CHAPTER TWO

QUALITY MANAGEMENT IN THE DESIGN OF COMMON TASK ASSESSMENT

2.1 INTRODUCTION

The purpose of this chapter is to develop a common understanding of quality within the context of managing CTA and, in addition, to define quality and concepts relating to quality. Furthermore, because assessment is considered to be an important tool to ascertain the well-being and robustness of the education system, this chapter will examine the concepts related to quality and assessment, namely authenticity, validity and reliability of assessment in relation to CTA.

Through assessment, both educators and learners are able to determine whether the learning outcomes have been achieved (Du Toit & Du Toit, 2004:18).

Outcomes-Based assessment is not something that educators should think about at the end of a unit of work or at the end of a lesson, but must be an integral part of all planning, presentation and preparation. According to the Gauteng Institute on Education and Development (2004:241), assessment strategies are used for measuring knowledge, behaviour or performance, values or attitudes. Outcomes-Based assessment in the NCS for Grade R-9 involves on-going collecting of information regarding learners' performance. This information is then checked against SAQA assessment standards and further used to give feedback to stakeholders, including learners' parents (Department of Education, 2002a:27).

Killen (2005:20) defines quality in assessment as those practices that provide reliable information about the standard of learners' learning that has been achieved so that inferences can be made about how well learners understand and can apply the things they have been learning.

According to Thomas (2003:234), theoreticians have struggled to come up with a variety of definitions for the term *quality*, including that of quality being

defined by the degree to which set objectives are achieved, the fitness of purpose, added value and client satisfaction.

Heyns (2001:2) asserts that education and training providers are the basis of an education and training system and in that way they are actually organizations in teaching, learning and assessment and deal directly with learners whom the education system is to serve as clients. Heyns (2001:3) indicates further that it is important for this service-provider to develop quality management systems that are supported in order to operate in the National Qualifications Framework.

A different, but complementary aspect of the framework for assuring quality outcomes in any organization is quality improvement or quality management. Quality improvement is part of the overall management function of the institution. Furthermore, Heyns (2001:3) points out those key elements in *quality management* would include *strategic planning operations* and evaluations. In particular, quality improvement rests on an ethos of continuous improvement in relation to the user or client's service requirements and to the organization's ability to meet these needs.

It is commonly viewed that general productivity in South Africa ranks amongst the lowest in the world. The World Competitiveness Yearbook (Havenga, 2010:79) ranks South Africa's overall comparative competitiveness as 42nd out of 48 industrialized countries. According to this competitive index, the ranking of management is 40 for Total Quality Management and 41 for customer orientation. What is alarming, is the fact that South Africa's education system is ranked 46th, which reveals that the system does not meet the needs of a competitive economy. This position may be linked to the fact that the lack of effectiveness in education could be ascribed to factors such as, among others:

- historical political developments;
- the collapse of a culture of teaching and learning at schools;
- under-qualified educators and educational managers;
- a lack of commitment and low educator morale;

- learner commitment and discipline;
- parental involvement; and
- limited teaching and learning materials, poor infrastructure and high levels of poverty at rural schools (Van der Westhuizen *et al*., 1999:315).

According to the International Organization for Standardization, (2010) quality management can be described as a provision of principles and the methodological frame for operations which co-ordinates activities to manage and control an organization with regard to quality.

Efforts have been made by the education department to engage schools in ceremonies to promote a culture of learning and teaching (COLTS) (Van der Westhuizen *et al*., 1999:315). The present government has taken the lead to ensure effective education by introducing its Five Year Implementation Plan: 2000-2004, of which the entire Programme 2 focuses on the effectiveness of schools (Department of Education, 2000a:14-17).

The Task Team on Education Management Development mentions managers of schools' lack of managerial skills and experience in their report (Department of Education, 2000a:14-17). In practice, it occurs that some educators, such as senior educators, are promoted to the position of principal without having the appropriate managerial and leadership skills (NW Department of Education, 1997:6). At the same time, the report of the Task Team provides guidelines on how managers could be developed to manage change. Hence the challenge in education lies in a development-orientated approach, emphasizing aspects such as leadership, organizational development and total quality management (Department of Education, 1995:10-27).

It is important for school managers to be able to manage the quality of teaching, learning and assessment in their schools, to ascertain that the tasks given to learners are of sound quality, addressing the correct teaching and learning outcomes as stated in the assessment policy. The researcher of this thesis will determine how quality in the designing and implementation of CTA is managed.

The chapter unfolds according to the following structure:

- Conceptualizing and defining quality management (*cf*. 2.2)
- Conceptualizing and defining quality assessment (*cf*. 2.3)

2.2 QUALITY MANAGEMENT CONCEPTUALIZED AND DEFINED

This chapter deals with quality in the management of CTA with regard to assessment and the impact the quality in the management process of assessment is making. The study further explores a total quality management approach towards dealing with assessment, in particular by inter alia considering quality criteria such as validity, reliability authenticity, fairness, transparency, content coverage and content quality.

2.2.1 Background

Various terms are used to describe quality management concepts or phrases such as Total Quality Management (TQM), Quality Management System (QMS), Systems Management, Quality Improvement Programme (QIP), and Continuous Improvement Strategy (CIS) (Meyer, 1998:14-15). The acronym TQM is used as the overriding concept in this research.

According to the Global Report Card (2008:12), the international notion of the document trends that have influenced education development during the last decade indicates that TQM is an important element to be managed in organizations. The search for quality at schools requires an improvement in all aspects of education and consequently strives to achieve, among others, excellence in classroom assessment practices so that recognized and measurable learning outcomes are attained.

In the next few paragraphs the concept TQM will be expounded upon.

2.2.2 The concept *total* in quality management

TQM is a generic philosophy of quality improvement and not a specific management strategy. The TQM philosophy allows for the development of models of quality that serve the specific needs of the organization. TQM should, therefore, not be seen as the only means through which a school can achieve improved quality more especially with regard to assessment of learners (Heyns, 2001:15). Stark (2010:1) defines TQM as an approach to

the art of management. It entails describing the characteristics of firms' culture, attitude and organization that determine how its products and service meet customers' needs. According to Stark (2010:1), culture refers to the firm's operations, its need to do things correctly the first time and to eliminate errors and wastages.

The Assessment Reform Group (2006b:18) points out that the quality assurance of all summative assessments, including any tests that educators give, should be arranged so that decision-making within schools concerning the progress of learners is based on dependable information.

However, there is a concern about the many undefined or ill-defined concepts and practices associated with quality management at schools. The concerns revolve around the fact that philosophical orientation that has power for some might become so open to interpretation by others that its individual concepts become meaningless (Heyns, 2001:15).

TQM recognizes the contribution of every member, and hence every function and level of an organization, to the provision of goods and services to customers: school leadership, school operations, the classroom**,** and even the curriculum. It affects everyone who works at the school as well as all activities undertaken in the name of the school (Wong & Kanji, 1998:634). Moreover, *total* suggests close interactions and give-and-take interrelationships of an organization with both its micro and macro environments. The quest for quality is everybody's concern and can come from any of the parties in the environment: customers, partners, suppliers, stakeholders, and even non-stakeholders (Wong & Kanji, 1998:634).

This study explored whether consultation was done by involving all the education patrons concerned when discussing the environments, cultural backgrounds, barriers and resources that might have an impact on quality in the designing and implementation of CTA.

2.2.3 The concept *quality* in management

Prior to this study, other studies conducted by quality experts have shown that attainment of quality involves a continuous commitment towards excellence which relies on principles such as continuous teamwork and leadership

(Kanold, 2006:17). Knipe and Speck (2002:57) emphasize that sound leadership, in particular, is essential in attaining improvement in the quality of learning.

Quality is viewed as a way to make sure that the quality of service offered at all stages fits or goes beyond that which is defined according to the agreed standard (Umalusi, 2004). Quality assurance mechanisms in this study that focus on assessment are identified and explained as follows: moderation, verification and quality control (Reddy, 2004:32; Sithole, 2009:23).

At the same time, three fundamental definitions of the term quality are frequently accepted within, among others, the education sector (Murgatroyd & Morgan, 1993:19; Quong & Walker, in De Bruyn & Van der Westhuizen, 2007:288-289):

- Quality assurance (definition for conventional standards)
- Contract conformance (definition for particular standards)
- Customer-driven quality (definition for market-driven standards).

According to Smith and Ngoma-Maema (2003:346), **quality assurance** means that educational experts collaborate to design an evaluation tool that identifies the characteristics of effective educators. In Britain, evaluation was undertaken by a team of inspectors, whose expertise was noted to be sufficient for making an appropriate evaluation in line with teaching and learning standards.

The **contract conformance (particular standards) definition** states that some quality standards have been specified during the negotiation of forming a contract. What is distinctive about contract performance, as opposed to quality assurance, is that the quality specifications are made locally by the person offering service supplies and not the person receiving the service. This form of quality can also be regarded as provider-driven quality (De Bruyn & Van der Westhuizen, 2007:289).

Moreover, **contract conformance** refers to the negotiation of standards in the process of entering into a contract. In this case, the service provider – and not the recipient – outlines the quality-specifications. One example of contract

conformance is assignments set by educators, outlining to learners the expectations regarding product (content) and process/deadline (De Bruyn & Van der Westhuizen, 2007:290).

Customer-driven or market-driven quality refers to a notion of quality in which those who are to receive a product or service make sure their expectations for this product or service, or an alternative shorthand definition of quality as fitness purpose, has been adopted by a state-sponsored system of academic quality audit and assessment to allow judgment about moving towards accomplishing an institution's publically stated purpose (Woodhouse, 1999:32).

Houston (2007:9) asserts that the definitions of quality are problematic without some linguistic slippage or manipulation in several directions; hence one has to redefine customer/client, learners and the education process. As Luizzi (2000:360) indicates, the business model of the supplier-customer relationship is fundamental to TQM, but fails to capture the nature of specific roles, obligations and responsibilities in this particular case. A customer-driven approach focusing on the position of learners and other partners in education might seem difficult to achieve (Houston & Studman, 2001:475; Meiriovich & Romar, 2006:326).

Within the notion of *quality*, it is assumed that most organizations produce a product or service that is intended to satisfy the needs or requirements of users or customers. Therefore, it implies the total package of all the features of quality management and characteristics of a product or service geared towards satisfying stated or implied needs that require quality. Or it implies a philosophy and a methodology, which assists institutions to manage change and to set their own agendas for dealing with the changes of new external pressures.

Quality can, therefore, be described as *fitness for purpose*, where purpose is related to customer needs and where customers ultimately determine the level of satisfaction with the relevant product or service. This includes evaluating the extent to which the institution does what it says it is doing (Thomas, 2003:239; De Bruyn & Van der Westhuizen, 2007:290). Campbell and

Rozsnyai (2002:60) define *fitness for purpose* as one of the possible set standards for determining whether or not a unit meets quality, measured against what is seen to be the goal of the unit.

There have been difficulties in arriving at clear definitions of quality in the educational sphere. The debate continues between those who identify quality in education with *outstanding* or exceptional performance measured against some implicit *gold standard* (learner success, teaching) and those who accept a *fitness for purpose* definition whereby learners, for example, have a say in defining both *fitness* and *purpose*.

According to Vlalsceanu *et al*. (2004:47), quality fitness for purpose is about conformity to sectoral standards. Woodhouse (1999:29-30) indicates that fitness for purpose is a definition that allows institutions to define their mission and objectives, so *quality* is demonstrated by achieving these. The definition allows variability in institutions, rather than forcing them to be clones of one another. These discussions have opened the door to asking further questions related to fitness of purpose in education and from this point to engaging in discussions about the relationship between quality and educational standards (De Bruyn & Van der Westhuizen, 2007:290).

As pointed out by De Bruyn and Van der Westhuizen (2007:291), the systematic focus on quality is beginning to revolutionize the work of organizations. Such a focus is imperative for organizations to survive in an increasingly global market place. The basis of this focus on quality is a move to balance quality assurance with contract conformance and customer-driven quality, the new revolution places emphasis on customer-driven quality supported by contract conformance and quality assurance (De Bruyn & Van der Westhuizen, 2007:291).

Organizations therefore have to recognize that consumer stakeholders are becoming increasingly sophisticated and demanding about the products and services provided by the organization (De Bruyn & Van der Westhuizen, 2007:291). This occurs at the same time as governments are moving to an increasingly market-driven basis for the economy and for public and social services.

The fusion of these two forces causes stakeholders to expect more say in the activities of the organization, giving more emphasis to customer-driven quality than has been the case in the past (De Bruyn & Van der Westhuizen, 2007:290). To meet minimum expectations, organizations are increasingly required to meet quality assurance standards and to add value to these through contract conformance developed at a local level. Meeting such quality assurance standards and adding value to them changes the emphasis in thinking about quality: the emphasis then turns away from quality being established within the professional body or expert or knowledgeable people in that field towards balancing the three kinds of quality, so as to meet the expectations and requirements of stakeholders better. It is a major change in thinking, which requires major changes in the culture of organizations, in particular those managed by professionals (De Bruyn & Van der Westhuizen, 2007: 290-291).

In this study, the researcher (1) determined whether CTA met criteria of *fitness for purpose*; (2) established whether learners were given criteria beforehand when CTA is conducted; (3) determined whether the CTA administered was the appropriate instrument to assess Grade 9 EMS learners; and (4) determined how the quality in the design of the CTA was being managed at the participating schools.

In this study, focus was among other placed on moderation processes and procedures that are used at schools, clusters (districts), national/provincial level and verification and quality control. In the following section the researcher will elaborate on the three quality assurance mechanisms, namely moderation, verification and quality control.

2.2.3.1 Moderation

According to SAQA (2001:3), *moderation* is a term used to describe approaches for arriving at a shared understanding of standards and expectations for the broad general education. It involves educators and professionals as appropriate, working together, drawing on guidance and exemplification, and building on existing standards and expectations to plan learning, teaching and assessments.

The Department of Education (2003a:5) views moderation as the process of authenticating or making sure that the results of school-based and external assessment are correct or a true reflection.

SAQA (2001:10) regards moderation as an essential process that might guarantee quality standards. Quality standards comprise of learning activities in the classroom and assessments for inputs (which are teaching and learning programmes) and the processes that are outputs (referred to as assessments and reports) – which are upheld (Ramotlhale, 2008:15).

