

Orpwood, G.: 2001, The role of assessment in Science curriculum reform, *Assessment in Education* **8**(2), 135–151.

Osborne, J.: 2002, Science without literacy: A ship without a sail?, *Cambridge Journal of Education* **32**(2), 203–218.

Osborne, J., Ratcliffe, M., Collins, S. and Duschl, R.: 2006, Specifying curriculum goals: Less of an art, more of a science, *in* R. Millar, J. Leach, J. Osborne and M. Ratcliffe (eds), *Improving Subject Teaching: Lessons from Research in Science Education*, Routledge, London, pp. 27–43.

Otero, J. C. and Campanario, J. M.: 1990, Comprehension evaluation and regulation in learning from Science texts, *Journal of Research in Science Teaching* **27**(5), 447–460.

Pandor, N.: 2006, Making sense of the new Life Sciences curriculum. Address by the Minister of Education, Naledi Pandor, at the launch of learner support material for Natural and Life Sciences curricula, University of the Western Cape, Bellville, 2 March 2006.

Parkinson, J.: 2002, *Reflective Teaching of Science*, Continuum, London.

Peacock, A. and Weedon, H.: 2002, Children working with text in Science: Disparities with literacy hour practice, *Research in Science and Technological Education* **20**(2), 185–197.

Peña, B. M.: 2001, The importance of images in astronomy education, *International Journal of Science Education* **23**(11), 1125–1135.

Penney, K., Norris, S. P. and Clark, G.: 2003, The anatomy of Junior High School Science textbooks: A content analysis of textual characteristics, *Canadian Journal of Science, Mathematics and Technology Education* **3**, 415–436.

Pepin, B. and Haggerty, L.: 2003, Mathematics textbooks and their use by teachers: A window into the education world of particular countries, *in* J. Van den Akker, W. Kuiper and U. Hameyer (eds), *Curriculum Landscapes and Trends*, Kluwer Academic, Dordrecht, pp. 73–100.

Pingel, F.: 1999, UNESCO guidebook on textbook research and textbook revision, *Technical Report ED-99/WS/27*, UNESCO.

Posner, G. J., Strike, K. A., Hewson, P. W. and Gertzog, W. A.: 1982, Accommodation of a scientific conception: Toward a theory of conceptual change, *Science Education* **66**(2), 211–227.

Potenza, E. and Monyokolo, M.: 1999, A destination without a map: Premature implementation of Curriculum 2005?, *in* J. D. Jansen and P. Christie (eds), *Changing Curriculum: Studies on Outcomes-Based Education in South Africa*, Juta & Co, Kenwyn, pp. 231–245.

Powell, J. C. and Anderson, R. D.: 2002, Changing teachers' practice: Curriculum materials and Science Education reform in the USA, *Studies in Science Education* **37**, 107–136.

Pretorius, C.: 2001, To read or not to read: The question mark over textbooks is vital for all our pupils, *Sunday Times* (29 April 2001).

Pretorius, C.: 2009, 'Flawed' tender for textbooks, *Mail & Guardian* (4 December 2009).

Pryor, J. and Lubisi, C.: 2002, Reconceptualising educational assessment in South Africa, *International Journal of Educational Development* **22**, 673–686.

Rallis, S. F. and Bolland, K. A.: 2004, What is program evaluation? Generating knowledge for improvement, *Archival Science* **4**, 5–16.

Ratcliffe, M.: 1999, Evaluation of abilities in interpreting media reports of scientific research, *International Journal of Science Education* **21**(10), 1085–1099.

Ravitch, D.: 2003, What Harry Potter can teach the textbook industry, *Hoover Digest* **2003**(4). Available from: `http://www.hooverdigest.org/034/ravitch.html` [last accessed on 2005-08-16].

Reddy, V.: 2005, State of Mathematics and Science Education: Schools are not equal, *Perspectives in Education* **23**(3), 125–138.

Reddy, V.: 2006, Challenges to meet international Maths and Science standards, *HSRC Review* **4**(1).

Reed, R. F. and Johnson, T. W.: 2000, *Philosophical Documents in Education*, second edn, Longman, New York.

Renner, P.: 2005, *The Art of Teaching Adults: how to become an exceptional instructor and facilitator*, PFR Training Associates, Vancouver.

Reys, B. J. and Reys, R. E.: 2006, The development and publication of elementary Mathematics textbooks: Let the buyer beware!, *Phi Delta Kappan* **87**(5), 377–383. Available from: `http://goliath.ecnext.com/coms2/gi_0199-5400851/The-development-and-pub%lication-of.html` [last accessed on 2010-05-12].

Richaudeau, F.: 1980, *The Design and Production of Textbooks*, Gower, Hampshire.

Ridgeway, V. G. and Dunsion, P. J.: 2000, Content determines process: Seeing the familiar in new ways, *Reading and Writing Quarterly* **16**, 119–138.

Robinson, M.: 2002, Teacher reforms in South Africa: Challenges, strategies and debates, *Prospects* **32**(3), 289–299.

Rogan, J. M. and Grayson, D.: 2003, Towards a theory of curriculum implementation with particular reference to Science Education in developing countries, *International Journal of Science Education* **25**(10), 1171–1204.

Rojewski, J. W.: 2002, Preparing the workforce of tomorrow: A conceptual framework for career and technical education. Paper prepared for the 2002 NCTTEI "A Call for Renaissance in Career and Technical Teacher Education".

Roseman, J. A., Kesidou, S., Stern, L. and Caldwell, A.: 1999, Heavy books light on learning, *Science Books and Films* **35**(6), 243–247. Available from: `http://www.project2061.org/newsinfo/research/textbook/articles/heavy.ht%m` [last accessed on 2006-05-22].

Roseman, J. E., Kulm, G. and Shuttleworth, S.: 2001, Putting textbooks to the test, *ENC Focus* **8**(3), 56–59. Available from: `http://www.project2061.org/newsinfo/research/articles/enc.htm` [last accessed on 2009-10-02].

Roy, B.: 1996, *Multicriteria Methodology for Decision Aiding*, Kluwer Academic, Dordrecht.

Rubin, A.: 2009, *Statistics for Evidence-Based Practice and Evaluation*, Brooks/Cole, Belmont, California.

Rudolph, J. L.: 2000, Reconsidering the 'Nature of Science' as a curriculum component, *Journal of Curriculum Studies* **32**(3), 403–419.

Russell, S. J.: 1997, The role of curriculum in teacher development, *in* S. N. Friel and G. W. Bright (eds), *Reflecting on Our Work: NSF Teacher Enhancement in K-6 Mathematics*, University Press of America, Landham, Maryland, pp. 247–254.

Ruttimann, J.: 2006, Teachers envisage the Science textbooks of the future, *BioED Online* . Available from: http://www.bioedonline.org/news/news.cfm?art=2580 [last accessed on 2007-12-19].

Saaty, R. W.: 1987, The analytic hierarchy process – what it is and how it is used, *Math Modelling* **9**(3-5), 161–176.

Saaty, T. L., 1996, Mathematics and multicriteria decision making [online]. Available from: http://mathforum.org/mam/96/resouces/saaty.html [last accessed on 2001-03-30].

Sandoval, W. A.: 2003, Conceptual and epistemic aspects of students' scientific explanations, *The Journal of the Learning Sciences* **12**(1), 5–51.

Sarinjeve, D., n.d., Why OBE feels like the sword of Damocles in the Humanities and English Studies [online]. Available from: http://www.unisa.ac.za/default.asp?Cmd=ViewContent\&ContentID=11814 [last accessed on 2004-05-12].

Scantlebury, K., Tal, T. and Rahm, J.: 2007, "That don't look like me." Stereotypic images of Science: Where do they come from and what can we do with them?, *Cultural Studies of Science Education* **1**, 545–558.

School Renaissance Institute: 2000, The ATOS Readability Formula for books and how it compares to other formulas, *Technical report*, School Renaissance Institute.

Schramm, W. L.: 1977, *Big Media, Little Media*, Sage Publications, Thousand Oaks, California.

Schultz, T. W.: 1997, Why Johnny won't read – producing readable printed matter, *Frontiers in Education, Annual* **2**, 797–799.

ScienceTextCentral, n.d., Difficulty picking a text? [online]. Available from: http://sciencetextcentral.org/ [last accessed on 2005-08-12].

SciLinks, n.d., Connecting your textbook to the Internet! [online]. Available from: http://www.scilinks.org/tour/default.asp [last accessed on 2005-11-24].

Scriven, M.: 2003, Evaluation theory and metatheory, *in* T. Kellaghan and D. L. Stufflebeam (eds), *International Handbook of Educational Evaluation*, Kluwer Academic, Dordrecht, pp. 9–14.

Segers, M. and Dochy, F.: 2006, Enhancing student learning through assessment: Alignment between levels of assessment and different effects on learning, *Studies in Educational Evaluation* **32**, 171–179.

Sewall, G.: 2005, Textbook publishing, *Phi Delta Kappan* **86**, 498–502.

Shaw, L. J., n.d., Educational philosophies [online]. Available from: http://edweb.sdsu.edu/LShaw/f95syll/philos/phintro.html [last accessed on 2006-01-26].

Shen, B. S. P.: 1975, Science literacy: The public need, *The Sciences* **24**, 27–28.

Sherin, M. G.: 2002, When teaching becomes learning, *Cognition and instruction* **20**(2), 119–150.

Shiland, T. W.: 1997, Quantum Mechanics and conceptual change in High School Chemistry textbooks, *Journal of Research in Science Teaching* **34**(5), 535–545.

Shraw, G.: 1998, Processing and recall differences among seductive details, *Journal of Educational Psychology* **90**(1), 3–12.

Simon, H. A.: 2000, Observations on the science of Science learning, *Journal of Applied Developmental Psychology* **21**(1), 115–121.

Singer, M. and Tuomi, J.: 2003, Selecting instructional materials. A guide for K-12 Science, *Technical report*, National Research Council, Washington, D.C.

Sitte, C., 1999, Das Geographie-Schulbuch [online]. Available from: http://www.univie.ac.at/geographie/ifgr/stzw/lehramt/fachdidaktik/home/%chsSCHULBUCH [last accessed on 2003-03-27].

Skevington, J.: 1994, Assessing and evaluating in Science Education, *in* R. Levinson (ed.), *Teaching Science*, Routledge, London and New York, pp. 115–128.

Sleeter, C. E. and Grant, C. A.: 1991, Race, class, gender and disability in current textbooks, *in* M. W. Apple and L. K. Christian-Smith (eds), *The Politics of the Textbooks*, Routledge, New York, pp. 78–110.

Smith, C., Maclin, D., Grosslight, L. and Davis, H.: 1997, Teaching for understanding: A study of students' pre-instruction theories of matter and a comparison of the effectiveness of two approaches to teaching about matter and density, *Cognition and Instruction* **15**(3), 317–393.