According to the South African Qualification Authority (SAQA, 2001:12) and Ramotlhale (2008:15), moderation is not only linked with outputs which are outcomes of teaching and learning during assessment of learning, but moderation is also supposed to be conducted continuously and not as the last part of the recurring nature of quality. In this study, moderation was regarded as the process that ensured that there are quality standards for the inputs-process as well as the outputs-process. As Ramotlhale (2008:15) indicates, this process ensures that moderation takes place from the beginning of the process of teaching and learning.

The Department of Education (2004b:5) indicates moderation as the process of validating the outcome of school-based and external assessment. Gawe and Heyns (2004:162) and Ramotlhale (2008:23) indicate that organizations must visibly show their processes in internal moderation, and policies and procedures must be accessible and give significant feedback to learners and other professional or education bodies concerned. In the context of this study, as reflected in 2.2.4.6, school-based assessment refers to Continuous Assessment (CASS; assessment as a continuous process) and external assessment refers to CTA.

Gawe and Heyns (2004:172) and Ramotlhale (2008:23) outline the purpose of moderation as follows:

- To set up committees to regulate assessment and monitor the reliability of assessment outcomes.
- To verify the design of assessment, materials for appropriateness for the rationale of qualification, and specified learning outcome.

- To monitor the assessment process for quality and justice.
- To evaluate the assessor performance, and offer support, assistance and recommendation to improve competence and to advance assessor performance.

Ramotlhale (2008:23) is of the opinion that moderation must focus on aspects that improve teaching and learning practice, embrace the significance and reflect current changes in the curriculum. Badasie (2005:14) and Govender (2005:37) contribute to the discussion by outlining factors that seem to impact badly on the execution of external moderation that was measured and used throughout the moderation process, reporting that it was untrustworthy, as educators' marks remain unchanged by the moderation process (Govender, 2005:37). On the whole, feedback on the organization and outlines of the portfolios were received (Brombacher & Associates, 2003:12; Department of Education, 2004a:16; Govender, 2005:38).

Badasie (2005:18) indicates that not carrying out peer moderation in an acceptable manner may be due to a lack of educator proficiency. As a result of deficiency in the measures to carry out moderation, Badasie (2005:18) and Govender (2005:38) indicate that there must be competence development programmes on quality assurance in organizations, such as hands-on workshops prepared by education sectors to provide practitioners with an understanding of quality assurance and the ability to execute quality assurance in their organizations, as well as to offer enough resources which are vital for quality of education (Reddy, 2005:17). Thus, it is preferable that subject experts conduct the moderation, as they are experienced and competent (Sigh, 2004:15).

Ramotlhale (2008:24) and Luckett and Sutherland (2000:103) distinguish between two types of moderation processes, in terms of whether they are part of the quality promotion processes which are formative and aim to advance quality or whether they are part of quality promotion control mechanisms which are summative decisions about quality.

According to Ramotlhale (2008:34), the school and the district office are accountable in affecting the assessment system, in other words: the input and

output processes. District offices are responsible for ensuring that schools demarcated to them have adequate staff and provide support through dissemination of policies. In turn, schools should ensure that educators possess the necessary qualifications to deliver quality teaching and learning, as well as the moderation of EMS CTA tasks.

Input factors to enhance the quality of moderation/assessment

Ramotlhale (2008:36) describes *inputs* as the resources available to the system, for example buildings, books, number and quality of teachers and educationally relevant background characteristics of learners. The aspects to be studied under inputs for this study are (1) educators' qualifications and skills in relation to EMS at Grade 9 level; (2) support from the departmental head and the learning facilitator; (3) resources; and (4) staff development.

It is vital to explain briefly the concepts under input factors to provide an understanding of the context under which they will be used in this study. In the following discussion an overview and description of the input aspects will be highlighted.

- **Educator qualifications**

A highly qualified educator is someone who possesses a bachelor's degree, an accepted or full teaching certificate or licence, and who is proficient in every educational subject he/she teaches (Ingersoll, 2005:28; Glatthorn *et al*., 2006:45). Other ways of assuring fine performance by educators in teaching particular subjects are possession of an undergraduate or graduate major, or an advanced certificate in the subject (Ingersoll, 2005:35). Furthermore, Glatthorn *et al*. (2006:19) indicate that highly qualified educators must exhibit proficiency in three brief areas: quality learning (content and academic understanding of the discipline), the science of teaching (which entails the crucial abilities and subject expertise) and educator professionalism.

Additionally, Ramotlhale (2008:36) emphasizes the significance of content knowledge for creating understanding of the subject matter to learners, and pedagogical knowledge as the skill of making the subject comprehensible to learners. A highly qualified educator who has fundamental abilities, such as subject knowledge, content knowledge and expertise in the subject, will

exhibit a high degree of preparation, teaching and learning strategies, and assessment feedback. Subject expertise will be revealed by successful teaching abilities, such as subject matter delivery to make learners understand what is being taught and giving learners proper feedback on teaching and learning (Ramotlhale, 2008:36).

According to the researcher of this thesis, it will be easier for educators to moderate the work learners and their fellow colleagues or cluster educators have marked if they are experts in their fields. For example, an EMS educator needs to be qualified in: accounting, economic and business economics in order to be able to mark and moderate learners' tasks fairly and be consistent in their practices.

- **Support from the Head of Department (HOD)**

It is imperative that the HODs at schools should be able to offer support as well as advice and supervision to EMS educators, and in interpreting policies and explaining how moderation should be carried out.

Support is seen as something coming from those who offer advice, supervision and assistance (Ramotlhale, 2008:39). Therefore, coaching and counselling roles consist of the supervisor's offering information, views and ideas, supported by expert knowledge and ability. This entails the supervisor operating with professional practice and conduct (Ramotlhale, 2008:39). In this case**,** the HOD for EMS and the learning facilitator may take the role of supervisor, by providing educators with information through the dissemination of assessment policies (explaining and giving information outlined in assessment policies), the NCS policy and exemplar materials on how to develop and moderate high quality tasks.

The professional knowledge and skills of the HODs with regard to curriculum implementation and sound assessment practices can be demonstrated by helping educators to interpret and implement policies and circulars relating to assessment and curriculum implementation.

According to the researcher, in the context of this study, the SMTs need to play an important role in making sure that they provide support to educators.

The quality of the support given by SMTs to educators could enhance the quality of learning and assessment activities in the classroom.

The HOD must check the assessment tools for content validity and mark allocation per activity during the implementation of CTA and establish whether the mark allocation per activity is appropriate for the activity given (Ramotlhale, 2008:40). The HOD needs to ensure that learners are assessed fairly by the educator (Ramotlhale, 2008:40). The HOD must also monitor the implementation of both sections of CTA (Ramotlhale, 2008:40-41). Moderating assessments will ensure enhancement of the validity of CTA and CASS marks (Ramotlhale, 2008:41).

- **Resources**

In this study, the term *resources* refers to the EMS NCS Learning Area Policy Grade R-9, educator assessment plans and the National Protocol on Recording and Reporting (NPRR), the National Curriculum Statement Assessment Guidelines for the General Education and Training Phase and learning programme. The EMS NCS Policy is useful to educators as it contains all the Learning Outcomes (L0s) and Assessment Standards (ASs) that must be addressed at Grade 9-level. The NCS Assessment Guidelines for the EMS in the GET phase contain useful guidelines on how to develop learning programmes within a learning framework, using the work-schedule and lesson plans for EMS (Department of Education, 2007b:10).

The quality standards for the input that resources could have in the moderation process are quality control and monitoring. Quality control is seen as the process whereby products are tested and discarded if they fall lower than standards (Ramotlhale, 2008:40). Therefore it is the duty of the Learning Area district facilitator for EMS to ensure that the quality of the learning programmes developed by educators is controlled through the process of moderation to validate its contents. The facilitator must ensure that the content of the learning programmes and assessment plans are relevant, and that they address the relevant Learning Outcomes and Assessment Standards.

Furthermore, the district official, which in the context of this study is the EMS facilitator, must ascertain that the structure of the learning programmes is ready for classroom use by providing feedback to the schools. Quality assurance will focus on controlling the whole process which led to the development of the learning programme. It was intended to give assurance as far as the learning programmes were concerned and to make it easy for them not to be rejected at the quality control stage meaning that moderation and standardization would have been specified regarding the learning programmes. For instance, if the learning programmes do not meet the specified standards, they should be rejected (Ramotlhale, 2008:40).

- **Staff development**

This section on staff development is important to highlight as it is vital for the Department of Education to take staff development as a priority. In the opinion of the researcher, if the training programmes offered by the department can bridge the gap that exists between pre-service and in-service educator training, educators will automatically be able to use assessment methodologies and moderation tools properly. Then learner achievement could be improved. More literature on the importance of staff development is discussed later (*cf*. 2.2.4).

According to Ramotlhale (2008:49), *output* refers to all efforts schools undertake to accomplish, consisting of cognitive attainment of learners and efficient characteristics, such as positive and negative feelings that the learners acquire pertaining to their behaviour.

While the focus of this study was not as such on all the output factors, the researcher argues that authenticity, reliability and validity could also be regarded as output factors to improve learner performance. A summary of valid, reliable and authentic tasks in this study is discussed later (*cf.* 2.3.1' 2.3.2; 2.3.3).

The researcher is also of the opinion that there is a positive relationship between the input and output process in the context of this study. The following input factors, namely educator qualification; the development of high quality learning programmes; and staff development positively influence the

teaching and learning and assessment activities in the classroom. Highly qualified educators with adequate content knowledge of EMS will be able to develop teaching and assessment tasks of required standards, which could translate into improving the performance of learners. Educators will then not only possess the relevant qualifications, but also have the ability to impart knowledge and skills to their learners. Educators will understand that the impact they make in their classrooms also affects the community at large. There will be quality assurance standards from the input-process-output factors, which guarantee that the results attained on the performance of learners will be authentic, valid and reliable. Thus, the qualification of learners in EMS will be credible and authentic.

The next sections present verification and quality control as the remaining quality control mechanisms.

2.2.3.2 Verification

The Scottish Qualification Authority (2010:3) defines *verification* as a range of quality measures used by the Scottish Qualification Authority to confirm that assessment tasks and activities provide learners with more and valid opportunities to meet the standards. It is the term used to describe the approaches to ascertain that schools' assessment decisions are valid and reliable and in line with national standards.

In the South African context, according to Umalusi (2006:12), verification as a way of ensuring the accuracy or appropriateness of what has been achieved is employed in terms of information on effectively fulfilled measures. This becomes important in this study which seeks to determine whether the Grade 9 CTA assessments are valid and reliable.

2.2.3.3 Quality control

According to SAQA (2001:32), *quality control* is defined as an inspection of the product or service in order to make judgements regarding whether or not these will satisfy the customer's needs. Gawe and Heyns (2004:160) argue that quality is a different approach for the new education and training systems, and one that necessitates continuous regular monitoring and feedback.

Quality is only authentic when it deals with the advancement of classroom practice.

The discussion and definitions of the concepts moderation, verification and quality control are relevant and appropriate in the context of this study, as the study seeks to address the important aspects linked to quality of assessment. The study inter alia seeks to establish what quality assurance mechanisms in CASS and CTA at Grade 9 are utilized within the EMS Learning Area.

Assessment forms a fundamental part of educator-instruction and development with a view to improving the quality of teaching and learning. Each school has to develop sound assessment practices which will improve not only learners' learning, but also the quality of learning programmes (Marais *et al*., 2008:152). Improved assessment practices refer to performances as they progress towards achieving the desired learning outcomes. It is in this regard that principals need to create opportunities to improve educators' assessment competencies (Csizmadia, 2006:66).

De Bruyn and Van der Westhuizen (2007:290) assert that quality will not be achieved by accident or by management dictates; it requires a cultural change that will transform management behaviour and attitudes in general. This process of change must be upheld by managers who are fully committed to the task. There are, of course, many approaches that produce quality results. It is noted, however, that the TQM approach has the additional advantage of facilitating practices that promote both quality and sound management processes. Yet this approach is not easy to implement and to maintain, with some critics arguing that the failure rate of implementing quality practices at schools could be as high as 70% (De Bruyn & Van der Westhuizen, 2007:301).

In everyday language, quality refers to an acceptable standard of satisfaction for a given product, such as a car or, as an example in education, a process (Thomas, 2003:234). Campbell and Rozsnyai (2002:20-21) and Harvey (2004:1) discuss quality as transformation of education. The focus is on the ability of an institution to empower learners with the skills, values, knowledge and attitudes requisite for functioning in the knowledge society. Harvey's

definition is particularly focused on situations of socio-political recharge, and is therefore pertinent to South Africa, where large numbers of previously disadvantaged learners are gaining access to Higher Education.

It is argued that measuring a transformational quality approach involves the following four key elements (Harvey, 2004:1):

- It should be a process that aims to improve the learners' experience.
- It should employ a bottom-up strategy gaining buy-in from learners in on-going improvement.
- It should emphasize effective action.
- It should give emphasis to external monitoring.

The shortcoming of this approach, according to Lomas (2007:72), is that transformation as the acquisition of intellectual capital is difficult to assess.

2.2.4 Achieving quality in management

Quality in management can be achieved in the following ways:

- **Familiarizing stakeholders with the process of quality management**

According to Heyns (2002:6), *quality management system* is referred to as the combination of processes used to ensure that the degree of excellence specified is attained. This is a quality management system which sums up the activities and information an organization uses to enable it to deliver services better and more consistently to meet and go beyond the needs and expectations of its customers and beneficiaries, most cost effectively and cost efficiently.

Schools are already undertaking a way that reflects the quality management philosophy. These include, among others, the use of curriculum teams, the relatively high level of accountability which educators have for educational decision-making in their classrooms and the use of the school-based planning process. The Assessment Reform Group (2006a:13) asserts that the emphasis, however, cannot be attributed to TQM *per se*, as many schools have developed their own particular organizational culture without applying TQM.

It is the direct responsibility of School Management Teams to ensure that parents understand how assessment is assisting learning and how the set standards are used in reporting progress at given times during the year (Assessment Reform Group, 2006a:13).

- **Focus on continuous improvement**

According to Heyns (2002:6), *quality audits* indicate the activities undertaken to measure the quality of products or services that have already been made or delivered.

Quality management focus on continuous improvement of work processes may put the high regard for people and their achievements, which is associated with the TQM, into perspective. According to De Bruyn and Van der Westhuizen (2007:311), people feel better about themselves as work processes are improved continuously. Relationships among people in the organization are more open and honest, and school managers often feel less isolated, misunderstood and burdened. With organizational change come opportunities for personal and professional growth, along with pride and joy in their work.