Smith, E. L., Blakeslee, T. D. and Anderson, C. W.: 1993, Teaching strategies associated with conceptual change learning in Science, *Journal of Research in Science Teaching* **30**(2), 111–126.

Smith, R. and Moore, T.: 2005, The context of increasing accountability and what it means for independent schools. Lecture at the AISQ 2005 Biennial State Conference, Townsville, Queensland, 27–29 July 2005.

Solomon, J.: 1994a, Group discussions in the classroom, *in* R. Levinson (ed.), *Teaching Science*, Routledge, London and New York, pp. 76–84.

Solomon, J.: 1994b, The laboratory comes of age, *in* R. Levinson (ed.), *Teaching Science*, Routledge, London and New York, pp. 7–21.

Sorrels, R. C. and Britton, B. K.: 1998, What is the point? Tests of a quick and clean method or improving instructional text, *in* C. R. Hynd (ed.), *Learning from Text Across Conceptual Domains*, Lawrence Erlbaum Associates, Mahwah, New Jersey, Chapter 5, pp. 95–115.

Spady, W. G.: 1994, *Outcome Based Education: Critical Issues and Answers*, American Association of School Administrators, Arlington, Virginia.

Spady, W. G. and Schlebusch, A.: 1999, *Curriculum 2005: A Guide for Parents*, Renaissance, Cape Town.

Stahl, S. A.: 2003, Vocabulary and readability: How knowing word meanings affects comprehension, *Topics in Language Disorders* **23**(3), 241–247.

Stahl, S., Jacobson, M. G., Davis, C. E. and Davis, R. L.: 2006, Prior knowledge and difficult vocabulary in the comprehension of unfamiliar text, *in* K. A. D. Sathl and M. C. McKenna (eds), *Reading Research at Work: Foundations of Effective Practice*, The Guildford Press, New York, pp. 284–302.

Stein, M., Stuen, C., Carnine, D. and Long, R. M.: 2001, Textbook evaluation and adoption practices, *Reading and Writing Quarterly* **17**, 5–23.

Stern, L. and Ahlgren, A.: 2002, Analysis of students' assessment in Middle School curriculum materials: Aiming precisely at benchmarks and standards, *Journal of Research in Science Teaching* **39**(9), 889–910.

Stern, L. and Roseman, E.: 2001, Textbook alignment, *The Science Teacher* **68**, 52–57.

Steyn, G. M.: 2004, Problems of, and support for, beginner educators, *Africa Education Review* **1**(1), 81–94.

Steyn, P. and Wilkinson, A.: 1998, Understanding the theoretical assumptions of outcomes-based education as a condition for meaningful implementation, *South African Journal of Education* **18**(4), 203–208.

Stinner, A.: 1992, Science textbooks and Science teaching: From logic to evidence, *Science Education* **76**(1), 1–16.

Stinner, A.: 1995, Science textbooks: Their present role and future form, *in* S. M. Glynn and R. Duit (eds), *Learning Science in the Schools*, Lawrence Erlbaum Associates, Mahwah, New Jersey, pp. 275–296.

Stoffels, N. T.: 2005, 'Sir, on what page is the answer?' Exploring teacher decision-making during complex curriculum change, with specific reference to the use of learner support material, *International Journal of Educational Development* **25**(5), 531–546.

Stylianidou, F., Ormerod, F. and Ogborn, J.: 2002, Analysis of Science textbook pictures about energy and pupils' reading of them, *International Journal of Science Education* **24**(3), 257–283.

Sukhraj, P., Mkhize, T. and Govender, S.: 2004, Untrained teachers let loose on our kids, *Sunday Times* (8 February 2004), 1.

Sulaiman, J. and Dwyer, F.: 2002, The effect of varied instructional text design strategies on the achievement of different educational objectives, *International Journal of Instructional Media* **29**(2), 215–223.

Sursock, A.: 2001, Quality and innovation in Higher Education, *in* N. Baijnath, S. Maimela and P. Singh (eds), *Quality Assurance in Open and Distance Learning*, University of South Africa and Technikon South Africa, Pretoria, pp. 81–93.

Sweet, A. P.: 2005, Assessment of reading comprehension: The RAND reading study group vision, *in* S. G. Paris and S. A. Stahl (eds), *Children's Reading Comprehension and Assessment*, Lawrence Erlbaum Associates, Mahwah, New Jersey, pp. 13–69.

Tamir, P.: 1999, Self-assessment: The use of self-report knowledge and opportunity to learn inventories, *International Journal of Science Education* **21**(4), 401–411.

Tansey, M., n.d., Curriculum and the human race [online]. Available from: http://www.people.ex.ac.uk/PErnest/pome14/tansey.htm [last accessed on 2006-02-15].

Tarr, J. E., Chávez, O., Reys, R. E. and Reys, B. J.: 2006, From the written to the enacted curriculum: The intermediary role of Middle School Mathematics teachers in shaping students' opportunity to learn, *School Science & Mathematics* **106**(4), 191–201.

Task Team for the Review of the Implementation of the National Curriculum Statement: 2009, Final report on the implementation of the national curriculum statement, *Technical report*, South African Department of Education.

Taylor, N.: 2008a, It's OBE, but not as it should be, *Mail & Guardian* (15 April 2008).

Taylor, N., 2008b, What's wrong with South African schools? [online]. Available from: http://www.jet.org.za [last accessed on 2009-03-19].

Tella, S.: 1998, The concept of media education revisited: From a classificatory analysis to a rhizomatic overview, *Media Education Publication* **8**, 85–150.

The Centre for Curriculum Materials in Science: 2006, Preparing the next generation of curriculum material leaders: A proposal to the National Science Foundation for a Center for Learning and Teaching with a focus on instructional materials in Science. Available from: http://ed-web2.educ.msu.edu/CCMS/Webpages/Proposal.htm [last accessed on 2010-05-12].

The University of Sydney: n.d., Teaching resources and textbook research unit (TREAT). Available from: http://www.edsw.usyd.edu.au/research/current/centres/treat.shtml [last accessed on 2007-09-28].

Thiede, K. W., Anderson, M. C. M. and Thierriault, D.: 2003, Accuracy of metacognitive monitoring affects learning of texts, *Journal of Educational Psychology* **95**(1), 66–73.

Tobin, K., Tippins, D. J. and Gallard, A. J.: 1994, Research on instructional strategies for teaching Science, *in* D. L. Gabel (ed.), *Handbook of Research on Science Teaching and Learning*, Macmillan, New York, Chapter 2, pp. 45–93.

Törnroos, J.: 2005, Mathematics textbooks, opportunity to learn and student achievement, *Studies in Educational Evaluation* **31**, 315–327.

Tyson, H.: 1997, *Overcoming Structural Barriers to Good Textbooks*, Inverness Research Associates, Inverness, California.

Ulerick, S. L., n.d., Using textbooks for meaningful learning in Science [online]. Available from: http://www2.educ.sfu.ca/narstsite/publications/research/textbook2.htm [last accessed on 2003-10-30].

UNESCO: 2002, *Teacher Education Guidelines: Using Open and Distance Learning*, Unesco Publishing, Paris.

UNESCO Regional Bureau for Education in Africa: 2003, Education for all in Africa bulletin, *Technical report*, UNESCO. Available from: http://www.dakar.unesco.org/efa/bulletin/index_en.shtml [last accessed on 2007-12-31].

University of Alabama, n.d., Science textbook review [online]. Available from: http://bama.ua.edu/~walto005/586/science_textbook_review.htm [last accessed on 2005-05-02].

Unsworth, L.: 2001, Evaluating the language of different types of explanations in Junior High School Science texts, *International Journal of Science Education* **23**(6), 585–609.

Van de Grift, W.: 1989, Evaluatie van leermiddelen, *in* A. Reints, R. A. De Jong and N. A. J. Lagerweij (eds), *Om de Kwaliteit van het Leermiddel*, Zwijsen, Tilburg, pp. 163–166.

Van den Berg, G.: 2004, *Die Geleenthede wat uitkomsgebaseerde taalhandboeke bied vir die ontwikkeling van leerders se meervoudige intelligensies*, PhD thesis, University of South Africa.

Van Hout-Wolters, B. and Schnotz, W.: 1992, *Text Comprehension and Learning from Text: An Introduction*, Swets & Zeitlinger, Amsterdam.

Van Schalkwyk, A.: 1993, *Onderwysmedia in die formele onderrig- en leergebeure: 'n Histories-pedagogiese ondersoek en evaluering*, PhD thesis, University of Pretoria.

Vekiri, I.: 2002, What is the value of graphical displays in learning?, *Educational Psychology Review* **14**(3), 261–312.

Vercellis, C.: 2009, *Business Intelligence*, John Wiley and Sons, Chichester, West Sussex.

Verspoor, A.: 1991, Twenty years of world bank support for basic education: Presentation and evaluation, *Prospects* **21**(3), 313–329.

Viadero, D.: 2007, 'What works.' Reviewers find no learning edge for leading math texts, *Education Week* **26**(20), 1–2.

Villaverde, L. E.: 2003, *Secondary Schools: A Reference Handbook*, ABC-CLIO, Santa Barbara, California.

Vogler, K. E.: 2006, Impact of High School graduation examination on Tennessee Science teachers' instructional practices, *American Secondary Education* **35**(1), 33–57.

Vosniadou, S.: 2001a, Conceptual change – Teaching and learning process, *in* H. Behrendt, H. Dahncke, R. Duit, W. Gräber, M. Komorek, A. Kross and P. Reiska (eds), *Research in Science Education – Past, Present, and Future*, Kluwer Academic, Dordrecht, pp. 177–188.

Vosniadou, S.: 2001b, How children learn, *Technical report*, International Academy of Education.

Walford, G.: 1985, The construction of a curriculum area: Science in society, *British Journal of Sociology of Education* **6**(2), 155–171.

Walliman, N.: 2005, *Your Research Project*, second edn, SAGE, London.

Walsh, M.: 2006, The 'textual shift': Examining the reading process with print, visual and multimodal texts, *Australian Journal of Language and Literacy* **29**(1), 24–37.

Watkins, M.: 2000, Ways of learning about leisure meanings, *Leisure Sciences* **22**, 93–107.

Watts, D. and Simon, S.: 1999, Textual support for primary Science teaching?, *The Curriculum Journal* **10**(3), 385–401.

Weinbrenner, P.: 1992, Methodologies of textbook analysis used to date, *in* H. Bourdillon (ed.), *History and Social Studies – Methodologies of Textbook Analysis*, Swets & Zeitlinger, Amsterdam, pp. 21–34.

Weiten, W., Deguara, D., Rehmke, E. and Sewell, L.: 1999, University, community college, and high school students' evaluations of textbook pedagogical aids, *Teaching of Psychology* **26**(1), 19–21.