Stark (2010:2) asserts that continuous improvement of all operations and activities is at the heart of TQM. Once it is recognized that customer satisfaction can only be obtained by providing a high quality product, continuous improvement of quality of the product is seen as the only way to maintain a high level of customer satisfaction. There must be a way of recognizing the link between product quality and customer satisfaction, with TQM also recognizing that the product quality is the result of a *process* of quality. As a result, there must be a focus on the continuous improvement of a company's processes and on product quality, aiming at increased customer satisfaction (Stark, 2010:2).

- **Staff development and training**

As indicated by Heyns (2002:6), *quality control* is referred to as a process undertaken by the person/s making the product – or delivering the service – for internal purposes.

Rebore (2001:180) asserts that School Development Teams must provide educators with opportunities to update their skills and knowledge in a subject area while keeping abreast of societal demands. Therefore the School Development Teams and other members of staff should reach an agreement on areas that need the attention of staff development. The performance agreement contract and agreement indicating the area where staff still need development and support should be in place and signed by the employer and the staff member. Furthermore, quality management requires education and training of all personnel. Everyone in the organization is involved in *quality* education to equip him/her to apply the quality principles in his/her own work situations (Rebore, 2001:180).

This means that everyone learns to speak a common language of quality improvement and this makes it possible to create an organizational culture to support the process (Venter, 2003).

According to the researcher of this thesis, educators should be given proper development and training to administer CTA and understand the quality of skills expected from these tasks.

The Assessment Reform Group (2002:9) points out that it is important that professional development should involve the following:

- ✓ Extending awareness of both limited validity of tests and other assessments of learning and of the ways in which evidence from these tests can be used to guide learning.
- ✓ Recognizing how preparation for, involvement in and responding to tests and assessments of learning can impact negatively on learners' motivation.
- ✓ Devising strategies to minimize the negative impacts of tests and assessment of learning; understanding the differential impact of tests on learners, including, for example, how the negative impact on low achieving learners can be reduced.
- ✓ Discussing and helping the implementation of within-school strategies for emphasizing learning goals as distinct from performance goals.

- ✓ Teaching methods that contribute most to the attainment of these goals will also be a feature of such a discussion.

The continuous improvement strategy intends to set demand and support; offering specialized improvement prospects which must allow educators to be kept updated about their subject knowledge, expertise and teaching skills; raising the supply of teaching resources, but also holding educators accountable for their achievement in supporting and increasing the performance of their learners (McMahon, 2004:131; Ramotlhale, 2008:42).

- **The need for inspection**

According to Heyns (2002:6), *quality assurance* is described as the sum of activities that assure or ascertain the quality of products and services at the time of production or delivery. Quality assurance procedures are frequently applied only to the activities and products associated directly with goods and services provided to external customers.

Schools routinely evaluate their own performance and are subject to periodic inspection by external agencies. Indicators derived by combining the results of individual pupils have a significance role in self-evaluation and inspection. However, they can only be indicative of some aspects of a school's performance. The use of such results for these purposes is likely to affect the way in which tests are seen both by educators and by learners (Assessment Reform Group, 2002).

In this research, the use of criteria for the inspection of quality of assessments at schools was observed and the extent to which assessments contribute to learners' learning as indicator of learning were investigated.

- **Management of the change process**

It appears as if the implementation of CTA requires a transformation process resulting in radical changes for the schools. Renewal depends on at least three possible approaches. Firstly, De Bruyn and Van der Westhuizen (2007:335) suggest that the attitudes of managers and educators need to be changed as a prerequisite for change at the schools. Secondly, according to Van der Westhuizen (2007:335), the most effective way to change behaviour

is to put people into new contexts that impose new roles, responsibilities and relationships on them. Thirdly, it is enough to change employee attitude without rectifying the structure of the organization at the same time (De Bruyn, 2003:96).

There is no use for schools to focus on processes to improve attendance figures and pass rates, yet schools produce learners who are not equipped to take on the demands of the modern community (Molete, 2004:67). According to the researcher of this thesis, the learners should be equipped with skills to be competent in the careers they would follow. It is not sufficient merely to satisfy the internal and external customers of the school. The school should rather identify the real needs of the community as a whole, for example, the quality of life, environment issues, crime and matters related to health and welfare.

The systematic focus on quality is beginning to revolutionize the work of organizations. Such a focus is imperative for organizations to survive in an increasingly global market place. The basis of this focus on quality is a move to balance quality assurance with contract conformance and customer-driven quality. The new revolution places emphasis on customer-driven quality supported by contract conformance and quality assurance (Murgatroyd & Morgan, 1993:51).

Organizations therefore have to recognize that consumer stakeholders are becoming increasingly sophisticated and demanding about the products and systems they have not received in the past, so the Department of Education needs to be aware that the small part of the resources currently used to run external examinations in schools is not enough.

- **Avoiding a top-down approach**

A school-based assessment initiative is doomed to failure if a top-down approach is adopted. To secure educators' support, more *assessment training* and *resource support* for educators are essential. Under a school-based assessment system, educators are under pressure because they wear two hats: as facilitators of learning and as examiners. Where one role ends and the other begins could pose considerable problems, particularly for new

educators. The next difficulty is to ensure credibility for school-based assessment. The authority needs an effective and efficient quality assurance and quality control system to assure the users of examination results, such as employers and tertiary institutions, as well as the general public, of the reliability of this scheme of assessment. This is not a simple task (Choi, 1999:415).

In this study, based on the discussion above, the empirical investigation also focused on determining whether there is credibility in the management of school-based assessment and whether quality assurance – which involves contract conformance and customer-driven quality (fitness to use in the market: for the context of this study the market is learners at schools to be assessed by administering CTA) – and quality control (which involves moderation and verification) are in place and properly managed at schools with regard to CTA.

2.3 QUALITY ASSESSMENT CONCEPTUALIZED AND DEFINED

Striving for quality underpins assessment. The setting of minimum criteria for the achievement of outcomes determines certain *standards* against which learners can demonstrate mastery of an outcome. Continuous, coherent and progressive assessment is seen as *one of the key elements in the quality assurance system* (Department of Education, 2002b:4). This section explores a number of key features of quality in assessment. These key features are well aligned with the policy that guides the implementation of assessment (*cf*. 3.2.1). One of the key features of quality in assessment is validity.

2.3.1 Validity in assessment

According to Du Toit and Vandeyar (2004:133), assessment is valid when it assesses the learning outcomes which it is supposed to assess.

Validity is one of the most important aspects of sound assessment practices. It is important that educators understand the concept and know how to use it as a quality control measure. Moreover, it is vital that validity assumes pride of place as the most fundamental consideration in developing evaluation tests (Stobart, 2008:12). Furthermore, according to the Department of Education (2002a:3), assessment in OBE gathers valid and reliable information about

the performance of the learner on an on-going basis (CASS) against clearly defined criteria using a variety of methods, tools, techniques and contexts, recording the findings, reflecting and reporting on them by giving positive, supportive and motivational feedback to learners, educators, parents and other stakeholders. Effective application of these concepts depends upon a deep understanding of their meaning and implications.

Validity refers to the integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores or to the modes of assessments (Killen, 2003:1; Reddy, 2004:35; Ebbutt, 2006:4). Airansian (2001:423) and Vandeyar and Killen (2006:386) express validity as the degree to which assessment information permits correct interpretations of the desired kind. The difference is not just one of detail; it is a significant change in emphasis, from validity being a property of a test item or assessment task, to validity being a value judgment about inferences and actions made as a result of assessment. According to Mertens (2010:383-384), validity is concerned with the appropriateness, usefulness and meaningfulness of inferences made from the assessment result. Validity therefore refers to measuring what is supposed to be measured, be it knowledge, understanding, subject content, skills, information or behaviours (SAQA, 2001:17; Jacobs *et al*., 2004:283; Reddy, 2004:35; Falchikov, 2005:29).

- **Evidence of validity**

Several books on assessment define validity, rather loosely, as the extent to which *a test measures what it is meant to measure* (Jacobs *et al*., 2004:284; Reddy, 2004:35; Falchikov, 2005:30-31). According to Brady and Kennedy (2001:55), it is when the assessment tasks *measure what educators want them to measure* that they are regarded as *valid tasks*. There is some appeal in the simplicity of this definition, because it can serve as a useful starting point for discussion on test items or assessment tasks, particularly in OBE where the educator can ask, *Is this item testing the outcome I want to test?* (Brady & Kennedy, 2001:55).

This simple view of validity as an inherent property of an item or test can be misleading and counterproductive. The reasons for this claim will be explored later in this chapter, but first some of the historical developments in the concept of validity will be briefly reviewed.

There have been a number of significant stages in the evolution of the concept of validity. However, it seems that the ideas emerging in each new stage have not always resulted in the majority of the educators changing their assessment practices. In fact; there is considerable evidence that many current assessment practices are still guided by vague conceptions of validity that are based on measurement theories (Stiggins, 2002:758).

Validity, however, is not a simple concept and various forms of it are identified according to the basis of the judgment of validity. Validity comprises six criteria, namely (1) content, (2) construct, (3) concurrent, (4) face, (5) criterion-related, and (6) consequential validity. Each of these criteria will be discussed briefly.

Evidence relating to the *content validity* of an assessment would result from comparing the content assessed with the content of a curriculum it was intended to assess. Content validity is an indication of how relevant the content of an assessment task is, and how representative it is of the domain that is purported to be tested (Killen, 2003:3; Reddy, 2004:35; Le Grange & Beets, 2005:115; Cohen *et al*., 2007:109).

It is essentially this concept of content validity that leads to claims such as *validity defines whether a test or item measures whatever it has to measure* (Brady & Kennedy, 2001:13; Van der Horst & McDonald, 2001:185).

It is also important to acknowledge the conditions under which the test is administered, the effect that learners' characteristics will have on their responses and the responsibility that educators have to interpret the test results in defensible ways. Decisions about validity should not overlook the fact that determining what a test measures, requires more than considering just content relevance and representativeness (Yung *et al*., 2008:11).

Item relevance and content coverage describe the potential of a test to provide information from which valid inferences can be drawn. If a test cannot

be related to the curriculum content, then it cannot produce useful evidence of learners' learning. If the total assessment task does not test a suitably representative sample of important curriculum content, then no inferences can be drawn about the curriculum content (Le Grange & Beets, 2005:115-116).

However, although item relevance and content coverage are necessary, they are not sufficient to guarantee that valid inferences are drawn. The researcher of this thesis wished to establish whether the CTA items are relevant and representative of the EMS content domain and whether appropriate inferences can be drawn from learners' answers to the CTA questions (Le Grange & Beets, 2005:116).

The second type of evidence of validity is that of *construct validity***:** a judgment of how well the assessment calls upon the knowledge and skills or other constructs that are supposedly assessed. Therefore this researcher wished to determine whether there is clarity on the domain being assessed, and if there was evidence that, in the assessment process, the intended skills and knowledge are used by the learners (Reddy, 2004:35).

Construct validity involves seeking evidence that the assessment task is actually providing a trustworthy measurement of the underlying construct in which the examiner was interested (Reddy, 2004:35).

If this can be established, then construct–related evidence that the inference made can be based on the test, has a possibility of being valid. However, it is still necessary to consider whether or not the inferences actually are valid (Reddy, 2004:35).

Le Grange and Beets (2005:116-117) assert that *concurrent validity* is derived from the correlation of the outcomes of one assessment procedure with another that is assumed to assess the same knowledge or skill. Furthermore, it can be used as a parameter in sociology, psychology and other psychometric or behavioural sciences. Concurrent validity is demonstrated where a test correlates well with a measure that has previously been validated. The two measures may be for the same construct or for different, but presumably related, constructs (Brady & Kennedy, 2001:19; Killen, 2003:2; Le Grange & Beets, 2005:115).

For example, a measure of job satisfaction might be correlated with work performance. Note that with concurrent validity, the two measures are taken at the same time. This is in contrast to predictive validity, where one measure occurs earlier and is meant to predict some later measure (Williams, 2001:3).

Face validity is based on expert judgment of what an assessment appears to assess: whether it assesses what it is supposed to assess (Pietersen & Maree, 2007b:217). Yung *et al*. (2008:11) refer to face validity as a check on face validity, by sending test surveys and memos to moderators to obtain suggestions for modification.

In this research, the researcher determined whether educators are given such a chance to evaluate the CTA and gave educators the opportunity to give their own opinions about the instrument in terms of face validity.

In addition to considering the relevance and content coverage of the assessment items, it is necessary to consider the extent to which the inferences drawn by the educators can be justified. One way to address this shortcoming is also to consider *criterion-related validity*. Historically, the criterion-related validity of a test was determined by comparing the test scores with one or more external variables (called criteria) that were considered to provide a direct measure of the behaviour or characteristics in question (Delport & Roestenburg, 2011:174). The comparisons were usually made by calculating correlations or regressions.

A similar situation exists when we consider predictive *criterion–related validity*. This concept was used historically to describe the correlation between a test score and some criterion measurement made in the future (Cunningham *et al*,, 1994:654). Again, this was a useful concept when the external measure was established as a direct measure of the quality of interest and if the earlier test was standardized and used repeatedly with different groups of learners. However, this is not the situation with most tests like CTA and other educator developed tests (Cunningham *et al*., 1994:643. In lieu of the purpose of this study, the researcher chose not to determine criterion-related validity.

By definition, the predictive validity of a test cannot be determined until the subjects have been tested on the (future) criterion test. For normal testing

purposes in classrooms, this means that the predictive validity of educator-developed tests cannot be determined in a timeframe that makes the exercise worthwhile. Thus, criterion-related validity is concerned with specific test-criterion correlation and is not a particularly useful concept for classroom educators (Li, 2003:90).

Williams (2006a:3) furthermore indicates that in criterion-related validity, a prediction is usually made about how operationalization with performance-based assessment on the theory of construct will be conducted. There are many stations in which educators want to measure one thing and determine whether it is systematically related to something else. For example, educators might be interested to know whether learners' results in an assessment task completed at home are a sound indication of their future examination performance, also termed the criterion measure (Killen, 2003:10).

Therefore, as pointed out by Killen (2003:10), educators are interested in the criterion-related evidence that inferences they make about the relationship between the assessment task and the examination results are valid. Again, it is not a simple process of judging the validity of the assessment task, the criterion measure and the relationship between the two (Killen, 2003:10). Most significantly, it is necessary to examine the evidence that the information on the tasks and their relationship has been used in appropriate ways to draw defensible conclusions relating to learners' learning.

A further form of validity of increasing interest and relevance is *consequential validity*. Reckase (2008:13-15) proposes that what is to be validated is not the test or observation device as such, but the inferences derived from test scores or other indicators – inferences about score meaning or interpretation and about the implications for action that interpretation entails. In other words, the uses of and consequences of the uses of a test determine its validity. If inappropriate use is made of tests which make them unfair in an ethical and social sense, irrespective of their *technical* validity, the tests lack *consequential* validity (Reckase, 2008:16).

In terms of conceptualizing validity of CTA, the following aspects were explored: content; concurrent; construct; face; consequential and criterion-

related validity. The aforementioned aspects were addressed in Section B of the learner questionnaire in B10, B15, B16, B17 and B19 (*cf*. Appendix I).

The next section will highlight reliability as a feature of quality in assessment. This is done in order to determine ways in which management procedures should secure that the CTA instrument complies with reliability criteria.

2.3.2 Reliability in assessment

Reliability implies consistency in terms of how far the same test would give the same results if done by the same learners under the same conditions (Vandeyar & Killen, 2006:389). According to Du Toit and Vandeyar (2004:133), reliability means that the same assessment task, administered at different times by different persons, produces comparable results.

A reliable test thus makes it possible for one to make reliable comparisons. Comparisons may be between the performance (norm-referenced) of learners and the attainment of outcomes (criterion-referenced).

Reliability in a sense implies consistency in assessment. Reddy (2004:36) indicates that problems with reliability and consistency occur for both assessors and learners. Assessors assess the same work differently and even individuals seem to assess the same work differently at times and in different contexts when involved with the same assessment. It therefore appears that reliability in assessment is difficult to achieve and that the ideal of 100% reliability is illusory (Reddy, 2004:36). However, ways in which assessment can be made more reliable are suggested.

These ways to make assessment more reliable include the following aspects (Reddy, 2004:34):

- Creating and communicating clear criteria against which learners' performance is measured. They argue that a few good, explicit criteria that are understood by assessors and learners lead to greater reliability than complicated marking tasks. Creating agreement on sound and usable criteria is thus an important aspect of improving reliability.

- Assessors in the examination setting mark some scripts and then convene to discuss the criteria and often make adjustments to these. This ensures more reliable approaches to assessment by the assessors.
- Another approach to ensure greater reliability in assessment is to have assignments or tests marked by two assessors, so-called double marking. This is, however, a time-consuming practice and is not always possible to carry out.
- Where a wide range of assessment techniques is used, a wide range of evidence is generated about the competence or performance of a candidate.
- Triangulation is a route to greater reliability. In this way, direct evidence gained in different ways from the candidate can be compared, and statements about the candidate from third parties can also be considered. The more these perspectives coincide, the more confident one can be about the judgments that result from the assessment (Reddy, 2004:34).

Assessment is an integral part of the learning process and there are excellent reasons for assessment. While assessment is a wonderful tool for educators to assist and gauge learning, it is not without its limitations. The tools of assessment, according to Reddy (2004:36), are crude and imperfect, and they often deal with factors that are intangible and difficult to measure. Different viewpoints around assessment should be prevalent during teaching and learning instead of singling out a certain preference for a particular type of assessment.

Cunningham *et al*. (1994:644) indicate that discussions of reliability in many textbooks are based on the notion that assessment takes place at a single time and that summary decisions are made about examinees based on single testing events. In the classroom, educators are engaged in on-going assessment over time and across many dimensions of behaviour. While individualization of instruction may result in better achievement and motivation, it means that standardization is difficult.

In a strict sense, reliability refers to the degree to which test scores are free from errors of measurement (Cunningham *et al*., 1994:643). Reliability in assessment is about consistency and the extent to which the same judgments

can be made in similar context in order to analyse the results statistically (SAQA, 2002:18).

The researcher explored the following criteria linked to reliability, namely consistency and double-marking.

2.3.2.1 The relationship between reliability and validity

It is sometimes said that validity is more important than reliability. In this sense there is no point in measuring something reliably unless one knows what one is measuring. After all, that would be like: *I've measured something, and I know I am doing it right, because I get the same reading consistently, although I do not know what I am measuring*. On the other hand, reliability is a pre-requisite for validity. No assessment can have any validity at all if the mark a learner gets varies radically from occasion to occasion, even if it depends on who does the marking (Williams, 2001:9-10). The reliability and validity relationship is one of the focuses for the given amount of testing time; one can get little information across a broad range of topics as occurs in the case of national curriculum tests, although the trade-off here is that the scores for individuals are relatively unreliable (Williams, 2001:13).

Smith and Ngoma-Maema (2003:345-347) suggests that reliability should depend on how well learners do their task, rather than finding out how well the learner has performed in relation to others. Reliability can be checked through the collection of sufficient observation of data over many tasks, and the impact on classroom assessment can be evaluated through a consideration of the intended and unintended consequences of educators' decisions (Smith & Ngoma-Maema, 2003:349). These new ways of looking at validity and reliability demonstrate that new ways of thinking are emerging with regard to formative assessment.

The recognition of the interaction between validity and reliability means that, while it is useful to consider each separately, what matters in practice is the way in which they are combined. This has led to the combination of the two in the concept of dependability, and James (1998:159) expresses this as follows:

Reliability + Validity = Dependability

Hence, for assessments to be reliable, educators must ensure that they have sufficient information from which to make dependable decisions about a learner's performance. Given this framework, evidence for the validity of assessments used to make decisions should be the most important factor for educators to take into consideration (James & Peddler, 2006:125). Reliability of assessment decisions depends on the quality of the assessments.

The researcher is of the opinion that if attention is given to evidence for validity, then educators can begin to ask themselves whether there is sufficient information from which to make dependable decisions.

It is imperative to adhere to the principles of high quality assessment when implementing CTA as an assessment tool (Killen, 2002:5; Vandeyar & Killen, 2003:126). The content of an assessment task must, like any question paper, meet certain requirements. These will substantially enhance the quality of the assessment tasks in CTA.

In order for CTA to be valid and reliable, the following aspects have to be managed well:

- The assessment procedures should focus clearly on the *outcomes* to be tested so that valid inferences can be drawn about learning.
- The assessment procedures should be *reliable*. There should be a conscious effort to minimize measurement errors and allowing learners to demonstrate their understanding should not be influenced by any irrelevant factors such as the learner's cultural background.
- Assessment should reflect the *knowledge and skills* that are most important for learners to acquire (that is, the building blocks for the achievement of long-term outcomes).
- Assessment should *challenge learners to limits* of their understanding ability to apply their knowledge. It will, therefore, discriminate between those who have achieved high standards and those who have not.
- Assessment tasks should be *authentic and meaningful* so that they support every learner's opportunity to learn and, because learners are individuals, assessment should allow these individual efforts to be demonstrated (Killen, 2002:5; Vandeyar & Killen, 2003:126).

In this study, the researcher investigated how well all the above-mentioned aspects were managed. These aspects were addressed in Section C of the learner questionnaire in C32, C33, C34, C36 and C37 (*cf*. Appendix I). In the educator questionnaire these aspects were dealt with in Section B18, B19, B20, B21 and B22 (*cf*. Appendix H).

The following section explores authenticity as a feature of quality assessment.

2.3.3 Authenticity in assessment

Authenticity is an important element of new modes of assessment. (Gulikers *et al*., 2008:74) In this section, the review of literature on the authenticity of assessment will be dealt with, along with the presentation of a five dimensional framework for designing authentic assessment and a discussion of implications of this five dimensional framework (Gulikers *et al*., 2008:73-86).

Traditional frontal classroom instruction for learning facts, assessed through short-answer or multiple-choice tests, is an example of such an alignment of classroom instruction for learning and assessment. Moreover, traditional teaching practices are characterized by knowledge transmission and rote memorization, and assessment is also termed *testing culture* (Birenbaum, 2003:15). Assessment in the traditional teaching approach consists primarily of decontextualized, psychometrically designed items in a choice-response format to test for knowledge and low-level cognitive skill acquisition.

In addition to this, OBE places emphasis that focuses on the acquisition of higher-order thinking processes and competencies instead of mere factual knowledge and basic skill. The function of this change in assessment is more of a summative rather than that of a formative nature and aims even more at promoting and enhancing learners' learning than is the case in traditional teaching. This view requires alternative assessment methods, because standardized, multiple-choice tests are not suitable for this approach (Segers *et al*., 2001:573).

Authentic assessment is characterized by the following (Vandeyar & Killen, 2003:121; Falchikov, 2005:71):

- learners take responsibility for their own learning;

- they reflect and collaborate; and
- conduct a continuous dialogue with the educator.

Authentic assessment, according to Montgomery (cited by Vandeyar & Killen, 2003:121 & Falchikov, 2005:71), requires learners to answer important questions, solve real problems and engage in non-routine and multistage tasks that require high-quality performance. Authentic assessment practices often include investigations conducted collaboratively, hands-on solving of real problems, performances completed over extended periods of time and the presentation of evidence of learning through portfolios or non-written products (Vandeyar & Killen, 2003:132).

Assessment involves interesting real life challenges that require learners to apply their relevant skills and knowledge or authentic tasks and contexts as well as multiple assessments opportunities to reach a profile score determining learners' learning or development (Gulikers *et al*., 2004:73). On the other hand, Muller (1998) asserts that authentic assessment is a form of assessment in which learners are asked to perform real world tasks that demonstrate meaningful application of essential knowledge and skills.

The two most important reasons for using authentic competency-based assessments are (1) their *construct validity* and (2) their impact on learners' learning, also called *consequential validity* (Dierick *et al*., 2001:14; Gielen *et al*., 2001:51; Gulikers *et al*., 2004:74). According to Gielen *et al*. (2001:51) and Guilkers *et al*. (2004:74), construct validity of an assessment is related to whether an assessment measures what it is supposed to measure. With respect to competency assessment this means that (1) tasks must appropriately reflect the competency that needs to be assessed; (2) the content of an assessment involves authentic tasks that represent real-life problems of the knowledge domain assessed; and (3) the thinking process that experts use to solve the problem in real life is also required by the assessment task (Gielen *et al.*, 2001:49).

The definition of authentic assessment as used in this study relates to assessment requiring learners to use the same competencies or combinations of knowledge, skills and attitudes that they need to apply in the real life

situation in their professional life (Muller, 1998). The level of authenticity of assessment is thus defined by its degree of resemblance to the criterion situation. This idea is extended and specified by the theoretical framework that describes that an assessment can resemble a criterion situation along a number of dimensions.

Complicating matters is the fact that authenticity is subjective (Honebein *et al*., 1993:88; Pretaglia, 1998; Huang, 2002:29) and is dependent on perceptions. This implies that what learners perceive as authentic, is not necessarily the same as what educators and assessment developers see as authentic. If these perceptions do indeed differ, then the fact that educators usually develop authentic assessments according to their own view causes a problem. Although they do their best to develop authentic assessments, this may all be for nothing if the learner does not perceive them as such. This process, known as pre-authentication (Pretaglia, 1998; Huang, 2002:29), can be interpreted either as that it is impossible to design an authentic assessment or that it is very important to examine the experiences of the users of authentic assessments carefully (Nicaise *et al*., 2000:84).

The elements of authentic instruction, authentic assessment / internal processes within the learner / authentic learning / perception of authenticity and authentic achievement led to the framework indicated below (*cf*. Figure 2.1). In this figure, the elements indicate the place of authentic assessment in educational practices.

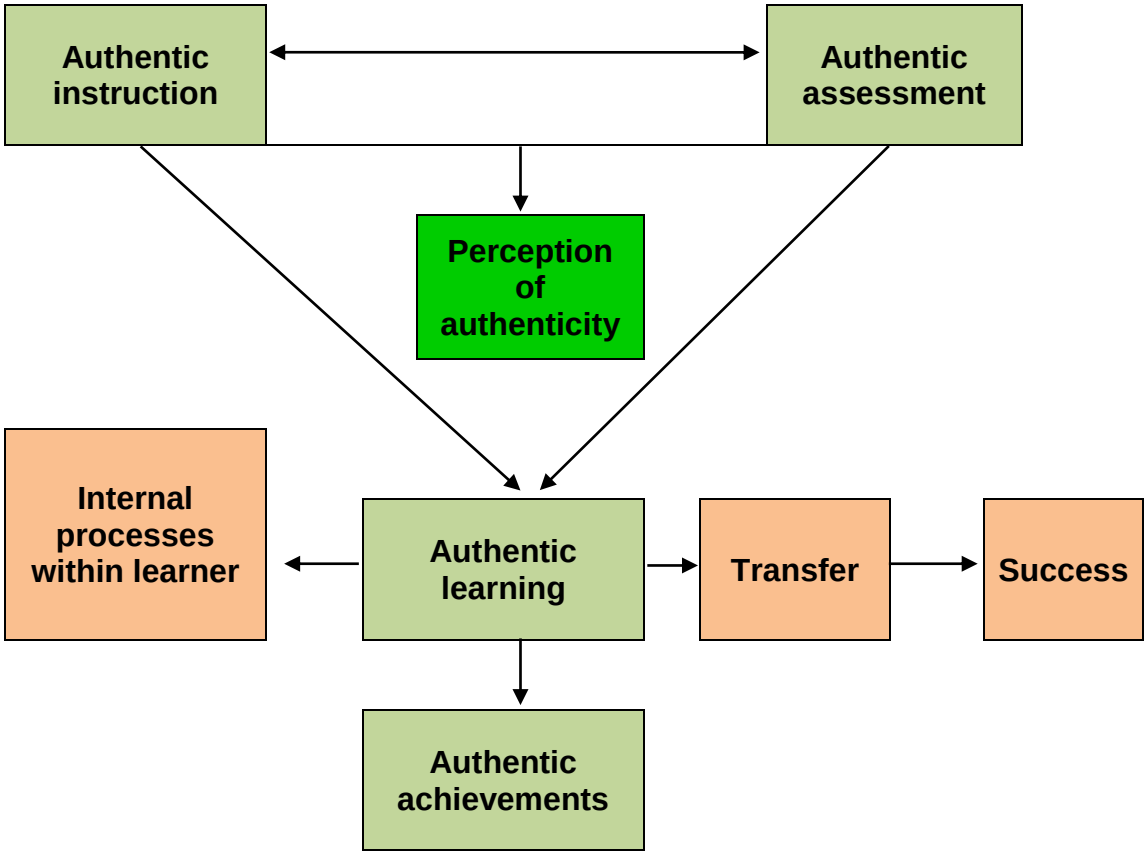

Figure 2.1: General framework for authentic assessment

Figure 2.1 summarizes the important elements of authentic assessment. To influence learners' learning positively, authentic assessment should be aligned to authentic instruction. Authentic assessment requires learners to demonstrate their competencies in a situation that resembles professional practice; authenticity is subjective, which makes it important to take learners' perceptions into account when designing authentic assessment.