Wenham, M.: 1995, *Understanding Primary Science. Ideas, Concepts and Explanations*, Paul Chapman, London.

Western Cape Department of Education: 2005, Information regarding the supply of selected textbooks for grade 10 Mathematics, Science and Languages, Circular 0032/2005.

White, B. Y. and Frederiksen, J. R.: 1998, Inquiry, modeling, and metacognition: Making Science accessible to all students, *Cognition and Instruction* **16**(1), 3–118.

Whitman, D.: 2004, The mad, mad world of textbook adoption, *Technical report*, Thomas B Fordham Institute.

Wieman, C.: 2007, Why not try a scientific approach to Science Education?, *Change* (September–October 2007), 9–15.

Wikipedia, n.d.a, Assessment [online]. Available from: `http://en.wikipedia.org/wiki/Assessment` [last accessed on 2008-01-20].

Wikipedia, n.d.b, Educational Perenialism [online]. Available from: `http://en.wikipedia.org/wiki/Educational_perennialism` [last accessed on 2007-03-12].

Wikipedia, n.d.c, Existentialism [online]. Available from: `http://en.wikipedia.org/wiki/Existentialism` [last accessed on 2006-03-12].

Wikipedia, n.d.d, Pragmatism [online]. Available from: `http://en.wikipedia.org/wiki/Pragmatism` [last accessed on 2007-03-12].

Wilen, W., Hutchinson, J., Bosse, M. I. and Kindsvatter, R.: 2004, *Dynamics of Effective Secondary Teaching*, Allyn & Bacon, Boston, Massachusetts.

Wilson, D. R., n.d., Ways in which the language of school teaching materials can impede learning [online]. Available from: `www.specialeducationalneeds.com/sen/readability.doc` [last accessed on 2010-02-17].

Wilson, T.: 1997, Electronic publishing and the future of the book, *Information Research* **3**(2). Available from: `http://informationr.net/ir/3-2/paper39.html` [last accessed on 2008-11-12].

Wingate, L. A.: 2002, The evaluation checklist project: The inside scoop on content, process, policies, impact, and challenges. Paper presented at the *American Evaluation Association Conference* in Washington, D.C.

Winston, W. L.: 2004, *Operations Research: Applications and Algorithms*, fourth edn, Brooks/Cole, Belmont, California.

Wong, S. L.: 1991, Evaluating the content of textbooks: Public interest and professional authority, *Sociology of Education* **64**, 11–18.

Woodward, A. and Elliott, D. L.: 1990, Textbooks: Consensus and controversy, *in* D. L. Elliot and A. Woodward (eds), *Textbooks and Schooling in the United States*, University of Chicago Press, Chicago, Illinois, pp. 146–161.

Yerrick, R. K., Doster, E., Nugent, J. S., Parke, H. M. and Crawley, F. E.: 2003, Social interaction and the use of analogy: An analysis of preservice teachers' talk during Physics inquiry lessons, *Journal of Research in Science Teaching* **40**(5), 443–463.

Yore, L. D.: 2000, Enhancing Science literacy for all students with embedded reading instruction and writing-to-learn activities, *Journal of Deaf Studies and Deaf Education* **5**(1), 105–122. Available from: `http://deafed.oupjournals.org/cgi/reprint/5/1/105.pdf` [last accessed on 2004-01-13].

Appendix A

Readability formulæ

1. The Dale-Chall readability formula (Dale and Chall, 1948) considers both the semantic and syntactical aspects of the text.

   A passage of 100 words is selected and the average sentence length (L) and the percentage of familiar words (F) are calculated. These familiar words are words that feature on the Dale-Chall list of 3000 words that should be familiar to fourth grade students (School Renaissance Institute, 2000). The Dale-Chall Readability Index is then calculated using the equation:

   Raw score = $(0,0496 \times \mathrm{L}) + (0,1579 \times \mathrm{F}) + 3,6365$

   This raw score is then converted to a grade level using the Dale-Chall table:

Raw score	Grade level
4,9 and below	4th grade and below
5,0 to 5,9	5–6th grade
6,0 to 6,9	7–8th grade
7,0 to 7,9	9–10th grade
8,0 to 8,9	11–12th grade
9,0 to 9,9	13–15th grade (College)
10,0 and above	16 (College graduate)

2. The Gunning-FOG readability test (Johnson, n.d.) analyses three samples of 100 words each. The syntactic difficulty is indicated by the average sentence length (L) and the semantic difficulty by the average number of words with three or more syllables (N). The grade level is then determined by $(L + N) \times 0,4$.

3. The Fry readability graph is one example of a technique where a graph is used in the place of a formula (Mikk, 2000). In samples of 100 words the av-

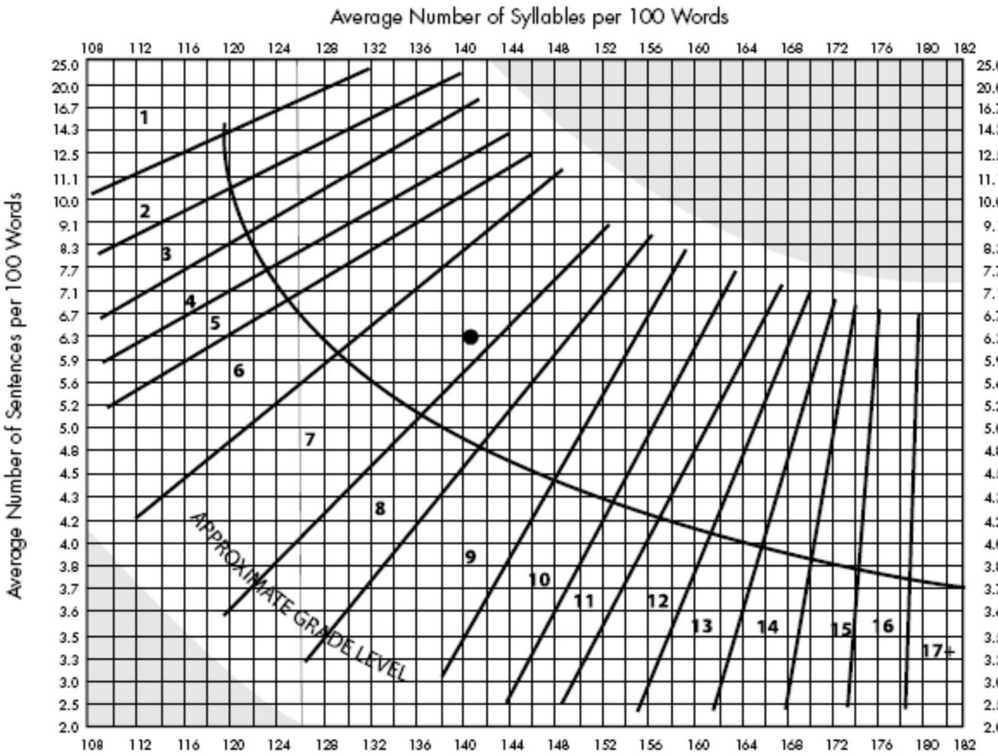


Figure A.1: Fry readability graph (Burns, 2006)

erage number of sentences per 100-word passage (y) (syntactical difficulty) and the average number of syllables per 100-word sample (x) (semantic difficulty) are calculated. These values are then used to read the reading age (in years) from the Fry graph in Figure A.1. The curve represents normal text. Points below the curve imply longer than average sentence lengths and points above the curve represent text with a more difficult vocabulary.

4. The Flesch reading ease test (Giles and Still, 2005) systematically selects 100-word samples. The number of syllables per 100 words (wl) (semantic difficulty) and the average number of words per sentence (sl) (syntactical difficulty) are calculated. The reading ease is then determined with the following equation:

   Reading ease $= 206,835 - 0,846wl - 1,015sl$

   The numerical value rendered by this formula correlates with the description of the readability on Flesch's scale:

Reading ease	Description
95 – 100	very easy
85 – 94	easy
75 – 84	fairly easy
65 – 74	standard
55 – 64	fairly difficult
40 – 54	difficult
0 – 39	very difficult

5. The Flesch-Kincaid formula was developed by the US Government Department of Defence (Johnson, n.d.). Both Microsoft Windows and Corel WordPerfect include the Flesch-Kincaid formula (Giles and Still, 2005). The average sentence length (L) and the average number of syllables per word (N) are calculated. Then the grade level is given by $0,39L + 11,8N - 15,59$.

6. The 'SMOG' (Simple Measure of Gobbledygook) formula of McLaughlin (Johnson, n.d.; Newton, 1990) is a readability formula that attempts to estimate the number of years of education needed to completely understand a text. It selects samples of 10 consecutive sentences from the beginning, middle and end of the text. The average number of words with three or more syllables (N) is calculated. Then the grade level is given by $1,043 \times \sqrt{N} + 3,1291$ and the reading age can be estimated by $\sqrt{N} + 8$ years. This test tends to give higher values than the other formula, because it is intended to predict the level necessary for 100% comprehension of the text (Johnson, n.d.).

7. The FORCAST formula was developed for US army technical manuals (Johnson, n.d.), and, since it is not based on a sentence count, can be used to determine the reading ease for texts with multiple choice questions or summaries in list format. Samples of 150 words are selected and the number of single-syllable words (N) is counted. Then the grade level is estimated by $= 20 - (N \div 10)$ and the reading age by $25 - N \div 10$.

8. The Lexile framework was developed by MetaMetrics, a large commercial company.(Fry, 2002; School Renaissance Institute, 2000). This is not an open standard, and it is not known exactly how the Lexile measure is calculated.

9. The Advantage-TASA Open Standard (ATOS) is a readability formula for books developed by the School Renaissance Institute (formerly Advantage Learning Systems) (Fry, 2002; School Renaissance Institute, 2000). They

did research that correlated text difficulty and the various variables (and combinations of variables) that readability researchers use as indicators of text difficulty. Their research showed that a combination of words per sentence, average grade level of words and characters per word are the best indicators of text difficulty. The entire text is analysed (Burns, 2006), and not just a sample. An expanded graded-vocabulary list (School Renaissance Institute, 2000) was compiled to enhance the accuracy of this formula.

Appendix B

Evaluation instruments

B.1 The evaluation instrument of Namibia

(National Institute for Educational Development of the Namibian Ministry of Education, 2005)

MINISTRY OF EDUCATION
NATIONAL INSTITUTE for EDUCATIONAL DEVELOPMENT

REPUBLIC OF NAMIBIA

Addendum 3: Textbook Evaluation Instrument (Natural Sciences)

TEXTBOOK EVALUATION INSTRUMENT
(Natural Sciences textbooks and teaching materials)

Subject: ……………………………………………………………

GRADE:……………… TITLE:……………………………………………..

All 4 categories should be completed to contribute to the final mark. Grading is between 5 (high) and 1 (low). Completed forms are confidential documents. For every grading an explanation must be given.