Assessment needs to be largely self-directed, with an emphasis on discovery, and with particular value being ascribed to the learners' interpretation of events and facts. The learning strategies need to reflect upon assessment needs, so that the learners are able to master the content. Furthermore, educators need to recognize different ways of finding solutions to help learners to master the content**.** Teaching strategies need to enable learners to construct their own meaningful and conceptually functional representations of the external world (Gulikers *et al*., 2008:80).

Learners are encouraged to discover their learning through assessment tasks and be intrinsically motivated to complete assessment tasks on their own. In

learning that is inspired by constructivism, assessments should not only enable learners to analyse situations and knowledge, but also enable them to transfer skills learned in authentic instruction to the given assessment task (Gulikers *et al*., 2008:79-80).

In this study, the CTA Section A was evaluated to determine whether learners' assessment tasks were authentic or not.

Figure 2.2 below presents a five dimensional model for authentic instruction.

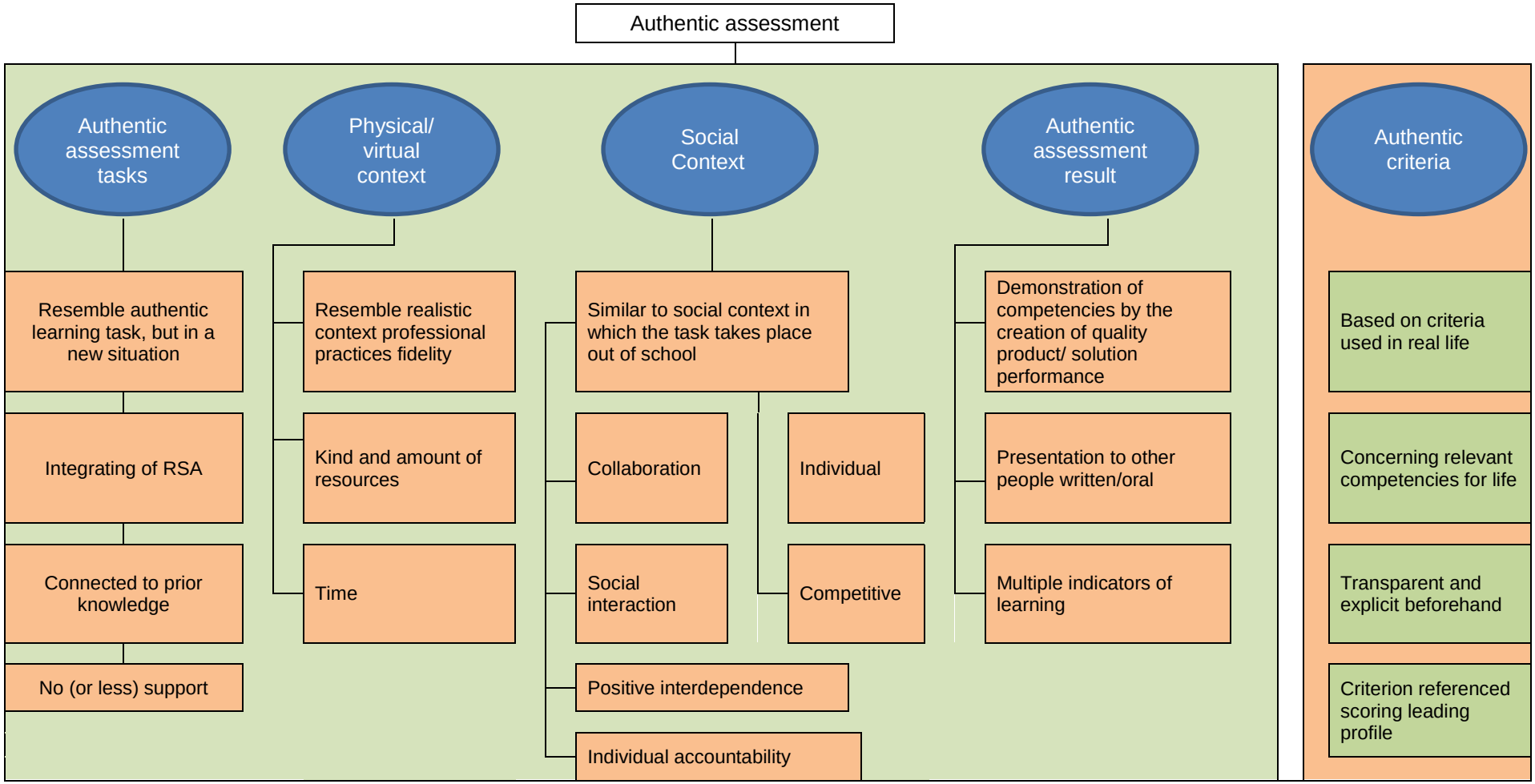

Figure 2.2: Five dimensional model for authentic instruction

(Newmann & Wehlage, 1993:10)

As stated, there is confusion and many differences of opinions exist about what authenticity of assessment really is, and which assessment elements are important for authenticity. Darling-Hammond and Snyder (2000:524) argue that authenticity deals with knowledge, skills and attitudes, and the capacity to apply them in situations. In their view, authentic assessment includes opportunities for the development and examining of learners' thinking actions. This implies that authentic assessment requires learners to demonstrate their learning. The notion of authenticity as continuum (Newmann & Wehlage, 1993:9) resulted in a conceptualization of the following five aspects as dimensions that can vary in their degree of authenticity, namely task, physical content, social context, assessment result and criteria or standards, as depicted in Figure 2.2.

- ✓ **Task**: An authentic task is a problem task that confronts learners with activities that are also carried out in professional practice. The fact that an authentic task is crucial for an authentic assessment is undisputed (Wiggins, 1993:343; Newmann, 1997:366; Herrington & Herrington, 1998:316), but different researchers stress different elements of an authentic task. The framework defines an authentic task as a task that resembles the criterion with respect to the integration of knowledge, skills, attitudes, its complexity, and its ownership (Strijbos, 2004:12). Furthermore, the users of the assessment task should perceive the task, including above elements, as representative, relevant, and meaningful.

  Gulikers *et al*. (2004:67) assert that the users of assessment tasks should perceive the task as representative, relevant and meaningful. This part is dealt with in the learner questionnaire in B16 and B18 (*cf*. Appendix I).

- ✓ **Physical context**: The physical context of an authentic assessment should reflect the way knowledge, skills and attitudes will be used in professional practice (Herrington & Oliver, 2000:25). Fidelity is often used in the context of computer simulations, which describe how closely a simulation imitates reality. Authentic assessment often deals with high fidelity contexts. The presentation of material and the amount of detail presented in the context are important aspects of the degree of fidelity (Gulikers *et al*., 2004:68).

Likewise, an important element of the authenticity of the physical context is that of the number and kinds of resources available (Segers *et al*., 1999:191) which mostly contain relevant as well as irrelevant information (Herrington & Oliver, 2000:30), and should resemble the resources available in the criterion situation (Segers *et al*., 1999:194). For example, McDowell (1998:335) argues that most school tests involve memory work, while out-of-school activities are often intimately engaged with tools and resources (calculators, tables, standards), making such a school test less authentic. Segers *et al*. (1999:195) are of the opinion that it is not appropriate for assessors to deprive learners from performing well by following requirements that are important for providing an authentic physical context when learners are given the opportunity to tackle the tasks. In this regard, Gulikers *et al*. (2004:78) point out that tests are normally administered in a restricted period require, for example, two hours during class time. In real-life, professional activities often involve more time scattered over days or, to the contrary, require fast and immediate reaction in a split second (Gulikers *et al*., 2004:78).

Segers *et al*. (1999:197) indicate that authentic assessment should not rely on unrealistic and arbitrary time constraints. In sum, the level of authenticity of the physical contexts is defined by resemblance of these elements to the criterion situation. Authenticity is addressed in the educators' questionnaire in C38 (*cf*. Addendum H) and in the learners' questionnaire in C29 and C30 (*cf*. Addendum I).

- ✓ **Social context**: Darling-Hammond and Snyder (2000:526) are of the opinion that not only the physical context, but also the social context, influences the authenticity of the assessment. In real life, working together is often the rule rather than the exception. They emphasize that learning and performing out of school mostly take place in a social system. Therefore, a model for authentic assessment should consider social processes in an equivalent situation in reality. The physical context of an authentic assessment should reflect the way knowledge, skills and attitudes will be used in professional practice (Herrington & Oliver, 2000:25-26). Fidelity is often used in the context of computer simulations,

which describes how closely a simulation imitates reality. Authentic assessment often deals with high-fidelity contexts. The presentation of the material and amount of detail presented in the context are important aspects of the degree of fidelity. This was addressed in Section C of the learner questionnaire (*cf*. Appendix I): C31 and C32; and in the educator questionnaire (*cf*. Appendix I) it was addressed in Section C32, Section F51 and F52.

The researcher of this thesis determined whether the activities in CTA cater for the social and physical context of learners to help them to apply the skills they have learned in the real life situation.

According to the five dimensional model for authentic instruction framework *(cf.* Figure 2.2), the physical context of an authentic assessment that reflects the way knowledge, skills and attitudes will be used in professional practice, indicates that if the real situation demands collaboration, then the assessment should also require collaboration processes such as social interaction, positive interdependency and individual accountability (Herrington & Oliver, 2000:24). When, however, the assessment is individual, the social context should stimulate some kind of competition between learners.

- ✓ **Assessment result or form**: Assessment involves an assessment assignment (in a certain physical and social context) that leads to an assessment result, which is then evaluated against certain assessment criteria (Moerkerke *et al*., 1999:121). The assessment result is related to the kind and amount of the content of the assessment (Gulikers *et al*., 2004:69).

  In the framework, an authentic result is characterized by four elements: (1) a quality product or performance-based task, which refers to learners being asked to produce a product in a real-life simulation. This product or performance should be (2) in a demonstration that permits making valid inferences about underlying competencies (Darling-Hammond & Snyder, 2000:537). Since the demonstration of relevant competencies is often not possible in one single test, authentic assessment should involve (3) a full

array of tasks and multiple indicators or learning, in order to come to fair conclusions (Darling-Hammond & Snyder, 2000:539). According to Uhlenbeck (2002:12), (4) the combination of different assessment methods should cover the whole range of competences when teaching learners adequately.

- ✓ **Criteria and standards**: Criteria are those characteristics of the assessment result that are valued references. Standards indicate the level of performance expected from various grades and ages of learners (Biggs, 1996:5). Sluijsmans (2002:34) indicates that setting criteria and making them explicit and transparent to learners beforehand is important in authentic assessment, because this guides learning and, after all, in real life, employees usually know on what criteria their performances will be judged. This implies that authentic assessment requires criterion-referenced judgment (Gulikers *et al*., 2004:68). Moreover, some criteria should be related to a realistic outcome, explicating characteristics or requirements of the product, performance or solutions that learners need to create. Furthermore, criteria and standards should concern the development of relevant professional competencies and should be based on criteria used in the real-life situation (Darling-Hammond & Snyder, 2000:539). This section is dealt with in the educators' questionnaire in B20 (*cf*. Appendix H).

Besides basing the criteria on the criterion situation in real-life, criteria of an authentic assessment can also be based on the interpretation of the other four dimensions of the framework (Gulikers *et al*., 2004:78). For example, if the physical context determines that authentic assessment of a competency requires five hours, a criterion should be that learners need to produce the assessment result within five hours. On the other hand, criteria based on professional practice can also guide the interpretation of the other four dimensions of authentic assessment. In other words, the framework argues for a reciprocal relationship between the criterion dimension and the other four dimensions (Gulikers *et al*., 2004:78).

What does all of this mean when educators or instructional designers try to develop authentic assessment? What do they need to consider?

The first consideration deals with *predictive validity*. If the educational goal of developing competent employees is pursued, then increasing authenticity of an assessment will be valuable (Gulikers *et al*., 2004:78). More authenticity is likely to increase the predictive validity of assessment because of the resemblance between the assessment and real professional practice. However, one should not throw the baby out with the bath water. Objective tests are still very useful for certain purposes as high-stakes summative assessments on individual achievement, where predicting learners' ability to function competently in future professional practice is not the purpose (Gulikers *et al*., 2004:78).

Another consideration in designing authentic assessment is that one should not lose sight of the educational level of the learners (Gulikers *et al*., 2008:83). Slow learners may not be able to deal with authenticity of a real, complex, professional situation. If they are forced to do this, it may result in cognitive overload and have a negative impact on learning (Sweller & Van Merrienboer, 1998:256). The question that immediately comes to mind in this context is: how is authentic assessment created for learners who are not prepared to function as beginning professionals? The answer is that authenticity of an assessment should be defined by its degree of resemblance to the criterion situation. Assessment can still be authentic as long as the abstracted situation requires learners to perform the whole range of competencies which are described in the National Curriculum Statement for EMS. The abstraction results from simplifying contextual factors that complicate the performance of the whole competency (Gulikers *et al*., 2008:82-83).

A third consideration also sheds light on the question stated in the previous sections, namely the subjectivity of authenticity. The perception of what authenticity is may change because of educational level, personal interest, age or amount of practical experience with professional practice (Honebein *et al*., 1993:29). This implies that the five dimensions that are depicted in the framework for authentic assessment are not absolute, but rather variable. It is possible that assessing professional practice requires more authenticity of the physical context than when assessing first year learners, who usually or often

take learners' changing perspectives into account when designing authentic assessment (Gulikers *et al*., 2008:80).

This researcher's view on CTA is that there is a need to engage learners in authentic assessment. Learners must take ownership of these assessment tasks given in CTA while they are tackling the problems, thus proving their originality. The physical context of the environment where learners are assessed should provide stimulation regarding resources such as the use of computers to search for information. Learners could be encouraged to work in groups rather than in isolation in some of the activities as part of the social context with regard to the CTA instrument: learning to work with others and to stimulate their learning environment so that they are able to learn from each other and identify where they lack information concerning performing assessment tasks. Furthermore, CTA design could be geared to yield results according to the content coverage in teaching and learning. Lastly, the criteria of CTA must be transparent: learners need to be informed beforehand on what is expected from this instrument.

In this study, CTA was evaluated to determine whether it is designed in a way in which it indicates the standard and level of performance expected from Grade 9 learners and whether criteria were made transparent and explicit. This was addressed in the learner questionnaire in Section B (*cf*. Appendices I) and in the educator questionnaire in B20 (*cf*. Appendix H).

The next section elaborates on flexibility as a feature of quality assessment.

2.3.4 Flexibility

The idea of a flexible assessment strategy that gives feedback to learners about their learning achievements underpins quality assessment practice. When learners are ready to demonstrate that they mastered an outcome, the ideal is that they should have the opportunity to do so. The focus on the individual progress of each learner towards achieving the learning outcomes implies that assessment should be flexible and should provide feedback useful to every individual learner about her/his progress (Draud, 2008:2).

Another important consequence of the principle of flexibility is the idea that a variety of assessment methods should be employed in the context of flexibility

for refinement and improvement (Sutherland & Peckham, 1998:98-103; Welch, 1999:23; Williams, 2006a:106).

Flexible assessment has two implications: educators should review all assessment procedures in use and ensure that they involve sufficient numbers and types of assessment that will result in deep learning for their learners (Sutherland & Peckham, 1998:103).

Flexibility will have important implications for implementation. An important misconception to be addressed at this point is that the call to use a variety of methods implies that every assignment/task should test knowledge, skills and attitudes (where appropriate). An educators' assessment strategy should demonstrate that a variety of strategies were used across a learning programme (Welch, 1999:23). Another misconception is the idea that summative examination is now taboo. In this context, if motivated, summative assessment has its place. Emphasis, however, should be on informal and formal on-going formative assessment, instead of being the most important assessment event. Summative assessment therefore becomes one of several other assessment strategies to use (Welch, 1999:24).

According to the new policy document (Department of Education, 2007a:18), it is important that learners who might experience barriers to learning and development are identified early, assessed and provided with learning support, and all assessment tasks should be adapted to accommodate learning needs.

In this study, flexibility in conducting CTA in terms of accommodating learners who experienced barriers to learning at mainstream schools in conducting CTA formed part of the empirical investigation to ascertain whether learners received support from school and from relevant district support teams. If learners did not get this support, it could point to the fact that CTA was not accommodating learners with learning barriers. This part was dealt with in the learner and educators' questionnaires in Section B19 and B20 (*cf*. Appendices I & H).

The next section explores expanded opportunity as a feature of quality assessment.

2.3.5 Expanded opportunity in assessment

Different types of assessment should be used to afford all learners different opportunities to be assessed in different ways (Du Toit & Du Toit, 2004:5; Vandeyar & Killen, 2006:388). Opportunity *to learn* refers to equitable conditions or circumstances within the school or classroom that promote learning for all learners. It includes the provision of curricula, learning materials, facilities, educators and instructional experiences that enable learners to achieve high standards. Expanded opportunity also relates to the absence of barriers that prevent learning (Vandeyar & Killen, 2006:388).

In connection with assessment, Williams (2006b:285) notes that opportunity to learn relates to the provision of adequate and timely instruction of specific content and skills prior to taking a test. She adds that opportunity to learn may be measured by time spent in reviewing, practising or applying a particular concept or by the amount and depth of content covered with particular groups of learners.

When learners are tested with high-stakes assessments, evidence must be provided that the learners have had adequate opportunity to learn the material on which they are being tested. Recent legislative proposals have called for the development of opportunity-to-learn standards that coincide with content standards and performance standards (National Department of Education, 2002; Department of Education, 2002c; Department of Education, 2007b; Department of Basic Education, 2011).

According to Du Toit and Du Toit (2004:5), educators must provide more than one opportunity to learners, if they are not successful, to demonstrate important learning. Rigid time frames and schedules must not restrict learning, although there must be limits to every expanded learning opportunity. The enhancement of opportunities by being focused, the creation of motivational channels, and the design of a clear path all contribute to the formulation of such limits. The application of expanded opportunity also implies that educators must change their teaching methods to ensure successful learning for all learners and that there will be no restriction on the number of successful learners. Furthermore, all learners must have the

opportunity to be exposed to a meaningful curriculum, quality learning experiences and multiple resources (Du Toit & Du Toit, 2004:5).

The researcher of this thesis investigated whether learners are given expanded opportunity with regard to CTA assessments, because it is essential that they should be given an opportunity to write a supplementary test if they fail to meet the required standards. This part is addressed in the learner questionnaire in Section B18 (*cf*. Appendix I) and in the educator questionnaire in B19 (*cf*. 2.2.4.5; Appendix H).

In the following section, assessment as a continuous process is explored as a feature of quality assessment.

2.3.6 Assessment as continuous process

An important principle concerning assessment is that it is a *continuous* and an *on-going* process. A statement that illustrates this is that assessment should be on-going, illustrating that a learner's progress will be monitored continuously (Department of Education, 2006c:71). This strongly links to the notion of a lifelong learning environment which implies that learners will be afforded several opportunities, over a period of time, to demonstrate the progress of their learning. In this study, assessment as a continuous process will be viewed as a course of action collecting valid and consistent data regarding the achievement of learners by means of a range of devices, processes, procedures and frameworks (Department of Education, 2002b:7; Umalusi, 2004:14; Ramotlhale, 2008:11).

A central theme in assessment being a continuous process seems the gathering of information or data on a regular basis. Ramotlhale (2008:11) identifies domains such as cognitive, affective, and psychomotor that should be integrated during teaching and assessment.

Greaney (2001:6) points out that continuous assessment (CASS) is the chief method by which assessment takes place in the Revised National Curriculum Statement. It covers all the OBE assessment principles and ensures that assessment is continuous.

The nature of continuous assessment can be described as follows (Greaney, 2001:6; Department of Education, 2006a:71):

- It takes place over a period of time and is on-going: Learning is assessed regularly and the records of learners' progress are updated throughout the year.
- It supports the growth and development of learners: Learners become active participants in learning and assessment. They understand the criteria that are used for assessment activities, are involved in self-evaluation, set individual targets for themselves, reflect on their learning, and experience raised self-esteem.
- It provides feedback from learning and teaching: Feedback is a crucial element in formative assessment. Methods of feedback include appropriate questioning; focusing the educator's oral and written comments on what was intended to be achieved by an assessment activity, and encouragement to a learner.
- It allows for the integrated assessment: This may include assessing a number of related Learning Outcomes within a single activity, and combing a number of different assessment methods. Competence in particular Learning Outcomes can be demonstrated in many different ways, and thus a variety of assessment methods and opportunities must be provided through which learners can demonstrate their ability.
- It uses strategies that cater for a variety of learner needs, including those of linguistic, physical, psychological, emotional and cultural nature: CASS allows educators to be sensitive to learners with special education needs and to overcome barriers to learning through flexible approaches. In any group of learners, there are different rates and styles of learning. All learners do not need to be assessed at the same time and in the same way.
- It allows for summative assessment: The accumulation of the results of continuous assessment activities provides an overall picture of all learners' progress at a given time. Summative assessment needs to be planned carefully from the beginning of the year to include a variety of assessment strategies – for example exercises, tasks, projects, school and class tests

– which will provide learners with a range of opportunities to show what they have learned.

- The choice of assessment strategies is a subjective one, unique to each educator, grade and school, and dependent on the educator's professional judgement. The availability of space and resources influences this decision, but even when resources are similar, educators differ in the way that they make their choices.

According to Sigh (2004:6), the methods chosen for assessment activities must be appropriate to the Assessment Standards to be assessed, and the purpose of the assessment must be clearly understood by all the learners and educators involved. Competence can be demonstrated in a number of ways. Thus a variety of methods is needed to give learners an opportunity to demonstrate their abilities more fully.

CTA should be linked to the school-based continuous assessment so that the educator cannot complain and raise negative perceptions that CTA requires a different curriculum than for continuous assessment. The design of this instrument should be in line with the learning programme offered at school and the assessment standards in the programmes. Furthermore, continuous assessment should be the benchmark for CTA assessment (Sigh, 2004:6).

Openness, transparency and accountability as features of quality assessment are explained below.

2.3.7 Openness, transparency and accountability

Coetzee-Van Rooy (2011:311) points out that this principle emphasizes that those learners should be informed regarding the criteria, method and context of assessment. Expectations should be made clear as well. Furthermore, the author indicates that the principle of accountability redirects the responsibility for learning from the institution to the learner. Learners in this context have to accept responsibility for their own learning progress. In this spirit, self and peer assessment are regarded as important assessment strategies.

Race (2003:73) explains transparency as the extent to which learners know where the goalposts are. The goalposts may be indicated by the intended

learning outcomes, matched nicely to the assessment criteria which specify the standards to which these intended outcomes are to be demonstrated by learners and also specify the forms in which learners will present evidence of their achievement of the outcomes.

The researcher is of the opinion that, when the CTA instrument is designed, specific characteristics of subgroups (learners, parents/caregivers, communities, District Officials and educational bodies) should be considered, in conjunction with standards. Assessment must be sensitive to various forms of diversity, including the cultural, within and across subgroups. It cannot be assumed that assessments will be effective and valid for all subgroups.

In this study, the researcher looked at whether learners are given assessment criteria beforehand. These aspects are addressed in the educator questionnaire in Section B7, B8, B9 and B10 (*cf*. Appendix H) and in the learner questionnaire in Section B13, B14, and B15 (*cf*. Appendix I).

In the following section, equity as a feature of quality assessment is addressed.

2.3.8 Equity

In the USA, the National Center for Research Evaluation Standards and Testing (Joan *et al*., 1991:1) defines equity in assessment as follows: Equity is the concern for fairness, implying that assessment is free from bias or favouritism. Assessment should be fair and enable all learners to show what they can do. At the minimum, all assessment should be reviewed to eliminate stereotypes, situations that may favour one culture over another, excessive language demands that prevent some learners from showing their knowledge and promote the potential to include learners with disabilities or limited English proficiency (Joan *et al*., 1991:1).

According to Scherman *et al*. (2006:174), equitable assessment practices allow for learners who learn in different ways. This implies that learners who have different backgrounds which act as unique learning frameworks, who may be at different developmental stages and develop different understanding of the instructional process, such as a learning difficulty or lateral thinking, are

all accommodated. This clearly illustrates the wide range of concepts incorporated in equitable assessment.

The researcher investigated whether the current CTA involves a range and balance of background contexts in which assessment items are presented, and whether a range and balance of types of assessment instruments and modes of response, including a balance and range of visual and linguistic material involve a range and balance of conditions. These aspects are addressed in the learner questionnaire in Section B20; and in Section C30 and C31 (*cf*. Appendix I).

When indicators of material, social, psychological or home backgrounds are used – such as parental values that build up pressure for the child to achieve well – one is more likely to find the relationship between home background and school achievement as factors that influence achievement (Greaney & Kellaghan, 1996:11). Concerning CTA, learners from disadvantaged families find it difficult because parents/caregivers can perhaps not afford to buy the necessary magazines or computers to enable them to complete their assessment tasks, especially at some schools where the Internet is not in use.

Concerning the principle of equity, it is important to remember that the current education system is in a transformation phase. Current South African education is a result of many years of inequitable distribution of money and human resources along racial lines (Greaney & Kellaghan, 1996:11). Achieving equity is a struggle that might continue for decades. Several South African schools and institutions of higher education are still disadvantaged (Scherman *et al*., 2006:174).

One example of inequity that prevails today is not considering the levels of English Second Language proficiency of learners and their possible effect on academic language achievement. Without the necessary cognitive academic language proficiency (CALP), learners could be severely hampered in trying to internalize a mass of established concepts (Scherman *et al*., 2006:174). Third language users should be provided with some sort of language support applicable to the programmes they are involved in (Scherman *et al*., 2006:174).

The following statement illustrates the principle of equity at work in the assessment context: assessment methods will not disadvantage individuals or groups by hindering or limiting them in ways unrelated to the evidence sought (Department of Education, 1998:19).

Assessment should be conducted in such a way that all learners have an equal chance to succeed. This is demonstrated in the notion that assessment criteria are communicated overtly at the onset of a module. It should enable every learner to monitor her/his progress towards mastering the outcomes involved (Department of Education, 1998:19).

Fairness as a feature of quality assessment is discussed next.

2.3.9 Fairness

SAQA (2001:13) explains fairness as taking account of and addressing issues pertaining to the inequality of opportunities, resources and appropriate teaching and learning approaches in terms of acquisition of knowledge, understanding and skills. Here, issues of bias in respect of ethnicity, gender, age, social class and race in the assessment approaches, instruments and materials are important. In addition, what is being assessed has to be clear (SAQA, 2001:16; Vandeyar & Killen, 2003:121; Vandeyar & Killen, 2006:392; Nitko & Brookhart, 2007:43).

According to Scherman *et al*. (2006:174), fairness in assessment obviously stretches well beyond mere cultural fairness. Fairness in assessment is created in terms of setting procedures for controlling intervening strategies.

Consideration for the cultural background of learners is very important. All learners have to have an equal opportunity to comprehend complex thinking and problem-solving skills that are the targets of the new assessment approach (Vandeyar & Killen, 2003:121). Looking at CTA, the researcher of this thesis believes that there is no fairness in this assessment because of language barriers and vastly different cultural backgrounds. English First Language learners are assessed through the same CTA tasks as learners who are doing English as Second Language. It becomes evident that fairness in assessment is often accomplished through accommodation of existing assessment where adjustments are made in terms of procedures and setting

or controlling of an intervening culture, in the assessment of a specific construct. These accommodations to establish fairness encompass change in scheduling, setting, equipment or technology, presentation and response (Vandeyar & Killen, 2003:121).

The following should also be taken into account, namely the example of the language benefits Afrikaans learners have when they are undergoing CTA in their mother tongue, while African language learners are disadvantaged by undergoing CTA in a different language. Language barriers could, according to the researcher, contribute to learners failing CTA.

Fairness also relates to issues of validity and reliability, which are intrinsically related to appropriate accommodation and adaptations of assessment. Elliot and Hufton (2001) regard accommodation as providing access to the instrument and assessing a learner without exposure to social practices. The latter would thus translate into an inequitable assessment practice and the learner would be limited in the use of reading strategies such as reading for meaning and utilizing context clues. Adaptation of CEM (Curriculum, Evaluation and Management Centre, 2006) instruments to conform to the South African context would thus need to consider exposure to specific contexts and cultural practices and even types of representation to ensure that these do not act as intervening variables and thus undermine the validity of the instrument in the South African context by confounding the underlying constructs being examined. The aspects of validity and reliability were addressed in the educator questionnaire in Section B12 and B16; Section C38, C41 and C42 (*cf*. Appendix H). In the learner questionnaire they are addressed in Section B21; Section C31, C36 and C38 (*cf*. Appendix I).