1.	**PHYSICAL CHARACTERISTICS**	
1.1	**Durability:** Quality of paper and binding. Is it likely that the book will be usable for 5 years? ……………………………………………… ……………………………………………… ………………………………………………(5 Yes or 1 No)	
1.2	**Typeface and size:** Is the type, including captions and labeling of illustrations clear and large enough for learners intended? ……………………………………………… ……………………………………………… ……………………………………………… (5 Yes or 1 No)	
1.3	**Layout and appearance:** Is the book aesthetically appealing? • aesthetically should be seen as relevant and usable (since the criteria is subjective) Look at general layout, width of margins, etc.……………………………………………… ……………………………………………… ……………………………………………… ………………………………………………(5 Yes or 1 No)	
1.4	**Cost:** Is the cost reasonable compared to similar materials? ……………………………………………… ……………………………………………… ……………………………………………… ………………………………………………(5 Yes or 1 No)	

Sub-total	➔	
% Mark: Subtotal x 100 / 20	➔	

10

2.	CONTENT (Please note 2.1 to 2.5 are counting out of 10)	
2.1	How consistent is the approach used in the book with the syllabus? : • Does the book integrate relevant learner activities and teacher demonstrations? ……… … ……… ……… ……… ………………………………………………………………………………(10 Fully to 1 Scarcely)	
2.2	To what extent are relevant knowledge objectives catered for in the book? • Does the book cover competencies (syllabus topics and objectives)? • Does the book follow a policy accepted pattern (LCA e.g. social constructivist approach/co-operative learning)? …… ……… ……… ……… …………………………………………………………………………….(10 Fully to 1 Scarcely)	
2.3	To what extent are relevant skills objectives catered for in the book? relevant skills seen as: • critical thinking, • practical skills, such as observation, recording, • processing information, • problem solving • interpretation, • graphing skills ……………………………………………………………………………………………………… ……… ……… ……… ………………………………………………………………………………(10 Fully to 1 Scarcely)	
2.4	To what extent does the content reflect current knowledge and culture? • Does the text book reflect some Namibian (science related) examples (such as mining, treatment of animal skins to make fur, beer-making etc.) • Does the book reflect current issues/controversies (renewable energy sources, environmental conservation, recycling, water treatment, food technology, population, genetic engineering etc.) ……… ……… ……… …………………………………………………………………………….(10 Fully to 1 Scarcely)	
2.5	To what extent is the content of the book factually accurate? ……… ……… ……… ……… ……………………………………………………………………………….(10 Fully to 1 Scarcely)	
2.6	How free is the book of biases unacceptable to teachers, learners, communities and MINISTRY OF EDUCATION policies? ……… ……………………………………………………...………………………………………… ………………………………………………………………………….(5 Entirely to 1 Scarcely)	
2.7	To what extent do the materials reflect the contributions and perspectives of various ethnic and cultural groups where appropriate? ……… ……… ……… ………………………………………………………………………………(5 Fully to 1 Scarcely)	

2.8	Is the book free of sex stereotypes? …………………………………………………………………………………………… …………………………………………………………………………………………… …………………………………………………………………………………………… ……………………………………………………………………………………(Yes 5 or No 1)	
2.9	To what extent does the material encourage a positive attitude towards gender? • E.g. stereotyping occupation or use of gender bias words like chairman instead of chairperson …………………………………………………………………………………………… …………………………………………………………………………………………… …………………………………………………………………………………………… ………………………………………………………………………………(5 Largely to 1 Scarcely)	
2.10	To what extent does the material encourage a positive attitude towards environmental issues? …………………………………………………………………………………………… …………………………………………………………………………………………… …………………………………………………………………………………………… ………………………………………………………………………………(5 Largely to 1 Scarcely)	
2.11	To what extent does the book encourage a positive attitude towards population issues? …………………………………………………………………………………………… …………………………………………………………………………………………… …………………………………………………………………………………………… ………………………………………………………………………………(5 Largely to 1 Scarcely)	
	Sub-total ➔	
	% Mark: <u>Subtotal x 100</u> ➔ ***80***	
3.	**PEDAGOGICAL ASPECTS**	
3.1	To what extent is the content of the book likely to be clearly understood by the learners? …………………………………………………………………………………………… …………………………………………………………………………………………… …………………………………………………………………………………………… ………………………………………………………………………………(5 Largely to 1 Hardly)	
3.2	How helpful are the tests and other assessment devices (such as practice exercises, end-of-chapter questions and other assessment devices, experiments, etc.) in the book to the **teacher**? …………………………………………………………………………………………… …………………………………………………………………………………………… …………………………………………………………………………………………… ………………………………………………………………………………(5 Very to 1 Not much)	
3.3	How helpful are the tests and other assessment devices (such as practice exercises, end-of-chapter questions and other assessment devices, experiments, etc.) in the book to the **learner**: …………………………………………………………………………………………… …………………………………………………………………………………………… …………………………………………………………………………………………… ………………………………………………………………………………(5 Very to 1 Not much)	
3.4	Does the design of the materials allow teachers to use them differently according to the needs of different learners …………………………………………………………………………………………… …………………………………………………………………………………………… …………………………………………………………………………………………… ………………………………………………………………………………(5 Very to Not much)	

☐

3.5	Does the book cater for skills development toward data analysis and problem solving ? …………………………………………………………………………………………… …………………………………………………………………………………………… …………………………………………………………………………………………… ………………………………………………………………………(5 Fully to 1 Scarcely)	
3.6	Is the use of the book or material easily manageable by the teacher? …………………………………………………………………………………………… …………………………………………………………………………………………… …………………………………………………………………………………………… ……………………………………………………………………(5 Largely or 1 Not easily)	
3.7	Does the book include activities that learners are capable of performing and will find stimulating, interesting and rewarding? …………………………………………………………………………………………… …………………………………………………………………………………………… …………………………………………………………………………………………… ……………………………………………………………………(5 Largely to 1 Scarcely)	
3.8	Does the book use appropriate tables, diagrams, charts, sketches and photographs to explain the content? …………………………………………………………………………………………… …………………………………………………………………………………………… …………………………………………………………………………………………… ……………………………………………………………………………………(5 Yes to 1 No)	
3.9	Does the book contain a table of content and an adequate index? : • table of contents, • adequate index, • comprehensive appendix on units and conversions …………………………………………………………………………………………… …………………………………………………………………………………………… …………………………………………………………………………………………… ……………………………………………………………………………………(5 Yes or 1 No)	
3.10	Is special equipment, **not** readily available at all schools , required in order to use the book effectively? …………………………………………………………………………………………… …………………………………………………………………………………………… …………………………………………………………………………………………… ……………………………………………………………………………………(5 No or 1 Yes)	
3.11	Does the book support learner-centred approach to teaching (assuming that this is the approach the national curriculum favours)? …………………………………………………………………………………………… …………………………………………………………………………………………… …………………………………………………………………………………………… …………………………………………………………………………………(5 Yes or 1 No)	
3.12	Is the book of appropriate length? …………………………………………………………………………………………… …………………………………………………………………………………………… …………………………………………………………………………………………… …………………………………………………………………………………. (5 Yes or 1 No)	

Sub-total	➔
% Mark: Subtotal x 100 **60**	➔

4.	**LANGUAGE LEVELS**	
4.1	Is the book written at an appropriate reading level for the learners ? …………………………………………………………………………………………………… …………………………………………………………………………………………………… …………………………………………………………………………………………………… ………………………………………………………………(5 Yes or 1 Too difficult or too easy)	
4.2	Is the book written at an appropriate comprehension level for the learners? …………………………………………………………………………………………………… …………………………………………………………………………………………………… …………………………………………………………………………………………………… ………………………………………………………………(5 Yes or 1 Too difficult or too easy)	
4.3	Are new and challenging concepts defined in a glossary or explained when they are first introduced in the text? …………………………………………………………………………………………………… …………………………………………………………………………………………………… …………………………………………………………………………………………………… ……………………………………………………………………………………… (5 Yes to 1 No)	
	Sub-total ➔	
	% Mark: <u>Subtotal x 100</u> ➔ ***15***	

Evaluator's name : ➾

____

Categories	**1**	**2**	**3**	**4**
TOTALS				
%				

Final mark: ____ / **175**

B.2 The evaluation instrument of the Gauteng Department of Education

(Mahlaba, 2006)

ADDENDUM A: EVALUATION INSTRUMENT

Evaluation Instrument for Learning Support Materials
For Foundation Phase

(REVISED NATIONAL CURRICULUM STATEMENTS)
2004

Learning area:	Grade:

Publisher:		Language:

Author/s :	

Title:		Number of components:	

Educators' Guide		Learners' Book		Work Book		Reader(s)	

Audiovisuals:	Charts		Poster		Models		Toys		Puzzles		Games	

Subtitles:	ISBN:	Price:

Evaluators:

Name	Signature

65

Team Leader(s)

Name	**Signature**

EVALUATION	**TOTAL**	**TICK**	**CROSS**
Generic Evaluation- Section A			
Learning Area Evaluation – Section B			

<u>FINAL DECISION</u>: Approve two ticks
Approved **Not approved**

Date: ____________________

SECTION A GENERIC EVALUATION

Rating scale:Very little = 1 Partially = 2 Mostly = 3 Very much so = 4

Circle the score that best describes the material/s.

1.Compliance with the RNCS	**Weight= 10%**
1.1 The Critical and Developmental Outcomes are covered in the LTSM.	1 2 3 4
1.2 All the Learning Outcomes of the Learning Area are covered effectively and comprehensively in terms of content and activities.	1 2 3 4
1.3 The materials are in line with the Constitution in terms of sensitivity to issues of poverty, inequality, race, gender, age, disability and HIV/AIDS.	1 2 3 4
1.4 The materials cater for learners with varied abilities, as stipulated in the policy on inclusion / Education White Paper 6.	1 2 3 4
$\frac{\text{Total Score}}{16} \times \frac{10}{1}$	

2.Assessment	**Weight= 10%**
2.1 The assessment activities accommodate diverse contexts, e.g. diverse cultures, environments and religions.	1 2 3 4
2.2 The assessment activities allow for learners to integrate and apply knowledge and skills.	1 2 3 4
2.3 Different assessment methods have been used, e.g. self, peer, group, educator, etc to cover the various aspects of the learner's performance.	1 2 3 4
2.4 The Assessment Activities assist the learners to set their own goals for progress.	1 2 3 4
2.5 The material assess thinking skills at different levels	1 2 3 4
$\frac{\text{Total Score}}{20} \times \frac{10}{1}$	

3. Integration of Learning Areas	Weight = 5%
3.1 The principles of the RNCS such as social justice, healthy environment, Human rights etc. are inherent in the LTSM and integrated across various Learning Areas.	1 2 3 4
3.2 The Learning Outcomes and Assessment Standards of different Learning Areas are integrated in the different activities.	1 2 3 4
$\frac{\text{Total Score}}{8} \times \frac{5}{1}$	

Rating scale:Very little = 1   Partially = 2   Mostly = 3   Very much so = 4

Circle the score that best describes the material/s.

4. Approach	Weight = 5%
4.1 The learning and teaching strategies used in the LTSM are appropriate for the content i.e. there is a proper balance between pair, group, class and educator – led activities.	1 2 3 4
4.2 There are various types of activities, e.g. discussions, debates, projects, etc to accommodate learners from different backgrounds, at different levels of development and with differing language proficiencies.	1 2 3 4
4.3 The various activities enhance learning and encourage learners to be active and constructive participants.	1 2 3 4
$\frac{\text{Total Score}}{12} \times \frac{5}{1}$	

5. Progression and Relationship	Weight = 5%
5.1 There is progression within and between activities eg. the activity on identifying dangers at home should precede a discussion on how to maintain safety at home.	1 2 3 4
5.2 There is a meaningful conceptual relationship between activities.	1 2 3 4
$\frac{\text{Total Score}}{8} \times \frac{5}{1}$	

6. Context	Weight = 10%
6.1 The learner's prior knowledge in terms of experience and context is sufficiently acknowledged as a starting point.	1 2 3 4
6.2 The content and activities provide opportunities for learners to discover a world beyond their own.	1 2 3 4
6.3 There is a balance between information familiar to the learners and new knowledge both in illustrations and in text.	1 2 3 4
6.4 Provision has been made for recognition of various social structures, e.g. if the material only makes reference to Christian beliefs, learners whose families practice other religions will feel excluded.	1 2 3 4
$\frac{\text{Total Score}}{16} \times \frac{10}{1}$	

Rating scale:Very little = 1   Partially = 2   Mostly = 3   Very much so = 4

Circle the score that best describes the material/s.

7. Content	Weight = 10%
7.1 The content is unbiased and free of prejudice. It should avoid gender stereotyping, racial inferences, offensive textual and visual material.	1 2 3 4
7.2 The activities geared towards the attainment of the Assessment Standards correlate with the age and level of the learner. (The activities should be neither too simple nor too complex.)	1 2 3 4
7.3 There is a fair representation of different cultural groups in illustrations, names of characters and text.	1 2 3 4
7.4 The content is accurate, up to date, coherent and conceptually sound.	1 2 3 4
7.5 The content is based in the South African context and also including updated information where relevant.	1 2 3 4
$\frac{\text{Total Score}}{20} \times \frac{10}{1}$	

8. Relationship between content, illustrations and activities	Weight = 5%
8.1 The activities are sufficiently related to the text and/or illustrations that coincide with them.	1 2 3 4
8.2 Content, illustrations and related activities address the same outcome. However, it is possible that the activities incorporate other Learning Outcomes through the principle of integration.	1 2 3 4
8.3 The activities prompt the learners to participate in different ways to avoid monotony, e.g. a reading text should not always be followed by questions to be answered orally.	1 2 3 4
$\frac{\text{Total Score}}{12} \times \frac{5}{1}$	

9. Language and Style	Weight = 10%
9.1 The vocabulary used in the materials is clear, accessible and appropriate for target group.	1 2 3 4
9.2 New concepts and information are introduced in a clear and logical manner.	1 2 3 4
9.3 The language used does not discriminate against certain races.	1 2 3 4
9.4 The language used is sensitive to issues of gender.	1 2 3 4
9.5 The language used does not denigrate certain religions and cultures.	1 2 3 4
9.6 The language used does not encourage stereotyping.	1 2 3 4
9.7 The language caters for learners with a variety of language backgrounds and abilities	1 2 3 4
$\frac{\text{Total Score}}{28} \times \frac{10}{1}$	

Rating scale:Very little = 1   Partially = 2   Mostly = 3   Very much so = 4

Circle the score that best describes the material/s.

10 Layout and Design	Weight = 10%
The page layout is suitable in terms of:	
10.1 clear headings and sub-headings.	1 2 3 4
10.2 clearly marked activities.	1 2 3 4
10.3 suitable print size and type.	1 2 3 4
10.4 adequate white space and margins.	1 2 3 4
The manner in which the design features are used:	
10.5 is appealing and attractive to the eye of the learner.	1 2 3 4
10.6 draws attention and holds interest for learners to interact continuously with the material.	1 2 3 4
$\frac{\text{Total Score}}{24} \times \frac{10}{1}$	

11.Pictures and Illustrations	Weight = 5%
11.1 The pictures and illustrations are relevant to the activities being dealt with, i.e. that pictures and illustrations are not there as gap-fillers or merely for decorative purposes.	1 2 3 4
11.2 The pictures and illustrations contain all the details mentioned in the activities	1 2 3 4
11.3 The pictures and illustrations are representative of different races and genders as well as people with special needs, where necessary and appropriate.	1 2 3 4
$\frac{\text{Total Score}}{12} \times \frac{5}{1}$	

12. Educator Support	Weight = 15%
12.1 There is clear reference and cross-reference to all the components of the package.	1 2 3 4
12.2 The educator is encouraged to adapt material to suit his/her individual situation and context	1 2 3 4
The educators' guide gives:	
12.2 Brief overview of the curriculum.	1 2 3 4
12.3 Guidance to the educator on how to facilitate activities.	1 2 3 4
12.4 Guidance to the educator on the application of different assessment methods.	1 2 3 4
12.5 Solutions to problems.	1 2 3 4
12.6 Guidance to educators to appropriate methods and learning strategies	1 2 3 4
$\frac{\text{Total Score}}{28} \times \frac{15}{1}$	

TOTAL SECTION A SCORE:

1. Add the scores to get a total

B.3 The evaluation instrument of South African National Department of Education

(Department of Education of South Africa, n.d.)

education
Department of Education
REPUBLIC OF SOUTH AFRICA

GUIDING CRITERIA FOR SELECTING TEXTBOOKS

ALL SUBJECTS

SECTION 1: CONTENT / CONTEXT	
1.1	The textbook covers all the Learning Outcomes and the Assessment Standards of the subject.
1.2	The textbook covers the suggested content and this is appropriately sequenced.
1.3	The content is suitably paced and the weighting of LOs is appropriate.
1.4	The content is current and up-to-date.
1.5	The content places learning in context i.e. integrates Assessment Standards within the subject to give learners an authentic learning experience.
1.6	There is clear integration of theory and applied competence.
1.7	The content is sensitive to diversity e.g. culture, religion, gender, etc.
1.8	The textbook provides a variety of meaningful activities for individuals, pairs and groups.
1.9	The level of the content is appropriate for the specific grade.
1.10	The language used and vocabulary are appropriate for the grade and language level.
1.11	Key concepts and terms are clearly defined.
1.12	The language and vocabulary are correct and appropriate for the subject.

SECTION 2: LEARNING ACTIVITIES & ASSESSMENT	
2.1	Learning activities and assessment tasks are derived from LOs and ASs.
2.2	The textbook presents the learner with assessment activities appropriate to the subject.
2.3	Assessment tasks are aligned to the Programme of Assessment as described in the Subject Assessment Guidelines.
2.4	A variety of learning activities and assessment tasks are used.
2.5	Assessment targets learner achievement at different levels of complexity.
2.6	Assessment tasks are clearly formulated and unambiguous.
2.7	Assessment tasks and learning activities provide for daily assessment.
2.8	Assessment tasks allow for expanded opportunities for learners.
2.9	Assessment activities reflect the integration of Assessment Standards.

SECTION 3: LAYOUT, DESIGN AND OVERALL QUALITY	
3.1	The text is structured, using headings and subheadings.
3.2	The font and typeface are clear and easy to read.
3.3	The illustrations and diagrams are clear and relevant, without bias.
3.4	The paper is of a good quality and bound securely[1].
3.5	The textbook has an index with clear reference to chapters and page numbers.

[1] If a draft copy/manuscript is submitted a clear indication of these must be given

1

SECTION 4: TEACHER GUIDE	
4.1	Provides clear and systematic guidance on the use of the textbook.
4.2	Provides examples of a subject framework and a work schedule.
4.3	Includes an exemplar assessment plan for the grade.
4.4	Provides memoranda, check lists, rubrics, etc. that match the assessment tasks in the textbook.
4.5	Provides suggested answers / solutions / memoranda / rubrics for learning activities / exercises.

PHYSICAL SCIENCES	
5.1	The textbook covers basic concepts extensively and accurately.
5.2	The textbook is consistent in the use of SI units and symbols.
5.3	The assessment tasks and learning activities promote the use and assessment of applied competence in Physical Sciences.
5.4	The textbook provides sufficient guidance on the 'hypothesis testing' approach to assessment in Physical Sciences.
5.5	The textbook provides learners with a framework to conduct research projects.
5.6	The content includes the latest trends in scientific inquiry and research.
5.7	The textbook is planned in such a way that learners are able to engage in practical activities.
5.8	Illustrations and diagrams have detailed captions and information.
5.9	The volume of content suggested is appropriate for the 4 hours per week allocated to the subject.

B.4 The evaluation instrument of the Western Cape Department of Education

(Western Cape Department of Education, 2005)

ADDENDUM B

WCED DIRECTORATE: CURRICULUM DEVELOPMENT
GR 10 TEXTBOOK CRITERIA FOR SELECTION

ADHERENCE TO NCS POLICY : Does the textbook comply with curriculum policy and guidelines?

The textbook:

1. is compatible with the NCS
2. covers the Grade 10 curriculum for the subject
3. supports attainment of the Critical Outcomes
4. encourages active learning
5. encourages critical thinking
6. presents teachers and learners with suitable assessment tasks

GOOD TEACHING MATERIAL : Does the textbook follow the conventions of good writing and good teaching material?

The textbook:

1. is interactive and interesting to read
2. has up to date content
3. organizes and structures the materials in a logical and coherent way
4. has accurate content
5. is written in a user-friendly way
6. has an attractive design, layout and graphics
7. has a concept level appropriate for the target group (the learners will understand the concepts that are assumed in the learning materials)
8. is sturdy and reusable
9. is not in the form of a workbook
10. has a vocabulary level appropriate for the target groups
11. has an available teaching guide that fully supports the use of the text at different levels
12. has some activities that encourage the use of visual and auditory media such as audio/video cassettes, dvd, computer-based programmes.
13. does not significantly exceed the WCED's target budget of R100.

TRANSFORMATIONAL ISSUES : Do the learning-support materials address issues of transformation?

The textbook:

1. is sensitive to the cultural groups in the Grade 10 classes, as well as to all cultural groups in the country
2. avoids racist innuendo, sexist stereotyping, and textual and visual material of an offensive nature
3. actively incorporates textual and visual materials representing all cultures of the WCED Gr 10 classes and of the country
4. acknowledges the prior experience of learners

1

7. gives a range of learning experiences and levels of achievement through the grade and beyond the indicated assessment standards towards Grade 9 as a bridging mechanism. presents learners with feedback on learning opportunities, activities and the various assessments

PHYSICAL SCIENCE:

A: CONTENT:

1. balances weighting of the three Knowledge strands i.e Earth& Beyond, Matter & Material, Energy & Change.