The following section addresses transferability and generalizability as features of quality assessment.

2.3.10 Transferability and generalizability

The results of assessment should be able to support accurate generalizations about learners' capability. Lorrie (2000:11) indicates that there is a close relationship between truly understanding a concept and being able to transfer knowledge and use it in new situations. In contrast to memorization and the

behaviourist assumptions that each application must be taught as a separate learning objective, true understanding is flexible, connected and generalizable (Lorrie, 2000:11-12).

Learners should be able to transfer the skills learned in class to deal with the demands of CTA. They should be able to connect the content taught in class to the assessment situation. They must be able to apply the knowledge and skills they learned to CTA tasks.

In this research, CTA was evaluated to see whether the instrument helped learners to transfer skills learned in Section A to Section B to determine whether there was a correlation or relationship between these two sections. These aspects are addressed in the learner questionnaire (*cf*. Appendix I) in Section B8, B9 and B18; and Section C32 and C33.

In the next section, the researcher elaborates on cognitive complexity as a feature of quality assessment.

2.3.11 Cognitive complexity

Segers *et al*. (2003:12) argue that cognitive complexity is grounded in the critical thinking process. It is important to recognize that cognitive complexity in assessment should focus on higher-order thinking processes as opposed to lower-order thinking processes.

In the next section, Bloom's Taxonomy will be expounded to get a clear understanding on how CTA can be designed to match different cognitive abilities.

2.3.11.1 Bloom's Taxonomy

Benjamin Bloom created a hierarchy of cognitive skills – called Bloom's Taxonomy – that is often used to categorize the levels of cognitive involvement (thinking skills) in educational settings (Nitko & Brookhart, 2007:25). The original taxonomy provides a good structure to assist educators in writing objectives and assessments. It can be divided into two levels: Level I (the lower level) contains knowledge, comprehension and application; Level II (the higher level) includes application, analysis, synthesis and evaluation (Nitko & Brookhart, 2007:25).

Bloom's Taxonomy is a multi-tiered model of classifying thinking according to six cognitive levels of complexity. Throughout the years, the levels have often been depicted as a stairway, leading many educators to encourage their learners to climb to a higher level of thought. The lowest three levels are knowledge, comprehension and synthesis, and evaluation. The taxonomy is hierachical: each level is subsumed by the higher levels. In other words, a student functioning at the application level has aslo mastered the material at the knowledge and comprehension level. (Wilson, 2006:2). One can easily see how this arrangement led to natural divisions of lower and higher level thinking.

In the revised version of the taxonomy, Bloom's major categories were changed from nouns to *verb* forms. Additionally, the lowest level of the original knowledge was renamed and became *remembering*. Finally, comprehension and synthesis were retitled to *understanding and creating*.

Figures 2.3 and 2.4 below present adapted visual illustrations of the revised version of Bloom's Taxonomy and the original educational objective version, respectively.

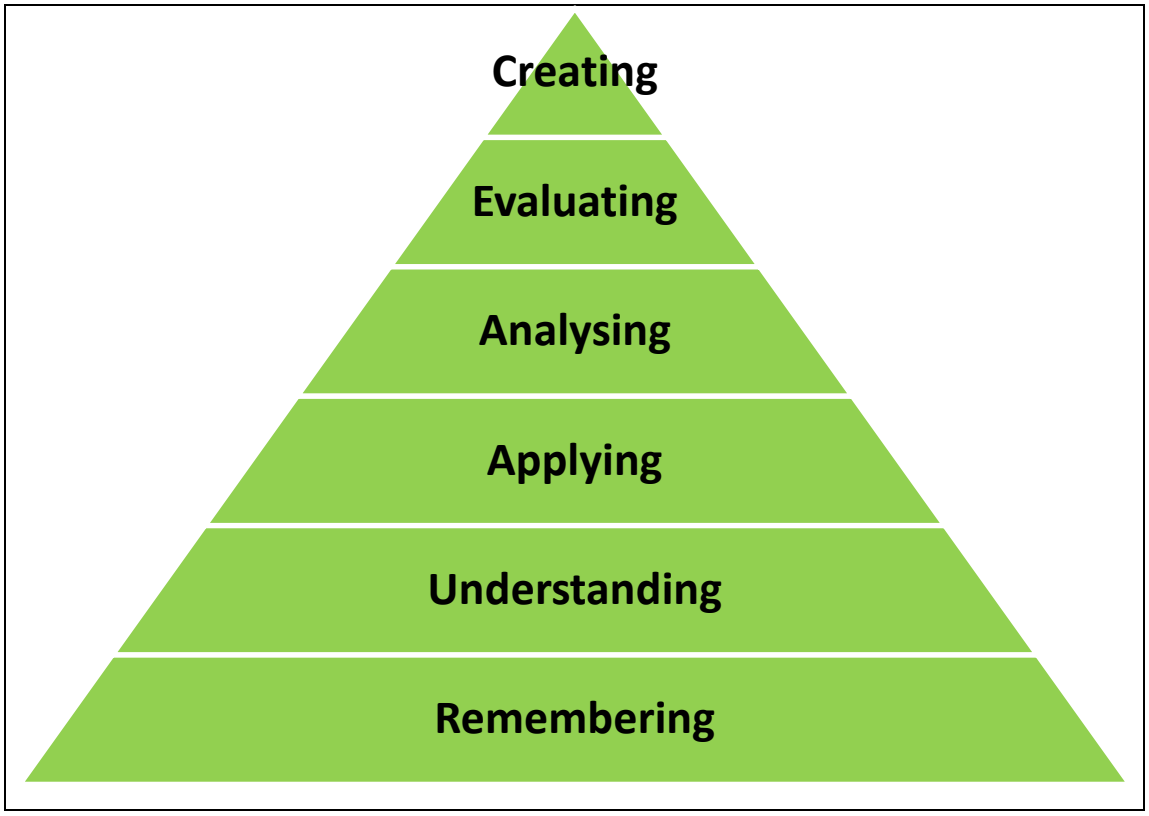

Figure 2.3: Bloom's Taxonomy – revised version

(Adapted from Wilson, 2006:3)

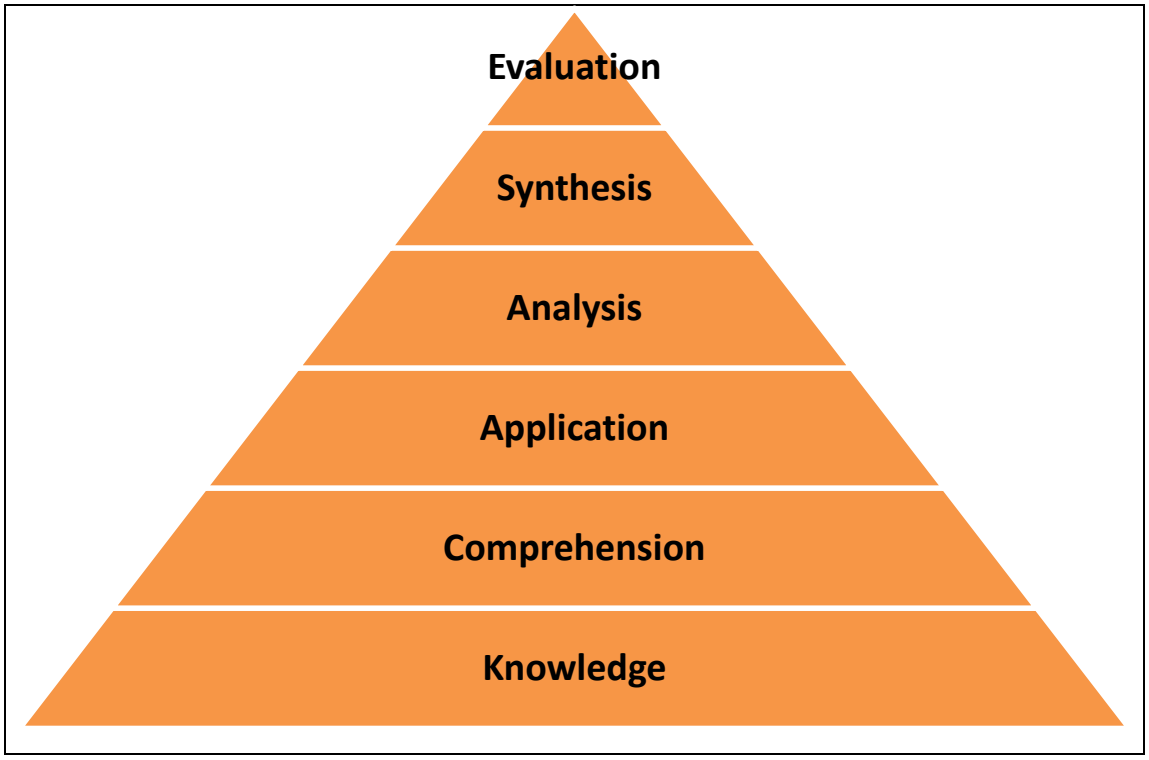

Figure 2.4: Bloom's Taxonomy – original version

(Adapted from Wilson, 2006:3)

The two models above are referred to as the most used cognitive domains: they can be viewed as a sequence of progressive contextualization of learning material. In this research it was established whether CTA tasks enable learners to achieve outcomes at a variety of cognitive difficulty levels.

The knowledge dimension comprises the following aspects (Anderson & Krathwohl, 2001:90-91):

- Factual knowledge, which is basic to specific disciplines. This dimension refers to essential facts, terminology details or elements learners must know or be familiar with in order to understand a discipline or solve a problem.
- Conceptual knowledge, which is knowledge of classifications, principles, generalizations, theories, models or structures pertinent to a particular disciplinary area.
- Procedural knowledge, which refers to information that helps learners to do something specific from a discipline, subject, and area of study. It also

refers to methods of inquiry, very specific or finite skills, algorithms, techniques and particular methodologies.

- Metacognitive knowledge, which is the awareness of one's own cognition and particular cognitive processes. It is strategic or reflective knowledge on how to go about solving problems and cognitive tasks, to including contextual and conditional knowledge and knowledge itself.

Anderson and Krathwohl (2001:67-68) indicate that assessment tasks should comprise of the following cognitive activities according to the new revised Taxonomy of Bloom:

- ✓ **Remembering**: Retrieving, recognizing and recalling relevant knowledge from long-term memory.
- ✓ **Understanding**: Constructing meaning from oral to written, and graphic messages through interpreting, exemplifying, classifying, summarizing, inferring, comparing and explaining.
- ✓ **Applying**: Carrying out or using a procedure through executing or implementing.
- ✓ **Analysing**: Breaking material into constituent parts, determining how the parts relate to one another and to an overall structure or purpose through differentiating, organizing, and attributing.
- ✓ **Evaluating**: Making judgments based on criteria and standards through checking and critiquing.
- ✓ **Creating**: Putting elements together to form a coherent or functional whole, reorganizing elements into a new pattern or structure through generating, planning or producing (Krathwohl, 2001:67-67; Wilson, 2006:2-5).

In this thesis, the framework of Bloom's Taxonomy was applied to determine whether the CTA instrument conforms to Bloom's Taxonomy. Assessment includes assessment tasks at lower and higher-order thinking levels. This framework is addressed in Section B of the learner questionnaire (*cf*. Appendix I) in B6, B13, B14, B16 and B20.

Content quality as a feature of quality assessment is clarified in the next section.

2.3.12 Content quality

When OBE was first introduced, a perception was created that learning content was no longer a significant component of the teaching-learning agenda. Some of the documents produced by the Department of Education tended to use constructs such as *outcomes-based* and *content-based* as if they were opposites, giving the impression that if one was in, the other was out. The Revised National Curriculum Statement (Department of Education, 2002c:12) describes one of its five principles as *a high level of skills and knowledge for all.* Although skills precede knowledge in this statement, it can be inferred that whatever skills learners need to acquire should involve processing, trying to comprehend particular knowledge or learning content selected from prescribed learning areas. The teaching methods and strategies which the educator applies, and the learning styles he or she accommodates among learners, determine the effectiveness with which learning outcomes will be achieved by learners (Department of Education, 2002c:12).

Chehore (2006:162) indicates that, generally, the curriculum contains what learners should know and be able to do (content), how it is taught (instruction) and how it is measured (assessment). Due to globalization and availability of vast sources of information, the content should include assessment tasks which should represent a full curriculum. Furthermore, Chehore (2006:162) asserts that content should include the following aspects:

- Different ways of knowing and validating information
- Approaches from multiple perspectives
- Identifying and connecting ideas, concepts and applications
- Communication
- Metacognitive abilities
- Awareness of contributions of different cultures to current knowledge

- Content that is relevant to the real world environment.

There has been a shift away from a largely content-focused and educator-centred curriculum to one which is learner-centred and emphasizes knowledge, skills, attitudes and values. Subject content has been modernized to make it more relevant to youngsters in the twenty-first century.

In this study, the researcher determined whether the content assessed in CTA is part of the learning programme for Grade 9 and complies with the three criteria for quality content coverage as explained. This aspect is covered in Section B of the learner questionnaire: B6, B10, B13, B16 and B18 (*cf*. Appendix I).

Meaningfulness in assessment is important to guarantee quality. Meaningfulness in assessment is therefore the focus of the next section.

2.3.13 Meaningfulness in assessment

Assessment should result in worthwhile educational experiences and greater motivation for performance (Barnes, 2002:55).

Barnes (2002:55) asserts that learning and meaning are interconnected. He further explains that to learn is to strive for meaning, and to have learned something is to have grasped its meaning. According to Marsh (2007:6), humans are dependent upon their ability to construct their own reality and are driven to use meaning as a way of understanding the reality they have created. Meaning is believed to be created and reshaped based on how people interpret and reinterpret what they have learned.

In this study, the researcher of this thesis will investigate whether CTA focuses on meaningful assessment tasks. This aspect is covered in Section B of the learner questionnaire in B16 and C32 (*cf*. Appendix I).

In the next section, cost efficiency as a feature of quality assessment is explained.

2.3.14 Cost efficiency in assessment

Efficient data-collection designs and scoring procedures are needed with the performance-based assessment or simulation tasks which learners need to carry out in CTA. Dating from the teaching experience of the researcher,

financial, educational and material resources are often difficult to find and are either not budgeted for at some schools or money allocated for the resources is often ill-timed.