B: INSTRUCTIONAL QUALITIES:

1. identifies a sense of purpose for the learning of Natural Sciences
2. builds on learners' ideas about science
3. engages learners in a scientific approach to inquiry
4. develops learners' scientific ideas
5. promotes learners' thinking about science as a social enterprise
6. provides for assessment of learners' progress in science
7. enhances the scientific learning environment.

B.5 The evaluation instrument of the Mozambique Department of Education

(Mozambique Ministry of Education, 2002)

Section 3 Criteria for the evaluation of textbooks and teacher's guides

1. Curriculum and content
   a. Conformity to syllabus
   b. Coverage and depth of topics
   c. Content relevance and accuracy
   d. Skills
   e. Assessment

2. Methodological approach and language
   a. Teaching approach
   b. Pupil centredness
   c. Appropriateness of activities to learner and environment
   d. Language and access (readability, appropriacy)

3. Values and cross-cutting issues (human rights, values and inclusivity)
   a. Stereotypes
   b. Nation-building values
   c. Respect for environment
   d. Positive attitudes

4. Structure and layout
   a. Organization of content
   b. Design and layout
   c. Artwork and illustrations

5. Physical specifications and material characteristics
   a. Extent
   b. Format
   c. Paper
   d. Type size and font
   e. Binding

B.6 U.S. Department of Education expert panel on Mathematics and Science Education

(Goldsmith et al., 2000)

U.S. Department of Education
Expert Panel on Mathematics and Science Education

Evaluation Criteria

A. Quality of Program

Criterion 1. ***The program's learning goals are challenging, clear, and appropriate for the intended student population.***

Indicator a. The program's learning goals are explicit and clearly stated.

Indicator b. The program's learning goals are consistent with research on teaching and learning or with identified successful practices.

Indicator c. The program's learning goals foster the development of skills, knowledge, and understandings.

Indicator d. The program's learning goals can include important concepts within the subject area.

Indicator e. The program's learning goals can be met with appropriate hard work and persistence.

Criterion 2. ***The program's content is aligned with its learning goals, and is accurate and appropriate for the intended student population.***

Indicator a. The program's content is aligned with its learning goals.

Indicator b. The program's content emphasizes depth of understanding, rather than breadth of coverage.

Indicator c. The program's content reflects the nature of the field and the thinking that mathematicians use.

Indicator d. The program's content makes connections within the subject area and between disciplines.

Indicator e. The program's content is culturally and ethnically sensitive, free of bias, and reflects diverse participation and diverse student interests.

Criterion 3. ***The program's instructional design is appropriate, engaging, and motivating for the intended student population.***

Indicator a. The program's instructional design provides students with a relevant rationale for learning this material.

Indicator b. The program's instructional design attends to students' prior knowledge and commonly held conceptions.

Indicator c. The program's instructional design fosters the use and application of skills, knowledge, and understandings.

Indicator d. The program's instructional design is engaging and promotes learning.

Indicator e. The program's instructional design promotes student collaboration, discourse, and reflection.

Indicator f. The program's instructional design promotes multiple and effective approaches to learning.

Indicator g. The program's instructional design provides for diverse interests.

Criterion 4. ***The program's system of assessment is appropriate and designed to inform student learning and to guide teachers' instructional decision.***

Indicator a. The program's system of assessment is an integral part of instruction.

Indicator b. The program's system of assessment is consistent with the content, goals, and instructional design of the program.

Indicator c. The program's system of assessment encourages multiple approaches and makes use of diverse forms and methods of assessment.

Indicator d. The program's system of assessment probes students' abilities to demonstrate depth, flexibility, and application of learning.

Indicator e. The program's system of assessment provides information on students' progress and learning needs.

Indicator f. The program's system of assessment helps teachers select or modify activities to meet learning needs.

B. Usefulness to Others

Criterion 5. ***The program can be successfully implemented, adopted, or adapted in multiple educational settings.***

Indicator a. The program provides clear instructions and sufficient training materials to ensure use by those not in the original program.

Indicator b. The program is likely to be successfully transferred to other settings.

Indicator c. The program specifies the conditions and resources needed for implementation.

Indicator d. The program's costs (time and money) can be justified by the benefits.

C. Educational Significance

Criterion 6. ***The program's learning goals reflect the vision promoted in national standards in mathematics education.***

Indicator a. The program's learning goals and subject matter content are consistent with national standards.

Indicator b. The program's pedagogy and assessment are aligned with national standards.

Indicator c. The program promotes equity and equal access to knowledge, as reflected in national standards.

Criterion 7. ***The program addresses important individual and societal needs.***

Indicator a. The program is of sufficient scope and importance to make a significant difference in student learning.

Indicator b. The program contributes to increases in teachers' knowledge of effective teaching and learning.

Indicator c. The program:

- is designed to improve learning for a wide spectrum of students *OR*
- serves to meet the special learning needs of under-served students *OR*
- serves to meet the special learning needs of students whose interests and talents go beyond core mathematics education.

D. Evidence of Effectiveness and Success

Criterion 8. The program makes a measurable difference in student learning.

Promising Programs, in addition to satisfying Criteria 1–7, must provide **preliminary evidence** of effectiveness in **one or more sites** for **at least one** of the indicators below:

Indicator a. The program has evidence of gains in student understanding of mathematics.

Indicator b. The program has evidence of gains in inquiry, reasoning, and problem solving skills.

Indicator c. The program has evidence of improvements in course enrollments, graduation rates, and post-secondary school attendance.

Indicator d. The program has evidence of improvements in attitudes toward learning.

Indicator e. The program has evidence of narrowing the gap in achievement or accomplishment between disaggregated groups.

Indicator f. The program has other evidence of effectiveness or success.

Exemplary Programs, in addition to satisfying Criteria 1–7, must provide **convincing** evidence of effectiveness in **multiple sites with multiple populations** regarding **two or more** of the indicators below. The items must include either both indicators from Part I or one indicator from Part I and one indicator from Part II. Providing evidence of two indicators from Part II is not sufficient.

Part I

Indicator a. The program has evidence of gains in student understanding of mathematics.

Indicator b. The program has evidence of gains in inquiry, reasoning, and problem solving skills.

Part II

Indicator c. The program has evidence of improvements in course enrollments, graduation rates, and post-secondary school attendance.

Indicator d. The program has evidence of improvements in attitudes toward learning.

Indicator e. The program has evidence of narrowing the gap in achievement or accomplishment between disaggregated groups.

Indicator f. The program has other evidence of effectiveness or success.

B.7 The AAAS criteria for judging textbooks

The AAAS 3.18 content analysis questions:

- Does the content address the substance of the specific benchmark(s) or only the benchmark's general topic?
- Does the content reflect the level of sophistication of the specific benchmark or are the activities more appropriate for targeting benchmarks at an earlier or later grade?
- Does the content address all parts of the learning goal?

The AAAS criteria for judging instructional effectiveness.

I	**Providing a Sense of Purpose** Conveying unit purpose Conveying lesson/activity purpose Justifying lesson/activity sequence
II	**Taking Account of Student Ideas** Attending to prerequisite knowledge and skills Alerting teachers to commonly held student ideas Assisting teachers in identifying their students' ideas Addressing commonly held ideas
III	**Engaging Students with Relevant Phenomena** Providing variety of phenomena Providing vivid experiences
IV	**Developing and Using Scientific Ideas** Introducing terms meaningfully Presenting ideas effectively Demonstrating use of knowledge Providing practice
V	**Promoting Students' Thinking about Phenomena, Experiences, and Knowledge** Encouraging students to explain their ideas Guiding student interpretation and reasoning Encouraging students to think about what they have learned
VI	**Assessing Progress** Aligning assessment to goals Testing for understanding Using assessment to inform instruction
VII	**Enhancing the Learning Environment** Providing teacher content support Encouraging curiosity and questioning Support all students

B.8 National Science Resources Centre assessment criteria

National Science Resources Centre assessment criteria (National Academy of Sciences, 1997, 73).

Criteria for Judging Pedagogical Appropriateness
Addressing the goals of elementary science teaching and learning 1. Do the materials focus on concrete experiences with science phenomena? 2. Do the materials enable children to investigate important science concepts in depth over an extended period? 3. Do the curriculum materials contribute to the development of scientific reasoning? 4. Do the materials stimulate students' interest and relate science learning to daily life? 5. Do the materials build conceptual understanding over several lessons through a logical sequence of related activities? 6. Does the instructional sequence include opportunities to assess children's prior knowledge and experience? Focusing on inquiry and activity as basis of the learning experience 1. Does the material focus on student inquiry and engage students in the processes of science? 2. Does the material provide opportunities for students to gather and defend their own evidence and express their results in a variety of ways? Using an effective instructional approach 1. Does the material include a balance of student-directed and teacher-facilitated activities as well as discussions? 2. Does the material incorporate effective strategies for the teacher aand/or students to use in assessing student learning? 3. Does the teacher's guide suggest opportunities for integrating science with other ares of the curriculum? 4. Do students have the opportunities to work collaboratively and alone?
Criteria for Judging Science Content
1. Is the science content current and accurately represented? 2. Does the content emphasise scientific inquiry? 3. Is the content of the science programme consistent with the National Science Education Standards? 4. Does the background material for teachers address the science content that is taught, as well as common misconceptions? 5. Is the treatment of content appropriate for the grade level? 6. Is the content free of bias? 7. Is the writing style for students and teachers interesting and engaging, and is scientific language used appropriately? 8. Is science represented as an enterprise connected to society?
Criteria for Judging Presentation and Format
1. Are the print materials for students well-written, developmentally appropriate, and compelling in content? 2. Are the directions for implementing activities clear in both the teacher's guide and the student material? 3. Are the suggestions for instructional delivery in the teacher's guide adequate? 4. Are the materials free of ethnic, cultural, racial, economic, age, and gender bias? 5. Are appropriate strategies provided to meet the special needs of diverse populations? 6. Are lists of materials for each activity provided, as well as a complete set of materials and information about reasonably priced replacement material? 7. Are safety precautions included where needed? 8. Are instructions for using laboratory equipment and materials clear and adequate?