According to Colclough (2005:36), it is difficult for local educators without the financial resources, the technical means and the intellectual milieu to invent performance assessment with necessary power and credibility. The process becomes much time-consuming, cumbersome and is also fraught with complications and difficulties (for example – the integration of Learning Areas). It is easier to set outcomes, but difficult to set criteria and establish performance levels.

Smith *et al*. (2000:158) report that educators encounter a lot of problems in implementing Outcomes-Based Assessment (OBA). These include purchase of materials for testing, the wording of questions, insufficient time limits, inadequately prepared scores, vague scoring rubrics and lack of or inadequacy of training. Stiggins (1995:11-19) suggests that for instruction and assessment practices to improve the quality of the education system, there is a need for educator professional development in a total environment that demands and supports quality assessment. Educators should be encouraged to engage in partnerships, and in an integrated professional development plan that will help with the allocation of resources to support quality in assessment practices.

Yet another problem is the question of allowing learners to progress at their own pace. Smith (2005:25) observes that schools can hardly make any provision for a learner who falls three months behind a class or who is four months ahead of a class. It is vital to know – rather than to speculate – to what extent this learner performance is paced with regard to promotion or progression, as is practicable in the South African curriculum context.

If the policymakers decide to amend the assessment instruments, they must work together with all stakeholders to ensure a smooth transition when there are changes in the implementation of policies, so that they make sure that the required resources are available and efficient (Smith, 2005:25). With regard to the design of CTA, it can be regarded as a sound instrument, but the

problem lies with the cost-related materials needed and resources for different Learning Areas' CTAs (Smith, 2005:25). This aspect is addressed in the learner questionnaire (*cf*. Appendix I) in Section C22 and C31; in the educator questionnaire (*cf*. Appendix H) it was addressed in Section F51and F53; and Section D43, D44, D45, D46 and D47.

Finally, balancing assessment of and assessment for learning is explored as a feature of quality assessment.

2.3.15 Assessment *of* learning versus assessment *for* learning

Harlen (2006a:116) points out that the most challenging part in assessment is balancing assessment of and assessment for learning. Assessment of learning (summative assessment) can be described as a way of summarizing what has been learned, or the process of interpreting information on learner assessments for planning future teaching (Harlen, 2007:121).

The same author, Harlen (2006b:119), asserts that assessment for learning (formative assessment) is essential for a continuing and repeated cycle of events in which the educator and learners use information from on-going activities to determine the next step in learning and how to take those steps.

Figure 2.5 below presents assessment for learning (formative assessment) as a cycle of events.

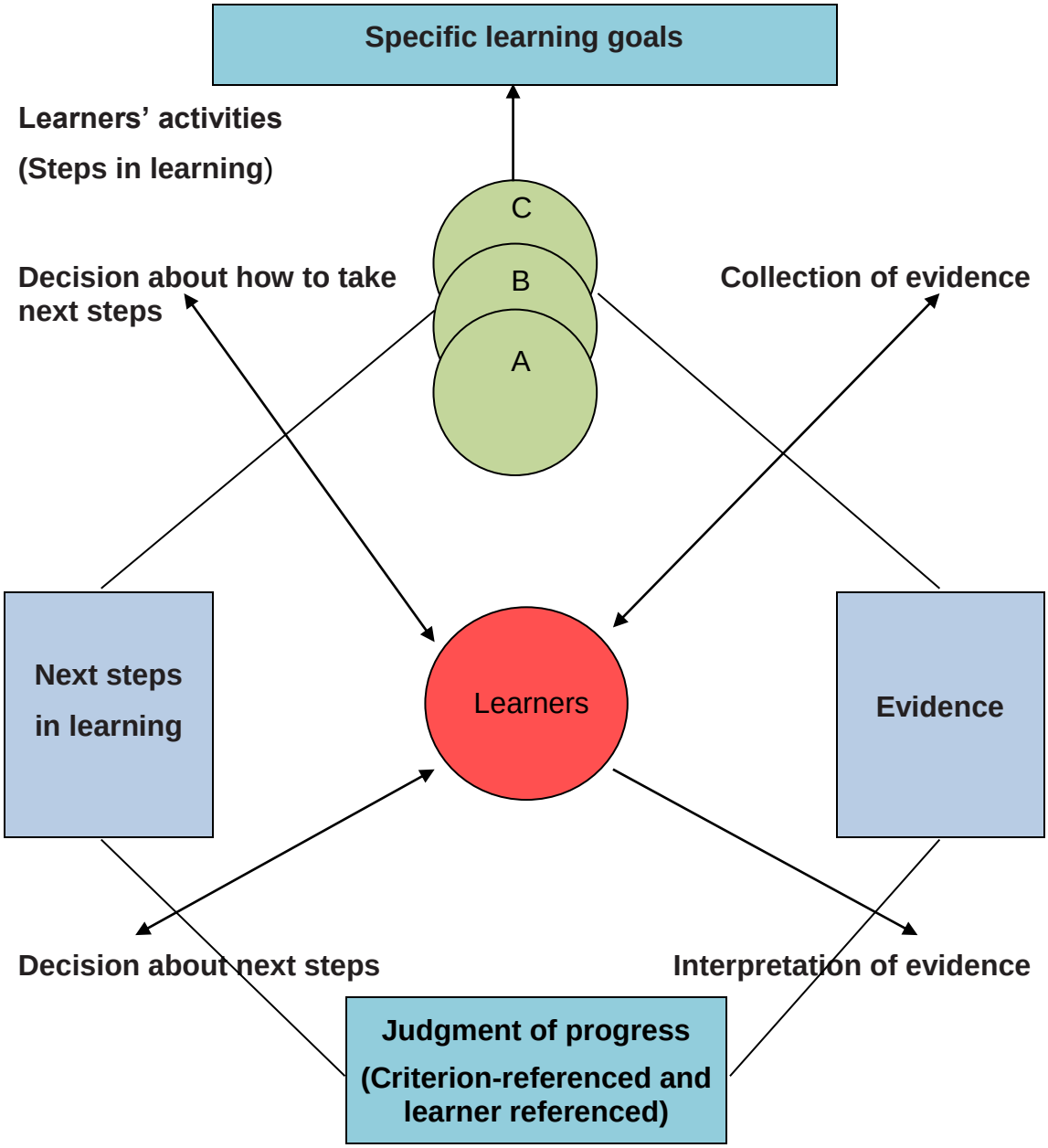

Figure 2.5: Assessment for learning as a cycle of events

(Adapted from Gardner, 2006:105; Harlen, 2006a:109; Harlen, 2007:120)

Figure 2.5 brings together the outcomes of the lesson and components of using assessment for learning. Activities A, B and C are directed towards the outcomes of the lesson, or series of lessons, on a topic. These outcomes, shared with learners by the educator, are expressed in specific terms, for example, in EMS, to plan and carry out market research on customers' preferences concerning a certain product. The learners work on activity A, directed at achieving the outcomes of a lesson, which provides an opportunity for both the educator and the learner to obtain evidence relating to

achievement of these outcomes. To interpret evidence, both the educator and the learner need to know what sound planning means, so learners must have some understanding of the criteria that need to be applied when assessing their work. For example, the learners have to establish whether the planned research they are conducting is taking into account all variables. What evidence will be gathered and how? The judgment will lead to a decision about the relevant next steps to be carried out and so on, to activity B, which is the result of deciding how to improve or move on (Gardner, 2006:105; Harlen, 2007:119).

Harlen (2007:120) indicates that learners are the centre of the assessment for learning process, since they are the ones who do the learning. The two-headed arrows linking learners to various parts of the assessment cycle indicate that learners both receive feedback from the educator and provide information and participate in decisions where appropriate.

Harlen (2007:120) asserts that in formative assessment, judgements about progress and decisions about the next step take into account circumstances, past learning and effort of individual learners, as well as what they are able to do in relation to the goals of work at a particular time. Thus the judgements are both learner- and criterion-referenced. As long as no comparisons are made between learners, which in formative assessment should not be the case, this approach supports learning far more than applying the same standards to all learners. The latter can lead to demotivation if lower-achieving learners are constantly judged by standards too far beyond their reach.

According to Earl (2003:67-68), striking an appropriate balance between the uneasy twins – assessment *of* learning and assessment *for* learning – might be especially tricky because what aids the former may be deleterious to the latter and vice versa. Some imbalances arise because assessment for and of learning are too weakly interconnected (Earl, 2003:68). One of the frequently observed side-effects is low-energy feedback: the scales that have been tipped towards assessment of learning, with too much emphasis on procedures for end-of-course grading and certifying of learners' performance, and a concomitant undervaluing of, or underinvestment in, feedback on the

part of staff (Earl, 2003:68-69). There may well be some opportunities for learners to gain practice in tasks on which they will later be formally assessed, but feedback is too sparse, too low in nutrients or comes too late, to benefit the quality of their learning significantly (Yorke, 2001:118; Higgins *et al*., 2002:59; Gibbs, 2003:31).

Imbalances can also occur where assessment ***for*** and ***of*** learning have been too tightly, rather than too loosely, interwoven (Gibbs, 2003:32-33). This too can have unwelcome, if unintended, consequences. One is a fractionalization of assessment, which arises from the conviction that learners will neglect or put too little effort into assigned work unless it carries marks. Over time, as marks get attached to an ever-widening pool of study activities, the weighting of any one task becomes smaller and smaller. Since everything seems to count, everything matters a little, but little matters a lot. Staff can find themselves with an unmanageable marking load, administrators have to run systems that count innumerable piles of small change, and learners may feel pressed to turn out and turn in the latest of their set work requirements, rather than necessarily doing it well or trying to learn from it (Earl, 2004:24).

A second unintended consequence of over-tautness between assessment *for* and *of* learning can be premature testing, where learners are graded on relatively unfamiliar tasks before they have had an opportunity to gain sufficient practice and confidence in doing them (Hounsell *et al*., 2006). This may occur particularly in modularized and semesterized curricula where course units run over a shorter span of weeks than hitherto, with the consequent risk of shrinking opportunities for learners to try their hand at an unfamiliar task, learn from feedback on it and practise it afresh, before being formally assessed on it (Gibbs, 2006:23). It can also arise where assumptions are made about what kinds of assignments learners will have experienced and learnt to do well at in previous course units, in a situation where the sheer range of course combinations which learners can opt for would make any such assumption precarious.

If assessment for learning, as described above, becomes standard practice only in classrooms of low-achieving, low socio-economic status learners, the achievement gaps that trouble us so deeply today would be erased. There is

no other school improvement innovation that can claim effects of this nature or size.

Assessment for learning involves learners in the learning process. These principles of assessment *for* learning (Stiggins *et al*., 2004:40; Chappell, 2004:22) that aim to enhance competency help testing look more like teaching. Some examples of what educators do when they apply those principles involving learners in the assessment process include the following aspects (Stiggins *et al.*, 2004:40; Chappell, 2004:22):

- Keep learners connected to a vision of quality as the learning unfolds, continually defining the learning expectations for learners.
- Use daily strategies in the classroom that require learners to think about their own progress and communicate their own understanding of what they have learned, and set goals to close the gap between where they are now relative to the intended learning and where they need to be in order to meet standards.
- Provide learners with descriptive feedback linked directly to the intended learning, giving them insight about current strengths and on how to do better next time, rather than evaluative feedback consisting only of marks and letter grades.
- Engage learners in activities that teach the skills of self-assessment, helping them collect evidence of their own progress.
- Gather accurate information about a learner's achievement on a regular basis in the classroom using high-quality, accurate assessments *for* learning (Stiggins *et al*., 2004:41; Chappell, 2004:22).

In this study, the above highlighted concept will be investigated with regard to how well assessment of and assessment for learning are balanced in the CTA tasks. This aspect was addressed in the educator questionnaire (*cf*. Appendix H) in Section B1O, B12, B15 and B19; and Section C29, C30 and C35.

2.4 SUMMARY

This chapter presented a perspective on quality management in the design of CTA, aiming at developing a common understanding of quality within this context by, among others, defining the term and its relevant concepts (*cf*. 2.2.1-2.2.3).

Achieving quality in management implies (1) familiarizing stakeholders with the process, (2) developing and training staff members, (3) evaluating performances through inspection, (4) managing the change process and (5) side-stepping a top-down approach (*cf*. 2.2.3.1).

As part of quality management in the design of CTA, the researcher of this thesis looked at quality in assessment (*cf*. 2.2.4) in terms of the following sub-categories: validity was named and discussed first as one of the most important aspects of sound assessment, including content, construct, concurrent, face, criterion-related, and consequential *validity* (*cf*. 2.2.4.1). Secondly, reliability as implying consistency in assessments was discussed by focusing on the relationship between reliability and validity (*cf*. 2.2.4.2.1). Thirdly, the researcher looked at authenticity in assessment (*cf.* 2.2.4.3), presenting a general framework (*cf*. Figure 2.1) and a five dimensional model for authentic instruction (*cf*. Figure 2.2) in order to present the reader with visual viewpoints. The latter comprised the task/physical context/social context/assessment result/criteria (*cf*. Figure 2.2; 2.2.4.3).

The reader was made aware of the need for educators to deal with *predictive validity*, the educational level of learners and subjectivity of authenticity when trying to develop authentic assessment (*cf*. 2.2.4.3).

In this chapter, the researcher addressed several other key features of quality assessment that comprised the following: flexibility which underscores feedback regarding learners' achievements (*cf*. 2.2.4.4); expanded opportunity in assessment so that learners may be assessed in different ways (*cf*. 2.2.4.5); assessment as a continuous process that refers to the ongoing monitoring of learners (*cf*. 2.2.4.6); openness, transparency and accountability that imply making the expectations clear to the learners (*cf*. 2.2.4.7); equity that concerns itself with being free from bias (*cf*. 2.2.4.8); fairness that

concerns itself with taking note of inequalities regarding opportunities, resources and teaching approaches (*cf*. 2.2.4.9); transferability and generalizability that point to learners being able to transfer classroom skills to assessment situations (*cf*. 2.2.4.10); cognitive complexity as being grounded in Bloom's Taxonomy (*cf*. 2.2.4.11.1); meaningfulness in assessment that concerns itself with worthwhile educational experiences (*cf*. 2.2.4.13); cost efficiency in assessment (*cf*. 2.2.4.14); and assessment *of* learning versus assessment *for* learning as it mediates between the learning needs in order to balance assessment (*cf*. 2.2.15).

The next chapter, Chapter Three, will focus on presenting an overview of managing the quality of the implementation of CTA.