Appendix C

Pairwise comparison questionnaire

Physical quality

Times more important

	5	4	3	2	|	2	3	4	5	
Affordability										Paper quality

Times more important

	5	4	3	2	|	2	3	4	5	
Affordability										Binding

Times more important

	5	4	3	2	|	2	3	4	5	
Paper quality										Binding

Nature of Science portrayal

Times more important

	5	4	3	2	|	2	3	4	5	
Human endeavour										Ongoing process

Learning Outcomes

Times more important

	5	4	3	2	|	2	3	4	5	
All LOs present and integrated										LOs appropriately balanced

Times more important

	5	4	3	2	|	2	3	4	5	
All LOs present and integrated										LOs clearly stated

Times more important

	5	4	3	2	|	2	3	4	5	
LOs appropriately balanced										LOs clearly stated

Core Knowledge

Times more important

All core knowledge addressed   5   4   3   2 | 2   3   4   5   Prior knowledge mentioned

Times more important

All core knowledge addressed   5   4   3   2 | 2   3   4   5   Addresses common misconceptions

Times more important

All core knowledge addressed   5   4   3   2 | 2   3   4   5   Progression (sequencing)

Times more important

All core knowledge addressed   5   4   3   2 | 2   3   4   5   Integration within and with other learning areas

Times more important

All core knowledge addressed   5   4   3   2 | 2   3   4   5   A world beyond

Times more important

All core knowledge addressed   5   4   3   2 | 2   3   4   5   SA context

Times more important

Prior knowledge mentioned   5   4   3   2 | 2   3   4   5   Addresses common misconceptions

Times more important

Prior knowledge mentioned   5   4   3   2 | 2   3   4   5   Progression (sequencing)

Times more important

Prior knowledge mentioned — 5 4 3 2 | 2 3 4 5 — Integration within and with other learning areas

Times more important

Prior knowledge mentioned — 5 4 3 2 | 2 3 4 5 — A world beyond

Times more important

Prior knowledge mentioned — 5 4 3 2 | 2 3 4 5 — SA context

Times more important

Addresses common misconceptions — 5 4 3 2 | 2 3 4 5 — Progression (sequencing)

Times more important

Addresses common misconceptions — 5 4 3 2 | 2 3 4 5 — Integration within and with other learning areas

Times more important

Addresses common misconceptions — 5 4 3 2 | 2 3 4 5 — A world beyond

Times more important

Addresses common misconceptions — 5 4 3 2 | 2 3 4 5 — SA context

Times more important

Progression (sequencing) — 5 4 3 2 | 2 3 4 5 — Integration within and with other learning areas

Times more important

Progression (sequencing) — 5 4 3 2 | 2 3 4 5 — A world beyond

Times more important

Progression (sequencing) — 5 4 3 2 | 2 3 4 5 — SA context

Times more important

Integration within and with other learning areas — 5 4 3 2 | 2 3 4 5 — A world beyond

Times more important

Integration within and with other learning areas — 5 4 3 2 | 2 3 4 5 — SA context

Times more important

A world beyond — 5 4 3 2 | 2 3 4 5 — SA context

Science

Times more important

Facts accurate — 5 4 3 2 | 2 3 4 5 — Facts up-to-date

Times more important

Facts accurate — 5 4 3 2 | 2 3 4 5 — Units and symbols correct

Times more important

Facts accurate — 5 4 3 2 | 2 3 4 5 — Equipment specified and readily available

Times more important

Facts up-to-date — 5 4 3 2 | 2 3 4 5 — Units and symbols correct

Times more important

Facts up-to-date — 5 4 3 2 | 2 3 4 5 — Equipment specified and readily available

Times more important

Units and symbols correct — 5 4 3 2 | 2 3 4 5 — Equipment specified and readily available

---

Activities

Times more important

Aimed at LOs — 5 4 3 2 | 2 3 4 5 — Encourage active participation (engagement)

Times more important

Aimed at LOs — 5 4 3 2 | 2 3 4 5 — Variation

Times more important

Aimed at LOs — 5 4 3 2 | 2 3 4 5 — In various social combinations

Times more important

Aimed at LOs — 5 4 3 2 | 2 3 4 5 — Laboratory work

Times more important

Aimed at LOs — 5 4 3 2 | 2 3 4 5 — Addresses all Bloom levels

Times more important

Aimed at LOs — 5 4 3 2 | 2 3 4 5 — Communication opportunities

					Times more important					
Encourage active participation (engagement)	5	4	3	2	|	2	3	4	5	Variation
Encourage active participation (engagement)	5	4	3	2	|	2	3	4	5	In various social combinations
Encourage active participation (engagement)	5	4	3	2	|	2	3	4	5	Laboratory work
Encourage active participation (engagement)	5	4	3	2	|	2	3	4	5	Addresses all Bloom levels
Encourage active participation (engagement)	5	4	3	2	|	2	3	4	5	Communication opportunities
Variation	5	4	3	2	|	2	3	4	5	In various social combinations
Variation	5	4	3	2	|	2	3	4	5	Laboratory work
Variation	5	4	3	2	|	2	3	4	5	Addresses all Bloom levels
Variation	5	4	3	2	|	2	3	4	5	Communication opportunities

Times more important

	5	4	3	2	|	2	3	4	5	
In various social combinations										Laboratory work

Times more important

	5	4	3	2	|	2	3	4	5	
In various social combinations										Addresses all Bloom levels

Times more important

	5	4	3	2	|	2	3	4	5	
In various social combinations										Communication opportunities

Times more important

	5	4	3	2	|	2	3	4	5	
Laboratory work										Addresses all Bloom levels

Times more important

	5	4	3	2	|	2	3	4	5	
Laboratory work										Communication opportunities

Times more important

	5	4	3	2	|	2	3	4	5	
Addresses all Bloom levels										Communication opportunities

---

Assessment

Times more important

	5	4	3	2	|	2	3	4	5	
Aimed at LOs										Regular formative assessment

Times more important

	5	4	3	2	|	2	3	4	5	
Aimed at LOs										Answers to formative questions

Times more important

Aimed at LOs — 5 4 3 2 | 2 3 4 5 — In different applications

Times more important

Aimed at LOs — 5 4 3 2 | 2 3 4 5 — Integration within and with other learning areas

Times more important

Aimed at LOs — 5 4 3 2 | 2 3 4 5 — Progression in formative assessment

Times more important

Aimed at LOs — 5 4 3 2 | 2 3 4 5 — Addresses all Bloom levels

Times more important

Regular formative assessment — 5 4 3 2 | 2 3 4 5 — Answers to formative questions

Times more important

Regular formative assessment — 5 4 3 2 | 2 3 4 5 — In different applications

Times more important

Regular formative assessment — 5 4 3 2 | 2 3 4 5 — Integration within and with other learning areas

Times more important

Regular formative assessment — 5 4 3 2 | 2 3 4 5 — Progression in formative assessment

Times more important

Regular formative assessment — 5 4 3 2 | 2 3 4 5 — Addresses all Bloom levels

Times more important

Answers to formative questions | 5 4 3 2 | 2 3 4 5 | In different applications

Times more important

Answers to formative questions | 5 4 3 2 | 2 3 4 5 | Integration within and with other learning areas

Times more important

Answers to formative questions | 5 4 3 2 | 2 3 4 5 | Progression in formative assessment

Times more important

Answers to formative questions | 5 4 3 2 | 2 3 4 5 | Addresses all Bloom levels

Times more important

In different applications | 5 4 3 2 | 2 3 4 5 | Integration within and with other learning areas

Times more important

In different applications | 5 4 3 2 | 2 3 4 5 | Progression in formative assessment

Times more important

In different applications | 5 4 3 2 | 2 3 4 5 | Addresses all Bloom levels

Times more important

Integration within and with other learning areas | 5 4 3 2 | 2 3 4 5 | Progression in formative assessment

Times more important

Integration within and with other learning areas | 5 4 3 2 | 2 3 4 5 | Addresses all Bloom levels

	Times more important									
Progression in formative assessment	5	4	3	2	|	2	3	4	5	Addresses all Bloom levels

Explanations and examples

	Times more important									
New concepts intelligible, plausible and fruitful	5	4	3	2	|	2	3	4	5	Examples in different applications

Scaffold meta-cognition

	Times more important									
Promote big ideas	5	4	3	2	|	2	3	4	5	Clear purpose visible throughout
Promote big ideas	5	4	3	2	|	2	3	4	5	Promote forming of connections
Promote big ideas	5	4	3	2	|	2	3	4	5	Summaries
Clear purpose visible throughout	5	4	3	2	|	2	3	4	5	Promote forming of connections
Clear purpose visible throughout	5	4	3	2	|	2	3	4	5	Summaries
Promote forming of connections	5	4	3	2	|	2	3	4	5	Summaries

Motivate learners

	Times more important									
	5	4	3	2	|	2	3	4	5	
Learner-centered examples & activities										Everyday relevance indicated
Learner-centered examples & activities										Appearance
Everyday relevance indicated										Appearance

Layout and design

(User friendly)

	Times more important									
	5	4	3	2	|	2	3	4	5	
Clear and logical structure										Print size and font
Clear and logical structure										Headings and signalling devices
Clear and logical structure										Index and table of content
Print size and font										Headings and signalling devices

	Times more important	
Print size and font	5 4 3 2 | 2 3 4 5	Index and table of content
Headings and signalling devices	5 4 3 2 | 2 3 4 5	Index and table of content

Pictures

	Times more important	
Correct and detailed captions	5 4 3 2 | 2 3 4 5	Relevant to text
Correct and detailed captions	5 4 3 2 | 2 3 4 5	Active functions
Correct and detailed captions	5 4 3 2 | 2 3 4 5	Technical quality
Relevant to text	5 4 3 2 | 2 3 4 5	Active functions
Relevant to text	5 4 3 2 | 2 3 4 5	Technical quality
Active functions	5 4 3 2 | 2 3 4 5	Technical quality

Language

	Times more important	
Appropriate for level	5 4 3 2 | 2 3 4 5	Scientific vocabulary appropriate

Overall

	Times more important									
	5	4	3	2	|	2	3	4	5	
Sensitivity to diversity										Promotion of values
Sensitivity to diversity										Physical quality
Sensitivity to diversity										Nature of Science portrayal
Promotion of values										Physical quality
Promotion of values										Nature of Science portrayal
Physical quality										Nature of Science portrayal

Content

	Times more important									
	5	4	3	2	|	2	3	4	5	
Learning Outcomes										Core Knowledge
Learning Outcomes										Science

Times more important

Core Knowledge 5 4 3 2 | 2 3 4 5 Science

Didactical aspects

Times more important

Activities 5 4 3 2 | 2 3 4 5 Assessment

Times more important

Activities 5 4 3 2 | 2 3 4 5 Explanations and examples

Times more important

Activities 5 4 3 2 | 2 3 4 5 Scaffold meta-cognition

Times more important

Activities 5 4 3 2 | 2 3 4 5 Differentiation

Times more important

Activities 5 4 3 2 | 2 3 4 5 Motivate learners

Times more important

Assessment 5 4 3 2 | 2 3 4 5 Explanations and examples

Times more important

Assessment 5 4 3 2 | 2 3 4 5 Scaffold meta-cognition

Times more important

	5	4	3	2	|	2	3	4	5	
Assessment										Differentiation

Times more important

	5	4	3	2	|	2	3	4	5	
Assessment										Motivate learners

Times more important

	5	4	3	2	|	2	3	4	5	
Explanations and examples										Scaffold meta-cognition

Times more important

	5	4	3	2	|	2	3	4	5	
Explanations and examples										Differentiation

Times more important

	5	4	3	2	|	2	3	4	5	
Explanations and examples										Motivate learners

Times more important

	5	4	3	2	|	2	3	4	5	
Scaffold meta-cognition										Differentiation

Times more important

	5	4	3	2	|	2	3	4	5	
Scaffold meta-cognition										Motivate learners

Times more important

	5	4	3	2	|	2	3	4	5	
Differentiation										Motivate learners

Presentation

	Times more important									
	5	4	3	2	|	2	3	4	5	
Layout and design (User friendly)										Pictures

	Times more important									
	5	4	3	2	|	2	3	4	5	
Layout and design (User friendly)										Language

	Times more important									
	5	4	3	2	|	2	3	4	5	
Pictures										Language

Textbook

	Times more important									
	5	4	3	2	|	2	3	4	5	
Overall										Content

	Times more important									
	5	4	3	2	|	2	3	4	5	
Overall										Didactical aspects

	Times more important									
	5	4	3	2	|	2	3	4	5	
Overall										Presentation

	Times more important									
	5	4	3	2	|	2	3	4	5	
Content										Didactical aspects

	Times more important									
	5	4	3	2	|	2	3	4	5	
Content										Presentation

	Times more important									
	5	4	3	2	|	2	3	4	5	
Didactical aspects										Presentation

Appendix D

Evaluation rubric

	3	2	1	0
Overall				
Sensitivity to diversity				
Sex/gender, cultural groups, religion	Positive representation of sex, gender, cultural groups and religion	Neutral representation of sex, gender, cultural groups and religion	Some stereotypical representations	Offensive text/ pictures
Promotion of values				
Democracy, social structures & justice, environments	Positive representation of democracy, social structures & justice, environmental issues	Limited positive representation of these values	Mixed messages about these values	Antagonistic to these values
Physical quality				
Affordability	Within budget	Up to 10% over budget	Up to 20% over budget	More than 20% over budget
Paper quality	High quality, strong, white	Good quality	Moderate quality	Newsprint
Binding	Should last 5 years	Should last at least 1 year	Flimsy	Already falling apart
Nature of Science portrayal				
Human endeavour	Frequent references to scientists and their work	Some reference to scientists and to their work	Some reference to scientists, no reference to their work	No reference to scientists or their work
Ongoing process	Reference to the development of scientific understanding and current research	Reference to the development of scientific understanding, but no reference to current research	Limited references to the development of scientific understanding	Facts and laws represented as final truths
Content				
Learning Outcomes (LO's)				
All LO's addressed throughout	All 3 LOs present and integrated throughout	LOs present, but concentrated in certain chapters	Not all LOs present and not well integrated	LOs not present
LO's appropriately weighted	LOs weighted appropriately	One LOs receives too much/little attention	Focuses only on one of the three LOs present	LOs weighted totally inappropriately
LO's clearly stated	All LO's clearly stated	Some LOs stated	LOs only implied	LOs formulated incorrectly

Core knowledge				
All core knowledge addressed	All core knowledge included	All topics present, but a few details are not included	All topics present, but many of the details are not present	Some topics not present
Prior knowledge mentioned	Prior knowledge mentioned wherever appropriate	Prior knowledge mentioned in most of the instances where it is appropriate	Prior knowledge mentioned in a few instances where it is appropriate	Prior knowledge never mentioned
Addresses common misconceptions a	Common misconceptions addressed wherever appropriate	Common misconceptions addressed in most of the appropriate instances	Common misconceptions addressed in a few instances where it is appropriate	Common misconceptions never addressed
Logical progression (sequencing)	All concepts logically sequenced	Very few concepts logically not optimally sequenced	A few concepts not logically sequenced	Serious sequencing problems
Integration (within and with other learning areas)	Exemplary integration: indicates many links with previous science topics and other learning areas	Adequate integration	Limited integration	No integration
A world beyond	Introduces many relevant ```new'' contexts or examples	Limited introduction of ``new'' contexts or examples	Only a few instances of the introduction of ``new'' contexts or examples	Only refers to contexts known to most learners
SA context	Content fully localized	Content mostly localized	Content contains only a few SA examples	No examples from SA context included
Science				
Facts accurate	All facts correct	Some minor mistakes	Several minor mistakes	Many mistakes or a few fundamental errors
Facts up-to-date	All facts up-to-date	Some outdated facts	Many outdated facts	Textbook gives an overall outdated impression
Units and symbols correct	All units and symbols present and correct	Some minor mistakes, inconsistencies or omitted units	Several minor mistakes, inconsistencies or omitted units	Many mistakes, inconsistencies or a few fundamental errors
Equipment specified and readily available	All equipment specified and standard issue or readily available	Most equipment specified and standard issue or readily available	Some equipment not specified or difficult to procure	Most equipment not specified or difficult to procure

Didactical aspects				
Activities				
Aimed at LO's	All activities contribute to achieving at least one of the LOs and a good balance of activities between LOs exists	All activities contribute to achieving at least one of the LOs, but the balance between LOs is incorrect	Some activities have no relationship to LOs **or** one of the LOs is not covered by activities at all	Some activities have no relationship to LOs **and** one of the LOs is not covered by activities at all
Encourage active participation (engagement)	All activities encourage active participation	Most activities encourage active participation	A few activities encourage active participation	No encouragement to participate actively
Variation	Exemplary variation in type of activity	Reasonable amount of variation in type of activity	Very limited variation in type of activity	No variation in type of activity
Various social combinations	Exemplary variation in social combinations (individual, pair, group etc.)	Limited variation in social combinations	Only a few variations in social combination	No variation in social combinations
Laboratory work	All laboratory activities are designed to contribute to achieving LOs and mastering core knowledge	Some laboratory activities do not contribute to achieving LOs and mastering core knowledge	Laboratory activities do not contribute to achieving LOs and mastering core knowledge	No laboratory activities included
All Bloom's levels included	All levels included: Knowledge, Understanding, Application, Analysis, Synthesis, Evaluation	Includes most levels	Limited to some levels	Limited to knowledge and understanding
Communication opportunities	Many and various types of communication opportunities, e.g writing, talking or reading answers, reports, presentations	Limited in number or variation	Limited in number and variation	No communication opportunities
Assessment				
Aimed at LOs	All 3 LOs are addressed in the assessment	One of the LOs is not addressed in the assessment	Assessment is limited to a single LO	No assessment opportunities
Regular formative assessment	Frequent formative assessment opportunities are provided	Formative assessment opportunities are provided at the end of chapters	Formative assessment opportunities are provided at the end of the book	No assessment opportunities
Answers to formative questions	All answers are provided	Some answers are provided	No answers are provided	No assessment opportunities

In different applications	Assessment in many different applications	Limited applications in assessment	No applications in assessment	No assessment opportunities
Integration (within and with other learning areas	Links to previous topics and other learning areas	Links to either previous topic or other learning areas	No links to previous topics or other learning areas	No assessment opportunities
Progression in formative assessment	Assessment progresses from lower to higher levels	Assessment shows limited progression in level	Assessment does not show progression in level	No assessment opportunities
All Bloom's levels included	All levels included: Knowledge, Understanding, Application, Analysis, Synthesis, Evaluation	Includes knowledge, understanding and application	Limited to knowledge and understanding	No assessment opportunities
Examples and explanations				
New concepts intelligible, plausible and fruitful	Exemplary introduction of new concepts	Adequate introduction of new concepts	Introduction of some new concepts fails on one or more of the mentioned criteria	Introduction of most new concepts fails on one or more of the criteria
Examples in different applications	Enough examples and examples in many different applications	Adequate number of examples, but limited number of different applications	Limited number of examples	No examples
Scaffolding meta-cognition				
Promote big ideas	The relationships between concepts and big ideas are indicated where relevant	Big ideas mentioned but individual concepts are seldom linked to it	Big ideas mentioned but individual concepts are never linked to it	Big ideas are not mentioned
Clear purpose throughout	Purpose is clearly and adequately indicated throughout	Purpose stated inconsistently or not clearly	Purpose stated inconsistently and not clearly	Purpose is never indicated
Promote forming of connections	Learners are effectively guided to form connections	The attempts are not enough or ineffective	Attempts to guide students are not enough and ineffective	No attempt to guide learners to form connections
Summaries	Frequency and quality of summaries appropriate	Either the frequency or the quality of summaries are inadequate	Both the frequency and the quality of summaries are inadequate summaries	No summaries
Differentiation				
Caters for varied ability	Exemplary support for slow learners and challenges for fast learners	Limited support for slow learners and challenges for fast learners are provided	Either support for slow learners or challenges for fast learners are provided	No attempt to cater for varied abilities

Motivate learners				
Learner-centered examples & activities	All examples & activities are learner-centered	Most examples & activities are learner-centered	Very few examples & activities are learner-centered	Learners will not relate to any of the example/activities
Everyday relevance indicated	Everyday relevance is indicated where appropriate	Either too few or inappropriate indications of everyday relevance	Iindications of everyday relevance are too few and inappropriate	Everyday relevance is never indicated
Appearance	The textbook appearance will appeal to the majority of the learners	The textbook appearance will appeal to some learners	The textbook appearance will definitely not appeal to learners	The textbook appearance will repulse many learners
Presentation				
Layout and design (User friendly)				
Clear and logical structure	The structure is very clear and logical	The structure is reasonable clear and logical	The structure is either unclear or illogical	The structure is both unclear and illogical
Print size and font	Exemplary	Adequate	Some difficulties	Unreadable
Headings and signaling devices	Exemplary	Adequate	Some difficulties	Unreadable
Index and table of content	Both present, correct and user-friendly	Both present,but either the index or table of content is incorrect or not user-friendly	Either the index or table of content is missing	No index or table of content
Pictures				
Correct and detailed captions	All captions correct	Some correct	Many of captions incorrect	No caption
Relevant to text	All pictures are relevant to the text	Some pictures are irrelevant	Many pictures are irrelevant	None of the pictures are relevant to the text
Active function	Most of the pictures have active functions	Some of the pictures have active functions	Few of the pictures have active functions	None of the pictures have active functions. Only used to fill empty space.
Technical quality	Exemplary	Adequate	Some problems	Many problems
Language				
Appropriate for level	Exemplary	Adequate	Some problems	Too difficult/easy throughout
Scientific vocabulary appropriate	Relevant necessary scientific vocabulary explained and used with care	Some scientific vocabulary missing, unnecessary or not explained	Many scientific terms missing, unnecessary or not explained	Scientific vocabulary used incorrectly